

N° d'ordre NNT : 2017LYSE1044

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale ED12
InfoMaths

Spécialité de doctorat : Informatique

Soutenue publiquement le 09 Mars 2017, par :
Huu-Nghi NGUYEN

Estimation de l'écart type du délai de bout-en-bout par méthodes passives

---

Devant le jury composé de :

Madame Tülin BERBER ATMACA, Professeur des universités, Télécom SudParis	Rapporteur
Monsieur Steven MARTIN, Professeur des universités, Université Paris-Sud	Rapporteur
Madame Katia JAFFRES-RUNSER, Maître de Conférences, Université de Toulouse	Examinatrice
Monsieur Congduc PHAM, Professeur des universités, Université de Pau	Examinateur
Madame Isabelle GUÉRIN LASSOUS, Professeur des universités, Université Lyon 1	Directrice de thèse
Monsieur Anthony BUSSON, Professeur des universités, Université Lyon 1	Co-encadrant de thèse
Monsieur Thomas BEGIN, Maître de Conférences, Université Lyon 1	Co-encadrant de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directeur Général des Services	M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur J. ETIENNE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. le Professeur Y. MATILLON
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

Table des matières

1 Introduction Générale **1**
- 1.1 Des réseaux de plus en plus complexes . . . 2
- 1.2 La solution SDN . . . 3
- 1.3 Résumé des contributions . . . 4
- 1.4 Structure du document . . . 4

2 Le délai de bout-en-bout et sa mesure **5**
- 2.1 Définitions . . . 6
  - 2.1.1 Différents types de délais . . . 6
  - 2.1.2 Le délai de bout-en-bout . . . 6
  - 2.1.3 Influence de la charge . . . 7
  - 2.1.4 Une autre mesure de délai : le RTT . . . 8
- 2.2 Caractérisation . . . 8
  - 2.2.1 Une quantité variable . . . 8
  - 2.2.2 Une description statistique . . . 8
  - 2.2.3 Description partielle : les deux premiers moments . . . 9
- 2.3 État de l'art . . . 10
  - 2.3.1 Le système GPS . . . 10
  - 2.3.2 Le protocole NTP . . . 10
  - 2.3.3 La mesure directe . . . 11
  - 2.3.4 Les mesures locales (sur un nœud) . . . 12
  - 2.3.5 Aller de mesures locales vers une mesure globale du délai . . . 12
- 2.4 Des délais corrélés sur les nœuds d'un chemin . . . 13
- 2.5 Conclusions . . . 14

3 Méthode par décomposition **17**
- 3.1 Introduction . . . 19
- 3.2 Scénario, hypothèses et notations . . . 19
- 3.3 L'espérance du délai de bout-en-bout . . . 23
- 3.4 Méthode triviale pour l'estimation de la variance du délai de bout-en-bout 24
- 3.5 Méthode par récurrence . . . 24
- 3.6 Méthode par décomposition . . . 26
  - 3.6.1 Décomposition de la variance . . . 26
  - 3.6.2 Description de la méthode par décomposition . . . 27
- 3.7 Premiers résultats numériques . . . 28
- 3.8 Conclusion . . . 30

4 Méthode utilisant l'espérance conditionnelle (méthode conditionnelle) 31
4.1 Introduction . . . 33
4.2 Problématique . . . 33
4.3 Méthode conditionnelle . . . 35
4.3.1 Bref rappel sur les espérances conditionnelles . . . 35
4.3.2 Scénarios et notations . . . 35
4.4 Méthode conditionnelle . . . 37
4.4.1 Sans flux secondaire f_2 . . . 37
4.4.2 Avec flux secondaire f_2 . . . 38
4.4.3 Cas général : heuristique . . . 38
4.5 Premiers résultats numériques . . . 39
4.6 Conclusion . . . 41

5 Évaluation et Discussion 43
5.1 Introduction . . . 45
5.2 Configuration . . . 45
5.2.1 Simulateur NS-3 . . . 45
5.2.2 Scénario . . . 46
5.3 Trace réseau réaliste . . . 46
5.3.1 Taille des paquets . . . 46
5.3.2 Arrivée des paquets . . . 47
5.3.3 Sélection des flux . . . 48
5.4 Impact du nombre des paquets . . . 49
5.4.1 Un très grand nombre de paquets observés . . . 49
5.4.2 Un petit nombre de paquets observés . . . 49
5.4.3 Nombre variable de paquets observés . . . 53
5.5 Impact de la taille du tampon . . . 55
5.5.1 500 paquets observés . . . 56
5.5.2 1000 paquets observés . . . 57
5.6 Impact de l'intervalle de temps . . . 59
5.6.1 Trace de l'Université de Stuttgart . . . 59
5.6.2 Trace CAIDA . . . 60
5.7 Conclusions . . . 60

6 Conclusions et perspectives 63
6.1 Résumé des contributions . . . 64
6.2 Travaux futurs . . . 65

Résumé : Depuis l'avènement du réseau Internet, le volume de données échangées sur les réseaux a crû de manière exponentielle. Le matériel présent sur les réseaux est devenu très hétérogène, dû entre autres à la multiplication des "middleboxes" (parefeux, routeurs NAT, serveurs VPN, proxy, etc.). Les algorithmes exécutés sur les équipements réseaux (routage, "spanning tree", etc.) sont souvent complexes, parfois fermés et propriétaires et les interfaces de supervision peuvent être très différentes d'un constructeur/équipement à un autre. Ces différents facteurs rendent la compréhension et le fonctionnement du réseau complexe.

Cela a motivé la définition d'un nouveau paradigme réseaux afin de simplifier la conception et la gestion des réseaux : le SDN ("Software-defined Networking"). Il introduit la notion de contrôleur, qui est un équipement qui a pour rôle de contrôler les équipements du plan de données. Le concept SDN sépare donc le plan de données chargés de l'acheminement des paquets, qui est opéré par des équipements nommés virtual switches dans la terminologie SDN, et le plan contrôle, en charge de toutes les décisions, et qui est donc effectué par le contrôleur SDN. Pour permettre au contrôleur de prendre ses décisions, il doit disposer d'une vue globale du réseau. En plus de la topologie et de la capacité des liens, des critères de performances comme le délai, le taux de pertes, la bande passante disponible, peuvent être pris en compte. Cette connaissance peut permettre par exemple un routage multi-classes, ou/et garantir des niveaux de qualité de service.

Les contributions de cette thèse portent sur la proposition d'algorithmes permettant à une entité centralisée, et en particulier à un contrôleur dans un cadre SDN, d'obtenir des estimations fiables du délai de bout-en-bout pour les flux traversant le réseau. Les méthodes proposées sont passives, c'est-à-dire qu'elles ne génèrent aucun trafic supplémentaire. Nous nous intéressons tout particulièrement à la moyenne et l'écart type du délai. Il apparaît que le premier moment peut être obtenu assez facilement. Au contraire, la corrélation qui apparaît dans les temps d'attentes des nœuds du réseau rend l'estimation de l'écart type beaucoup plus complexe. Nous montrons que les méthodes développées sont capables de capturer les corrélations des délais dans les différents nœuds et d'offrir des estimations précises de l'écart type. Ces résultats sont validés par simulations où nous considérons un large éventail de scénarios permettant de valider nos algorithmes dans différents contextes d'utilisation.

Abstract : Since the early beginning of Internet, the amount of data exchanged over the networks has exponentially grown. The devices deployed on the networks are very heterogeneous, because of the growing presence of middleboxes (e.g., firewalls, NAT routers, VPN servers, proxy). The algorithms run on the networking devices (e.g., routing, spanning tree) are often complex, closed, and proprietary while the interfaces to access these devices typically vary from one manufacturer to the other. All these factors tend to hinder the understanding and the management of networks.

Therefore a new paradigm has been introduced to ease the design and the management of networks, namely, the SDN (Software-defined Networking). In particular, SDN defines a new entity, the controller that is in charge of controlling the devices belonging to the data plane. Thus, in a SDN-network, the data plane, which is handled by networking devices called virtual switches, and the control plane, which takes the decisions and executed by the controller, are separated. In order to let the controller take its decisions, it must have a global view on the network. This includes the topology of the network and its links capacity, along with other possible performance metrics such delays, loss rates, and available bandwidths. This knowledge can enable a multi-class routing, or help guarantee levels of Quality of Service.

The contributions of this thesis are new algorithms that allow a centralized entity, such as the controller in an SDN network, to accurately estimate the end-to-end delay for a given flow in its network. The proposed methods are passive in the sense that they do not require any additional traffic to be run. More precisely, we study the expectation and the standard deviation of the delay. We show how the first moment can be easily computed. On the other hand, estimating the standard deviation is much more complex because of the correlations existing between the different waiting times. We show that the proposed methods are able to capture these correlations between delays and thus providing accurate estimations of the standard deviation of the end-to-end delay. Simulations that cover a large range of possible scenarios validate these results.

Remerciements

Je souhaite remercier Mme. Tülin Berber Atmaca, Professeur à Télécom SudParis et M. Steven Martin, Professeur à l'Université Paris-Sud d'avoir accepté d'être rapporteurs de ma thèse. Leurs suggestions m'ont permis d'améliorer le manuscrit de ma thèse Je souhaite également à remercier Mme. Katia Jaffres-Runser, Maître de Conférences à l'Université de Toulouse et M. Congduc Pham, Professeur à l'Université de Pau d'avoir accepté de juger ma thèse.

Je souhaite remercier Mme. Isabelle Guérin-Lassous, Professeur à l'Université Lyon 1, M. Anthony Busson, Professeur à l'Université Lyon 1 et M. Thomas Begin, Maître de Conférences à l'Université Lyon 1 d'avoir dirigé ma thèse. Ils m'ont enseigné non seulement des méthodologies pour faire de la recherche, mais également à rédiger des rapports, à être précis dans les terminologies utilisées, et à bien comprendre les notions fondamentales des réseaux. Je pense que cette thèse n'aurait pas vu le jour sans leur sympathie et patience.

Je souhaite remercier M. Paulo Gonçalves pour m'avoir donné de précieux conseils. Ses conseils touchant aussi bien la culture que le plaisir de faire la recherche m'ont permis de positionner mes objectifs pour le futur. Je remercie aussi les professeurs de l'Institut des Systèmes Complexes IXXI, pour leurs conseils et discussions précieuses.

Je souhaite remercier l'ensemble des secrétaires au laboratoire LIP de l'ENSL pour avoir, entre autres, préparé les dossiers et les papiers administratifs me concernant. J'ai une pensée toute particulière pour Mme. Laetitia Lecot et ses explications concrètes et claires sur des procédures complexes.

Je tiens aussi à saluer les doctorants, Mohammed, Sarrah, Jazmin, Jordan, Esteban, Marija, qui ont partagé mon bureau dans la salle 223 à l'IXXI, ainsi que les autres thésards de l'ENSL.

Enfin, je souhaite remercier mes parents et mes deux grandes sœurs. Leurs encouragements m'ont aidé à surmonter les difficultés rencontrées pendant les trois ans de la thèse.

Chapitre 1

Introduction Générale

Contents

1.1 Des réseaux de plus en plus complexes 2
1.2 La solution SDN 3
1.3 Résumé des contributions 4
1.4 Structure du document 4

1.1 Des réseaux de plus en plus complexes

Un trafic grandissant et sous contraintes de performances

Les réseaux informatiques sont utilisés pour échanger, transférer, et communiquer des informations depuis les serveurs de contenu vers les utilisateurs ou bien directement entre utilisateurs. Aujourd'hui, dans nos sociétés numériques, une part grandissante des échanges d'informations se fait par les réseaux (au détriment des canaux traditionnels). Ainsi, l'email remplace progressivement le courier postal, l'e-commerce et l'e-paiement se démocratisent, les vidéos en ligne (par exemple youTube) prennent du terrain sur la télévision, ... En conséquence, le volume de données échangées sur les réseaux croît de manière exponentielle.
À sa création, Internet (ou son ancêtre ARPANET) avait été pensé pour constituer un réseau de taille modeste (quelques centaines de nœuds) qui permettrait d'exécuter des applications distribuées élastiques (c'est-à-dire sans contraintes particulières sur les délais d'acheminement des paquets). Or en quelques dizaines d'années, Internet s'est étendu de manière prodigieuse et il héberge à présent une vaste gamme d'applications comprenant des applications élastiques (web, email, ftp) mais également des applications non élastiques qui nécessitent un acheminement rapide et fiable de leurs données (voix sur IP, visioconférence, vidéos à la demande). Ces dernières applications, bien que plus récentes, sont généralement très consommatrices en resource réseaux et représentent une part grandissante du trafic Internet. En 2013 Netflix seul représentait environ 33% du trafic aux heures de pointes sur les réseaux opérateurs nord-américains [40].
Pour accompagner cette croissance exponentielle du trafic Internet, les opérateurs ont renforcé leurs réseaux en augmentant leur capacité. Ainsi la capacité des liens réseaux est passée de quelques Mbps à plusieurs dizaines, voire centaines, de Gbps grâce au progrès technologiques (comme par exemple, le déploiement de la fibre optique).

Une gestion des ressources rendue complexe

Toutefois, les administrateurs réseaux se sont heurtés et se heurtent encore à certains obstacles.
Premièrement, le matériel présent sur les réseaux est très hétérogène. En effet, en plus des routeurs et des commutateurs, on trouve à présent une multitude d'autres équipements. Par exemple, pour sécuriser leurs réseaux et leurs communications entre sites distants, ils font appels à des pare-feux ("firewall") et à des tunnels VPN. Pour palier la pénurie d'adresses IPv4 disponibles (environ 4 milliards), ils déploient des mécanismes de NAT (Network Address Translation) aux interfaces de leur réseau. Pour contrôler les communications de leurs utilisateurs, ils peuvent mettre en place un serveur Proxy qui se substitue aux postes clients pour faire les requêtes aux serveurs distants. Pour améliorer certains problèmes de performances, ils déploient des équipements permettant l'équilibrage de charge sur plusieurs nœuds répliquant le même contenu, ainsi que des optimiseurs WAN permettant de mieux s'accommoder de liens longues distances (long délai de propagation). Tous ces équipements réseaux (pare-feux, tunnels VPN, NAT, équilibrage de charges, optimiseurs WAN, ...), dont le rôle principal n'est pas l'acheminement des paquets, sont généralement désignés sous l'appellation commune "middlebox". Bien qu'omniprésents et essentiels au fonctionnement d'un réseau, ces middleboxes participent grandement à l'accroissement de l'hétérogénéité du matériel.
Deuxièmement, le code exécuté sur les équipements réseaux (comme par exemple, le protocole de routage sur les routeurs, le "spanning tree" sur les commutateurs et autres algorithmes sur les middleboxes) est souvent complexe, fermé et propriétaire. Il devient

difficile alors pour les administrateurs réseaux de comprendre précisément le fonctionnement de ses équipements et a fortiori de son réseau. Par exemple, les protocoles de routage reposent généralement sur des algorithmes distribués mettant en jeu des centaines de routeurs. Les routeurs sont vendus par les constructeurs avec une implémentation particulière de l'algorithme, pas toujours entièrement compatible avec les implémentations des autres constructeurs (certains paramètres pouvant être calibrés différemment). Ces incompatibilités peuvent être sources de bien des difficultés et d'erreurs pour les administrateurs réseaux.
Troisièmement, les interfaces par lesquelles les administrateurs réseaux peuvent prendre le contrôle pour configurer les équipements réseaux diffèrent souvent d'un constructeur à un autre (voire d'un modèle à un autre).
Ces trois facteurs que sont l'hétérogénéité du matériel, la complexité des codes et l'hétérogénéité des interfaces de contrôle participent à rendre la gestion d'un réseau de plus en plus complexe. Cette complexité d'administration implique également une augmentation des coûts d'acquisition et d'exploitation d'un réseau, et représente également un frein à l'innovation technologique.

1.2 La solution SDN

Une gestion unifiée et centralisée des resources

Au début des années 2000, la communauté scientifique réseaux a défini un nouveau paradigme afin de simplifier la conception et la gestion des réseaux. Ce nouveau paradigme est connu sous le nom de SDN : "Software-defined Networking" [13]. Le SDN redéfinit profondément le fonctionnement des réseaux IP traditionnels et comporte principalement deux caractéristiques.
Premièrement, contrairement aux réseaux informatiques actuels, le SDN dissocie le **plan de contrôle** et le **plan de données**. Le plan de contrôle a pour mission de décider comment traiter le trafic, c'est-à-dire principalement de définir les politiques d'acheminement pour les paquets de données. Le plan de données s'occupe uniquement de l'acheminement du trafic, en appliquant les décisions prises par le plan de contrôle. Ainsi, le plan de contrôle décide, et le plan de données exécute et met en œuvre.
Deuxièmement, le SDN introduit une entité nouvelle, appelée le **contrôleur SDN**, qui a pour rôle de contrôler les nombreux équipements du plan de données (c'est-à-dire les commutateurs de paquets [1]). Habituellement ces commutateurs de paquets sont appelés **commutateurs virtuels** ("virtual switches"). Notons que ces commutateurs virtuels peuvent exécuter à la fois des opérations d'acheminement (à l'instar des routeurs et des commutateurs), de filtrage (à l'instar des pare-feux), de translation d'adresses (à l'instar des NAT), ... Ainsi, ces commutateurs virtuels sont polyvalents et devraient être l'unique équipement retrouvé dans les réseaux (mettant un terme à l'expansion des middleboxes). Enfin, les décisions prises par le contrôleur SDN sont communiquées aux commutateurs virtuels en utilisant un protocole tel qu'OpenFlow.

Le rôle crucial du contrôleur SDN

Typiquement, le contrôleur SDN est un programme exécuté sur un ou plusieurs nœuds du réseau. C'est lui qui calculera les routes à prendre par les paquets entrants dans le réseau et devra ensuite les communiquer à l'ensemble des commutateurs virtuels. Le contrôleur

1. Notons que ces commutateurs de paquets ne sont plus à proprement parler des routeurs car ils ne construisent pas eux-mêmes leur table de routage mais reçoivent leurs instructions du contrôleur.

SDN a bien entendu un rôle central pour le bon fonctionnement du réseau. Ses décisions ont un impact direct sur les performances ressenties des utilisateurs. Il est donc crucial que ses décisions soient prises au mieux.

Un tableau de bord des performances

Pour permettre au contrôleur de prendre ses décisions, ce dernier doit disposer d'une vue globale du réseau. A minima, le contrôleur doit en connaître la topologie, c'est-à-dire le graphe du réseau. Le contrôleur peut alors appliquer un algorithme de plus court chemin (tel que Dijkstra) pour déterminer le routage des paquets sur le réseau. Toutefois, il pourrait mieux faire si, en plus de la topologie, il dispose des capacités de transmission et des délais de propagation sur chaque lien. Enfin, s'il souhaite mettre en place un routage plus avancé, comme par exemple le routage multi-classes sous contraintes de performances, il est alors nécessaire d'enrichir sa connaissance du réseau en lui indiquant des métriques de performances telles que le taux d'utilisation sur chaque lien, la bande passante résiduelle sur chaque lien, le taux de perte sur chaque lien, les délais d'attente dans les tampons des commutateurs virtuels ou les délais de bout-en-bout pour chaque chemin.
Ainsi, le contrôleur SDN, et a fortiori le réseau tout entier, bénéficient largement d'une connaissance étendue du réseau et de ses performances. Le contrôleur aurait alors à sa disposition, en plus de la topologie du réseau et des capacités des liens, un ensemble de métriques de performances qu'il pourrait consulter avant d'exécuter ses algorithmes et de prendre ses décisions. La bande passante résiduelle (ou de manière équivalente le taux d'utilisation) d'un lien fait partie naturellement de ces métriques. Nous pensons qu'étant donné le nombre très important d'applications temps réel (voix sur IP, visioconférences, vidéos à la demande), le **délai de bout-en-bout**, qui est l'objet de ce manuscrit, est également un candidat naturel pour apparaître dans cette liste de métriques de performances.

1.3 Résumé des contributions

Ce manuscrit comprend deux principales contributions portant sur la faisabilité d'inclure le délai de bout-en-bout dans l'ensemble des métriques de performances mis à disposition du contrôleur SDN. D'abord, nous traitons rapidement le cas (simple) du délai moyen de bout-en-bout. Ensuite, nous présentons deux approches possibles pour estimer efficacement et avec précision l'écart type du délai de bout-en-bout.

1.4 Structure du document

La suite de ce document est organisée de la façon suivante. Dans le Chapitre 2, nous définissons précisément le délai de bout-en-bout et nous expliquons la difficulté à l'estimer mais également l'intérêt de le caractériser par sa moyenne et son écart type. Le Chapitre 3 propose une première solution pour estimer l'écart type du délai de bout-en-bout d'un chemin réseau. Nous présentons une autre solution, basée sur une approche différente, dans le Chapitre 4. Le Chapitre 5 présente la précision obtenue par chacun de ces deux solutions, et nous permet de connaître la meilleure. Le Chapitre 6 clôt ce manuscript.

Chapitre 2

Le délai de bout-en-bout et sa mesure

Contents

2.1 Définitions 6
2.1.1 Différents types de délais 6
2.1.2 Le délai de bout-en-bout 6
2.1.3 Influence de la charge 7
2.1.4 Une autre mesure de délai : le RTT 8
2.2 Caractérisation 8
2.2.1 Une quantité variable 8
2.2.2 Une description statistique 8
2.2.3 Description partielle : les deux premiers moments 9
2.3 État de l'art 10
2.3.1 Le système GPS 10
2.3.2 Le protocole NTP 10
2.3.3 La mesure directe 11
2.3.4 Les mesures locales (sur un nœud) 12
2.3.5 Aller de mesures locales vers une mesure globale du délai 12
2.4 Des délais corrélés sur les nœuds d'un chemin 13
2.5 Conclusions 14

2.1 Définitions

2.1.1 Différents types de délais

Localement, sur chaque nœud emprunté par un paquet, on distingue généralement quatre types du délai comme illustré sur la Figure 2.1 :

- **Le délai de traitement** qui est le temps passé par le nœud à examiner et éditer le paquet (changement de l'en-tête, interrogation de la table de routage, copie en mémoire, commutation sur les bus,. . .). Le délai de traitement est typiquement très faible, inférieur à quelques micro-secondes et peut donc être négligé dans le calcul du délai total d'un paquet. À noter que dans les réseaux de nouvelle génération où des équipements logiciels remplaceront les équipements matériels comme le préconise l'approche NFV (Network Function Virtualization) [18, 25], le temps de traitement des paquets devrait rester très faible (comparé aux autres types de délais). En effet, un équipement logiciel, une fois équipé d'une solution de switch logiciel comme Open vSwitch [32] accélérée par une librairie type DPDK [37], Netmap [38], OpenOnload [33] ou PacketShader [34], peut traiter un paquet en généralement quelques dizaines (voir quelques centaines) de cycles CPU [11], ce qui, avec des processeurs cadencés à 2 ou 3 GHz (temps de cycle inférieur à la nanoseconde,) représente un délai de traitement bien inférieur à la microseconde. Pour plus d'informations, le lecteur intéressé peut se référer à [7].
- **Le délai d'attente** qui correspond au temps que le paquet doit attendre dans un tampon du nœud avant de pouvoir entamer son émission. Ce temps correspond généralement à la transmission d'autres paquets précédemment arrivés sur le même nœud. Sa valeur est très difficile à prévoir car elle dépend du nombre de paquets déjà en attente lors de l'arrivée du paquet sur le nœud.
- **Le délai d'émission** qui correspond au temps que le nœud consacre à transmettre le paquet sur le lien. Sa valeur dépend de la taille du paquet et de la capacité de transmission (bande passante) du lien.
- **Le délai de propagation** qui correspond au temps pour le signal transportant le paquet d'atteindre le nœud suivant. Sa valeur dépend de la longueur du lien à parcourir et de la célérité du signal.

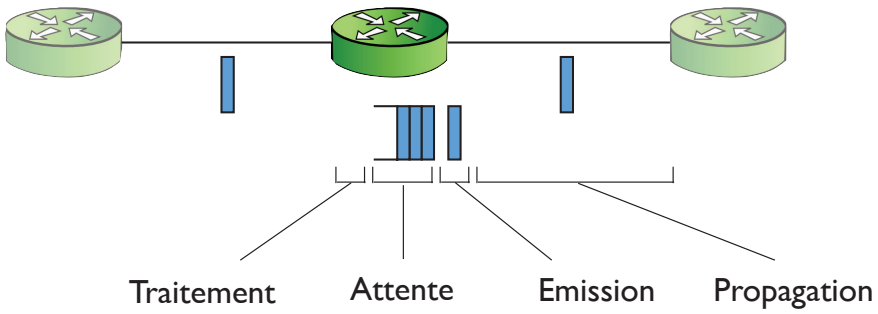


FIGURE 2.1 – Types de délais sur un nœud.

2.1.2 Le délai de bout-en-bout

Pour un chemin réseau qui permet l'acheminement de paquets entre un nœud expéditeur et un nœud destinataire, on définit la notion de délai de bout-en-bout. Comme discuté dans le chapitre précédent, le délai de bout-en-bout constitue avec le débit, le taux de perte et le taux d'utilisation une des principales métriques de performances en réseau.

Nous définissons le **délai de bout-en-bout** d'un paquet comme étant le temps qui s'écoule entre l'instant où le paquet est mis à disposition de la couche Réseau sur le nœud expéditeur et l'instant où le paquet est reçu par la couche Réseau sur le nœud destinataire. Par conséquent le délai de bout-en-bout d'un paquet comprend les temps de traitement, d'émission, de propagation et éventuellement d'attente sur le nœud expéditeur ainsi que sur l'ensemble des nœuds que le paquet devra traverser pour atteindre le nœud destinataire. La Figure 2.2 illustre la définition du délai de bout-en-bout pour un chemin composé de quatre nœuds. Sur cet exemple, il est à noter qu'un seul délai d'attente intervient (sur le deuxième nœud) et que nous avons négligé les délais de traitement.

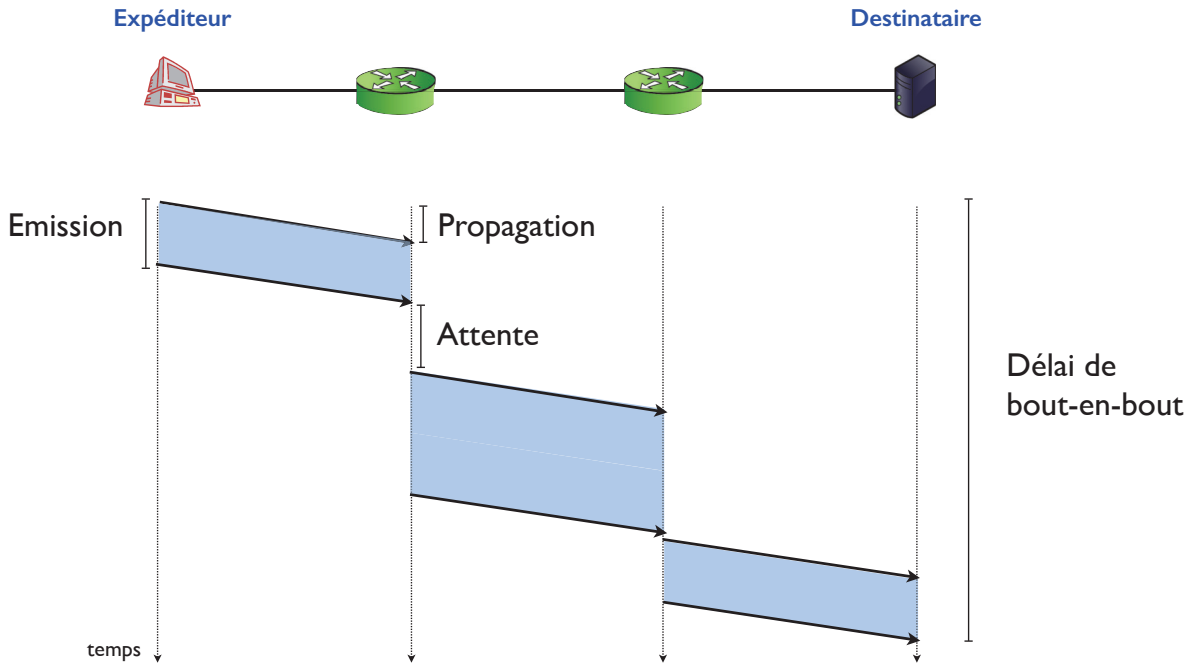


FIGURE 2.2 – Délai de bout-en-bout sur un chemin composé de 4 nœuds.

2.1.3 Influence de la charge

La valeur du délai de bout-en-bout pour un paquet dépend bien entendu du chemin emprunté (nœuds traversés) mais également du niveau de charge. Pour un chemin et un paquet donnés, les délais d'émission, de propagation et de traitement sur chaque nœud peuvent être prévus. En revanche, il en va différemment pour les délais d'attente. Les délais d'attente peuvent être nuls (par exemple si aucun autre paquet ne circule au même instant sur le même lien) ou très important (si les tampons des nœuds contiennent de nombreux paquets également en attente de l'accès aux liens). Ainsi, plus le niveau de charge sur le chemin réseau sera important, plus le délai de bout-en-bout sera grand. Toutefois, son élongation avec le niveau de charge n'est pas simple (et non linéaire). Notons que l'estimation des délais d'attente est particulièrement compliquée lorsque le niveau de charge est modéré ce qui correspond à des tampons pouvant être à tour de rôle presque vides ou presque pleins.
Les opérateurs réseaux surdimensionnent leur infrastructure en ayant des liens avec des capacités de transmissions souvent bien supérieures aux demandes de trafic des usagers. De par cette pratique des opérateurs, le niveau de charge sur les liens réseaux d'un opérateur est généralement faible ou modéré [6, 14, 31, 15, 17], avec un taux d'utilisation des liens ne dépassant pas 50%.

2.1.4 Une autre mesure de délai : le RTT

Il existe une autre mesure de délai commune pour un chemin réseau. Le RTT ("Round Trip Time") exprime le temps que mettrait un paquet partant du nœud expéditeur pour aller jusqu'au destinataire puis revenir à l'expéditeur (éventuellement en empruntant un chemin différent à l'aller et au retour). L'estimation de la valeur du RTT est essentielle au mécanisme de retransmissions des paquets par TCP. En effet la valeur du RTT permet de calibrer la durée au bout de laquelle un paquet non acquitté est considéré comme perdu, et donc nécessite une retransmission par TCP. Toutefois, contrairement au délai de bout-en-bout, cette mesure n'est pas très utile pour des applications avec des contraintes fortes en QoS. Notons que la mesure du RTT est généralement simple car elle s'effectue entièrement sur le nœud expéditeur. Une seule horloge suffit et il n'y a donc pas de problème de synchronisation des horloges. Notons qu'il n'est pas possible de déduire la valeur du délai de bout-en-bout d'un paquet à partir de la connaissance de son RTT (ne serait-ce parce que le chemin emprunté sur la voie retour peut être différent).

2.2 Caractérisation

2.2.1 Une quantité variable

Considérons un chemin réseau donné (déterminé par le(s) protocole(s) de routage) entre deux nœuds du réseau. Chaque paquet empruntant ce chemin réseau aura un délai de bout-en-bout différent, notamment car les conditions de charge sur les nœuds intermédiaires varient (l'occupation des tampons). Ainsi, il existe une multitude, voire une infinité de valeurs possibles pour le délai de bout-en-bout d'un chemin. Toutefois, certaines valeurs sont plus probables que d'autres.

2.2.2 Une description statistique

Si l'on souhaite caractériser cette variabilité du délai de bout-en-bout, on peut considérer le délai de bout-en-bout d'un chemin réseau comme étant une variable aléatoire. Chaque paquet traversant le chemin produit un échantillon. Pour décrire une variable aléatoire, on fait couramment appel à sa **distribution de probabilité** (également appelée loi de probabilité) qui décrit l'ensemble des valeurs possibles ainsi que leur fréquence d'occurrence (ou probabilité). Toutefois, en pratique, on caractérise souvent, et de façon partielle, une variable aléatoire par ses premiers moments. Le premier moment, plus connu sous le nom de **moyenne** ou d'**espérance**, indique la tendance centrale de la variable aléatoire. Le deuxième moment centré, plus connu sous le nom de **variance**, indique le degré auquel la variable aléatoire se disperse autour de son espérance. Une variance élevée indique que les valeurs sont très écartées les unes des autres. À l'inverse, une variance est nulle lorsque toutes les valeurs pouvant être prises par la variable sont identiques. Il s'agit alors d'une constante. La variance permet d'obtenir l'écart type, qui est la racine carrée de la variance. L'**écart type** est souvent plus simple à interpréter car ses valeurs sont exprimées dans les mêmes unités que la quantité observée (en secondes dans le cas du délai de bout-en-bout). La Figure 2.3 illustre les notions de distribution de probabilité, espérance et écart type pour une distribution normale d'espérance égale 10 et d'écart type égal à 3.
Notons qu'il existe beaucoup d'autres éléments pour caractériser le comportement d'une variable aléatoire tels que sa médiane, ses modes, les moments d'ordre supérieur (comme le skewness et le kurtosis), les quantiles, ...

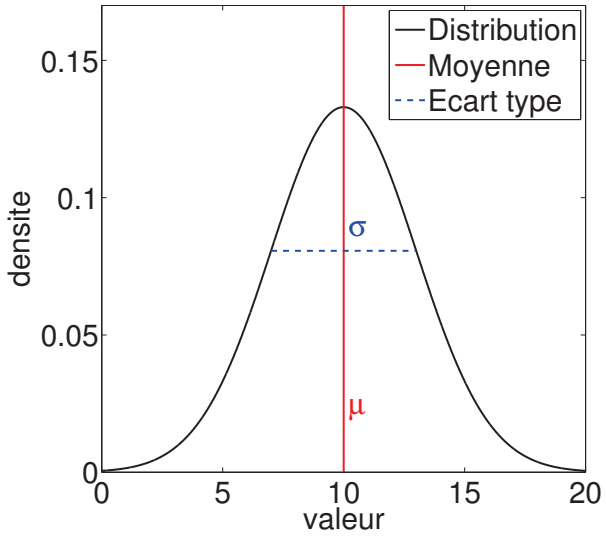


FIGURE 2.3 – Exemple de distribution de probabilité d'une variable aléatoire avec son espérance et son écart type.

2.2.3 Description partielle : les deux premiers moments

Le comportement d'une variable aléatoire est entièrement définie par sa distribution de probabilité. Par ailleurs, on peut également entièrement caractériser une variable aléatoire par la totalité de ses moments [22].
En pratique, une connaissance, bien qu'incomplète, des deux premiers moments d'une variable aléatoire permet généralement d'avoir une idée assez juste de son comportement. Il existe plusieurs façons d'exploiter la connaissance des premiers moments d'une variable aléatoire. Nous en présentons deux.

L'inégalité de Chebyshev

L'inégalité de Chebyshev fournit une borne probabiliste sur l'éloignement des valeurs prises par une variable aléatoire à son espérance. La singularité de ce ce résultat est qu'il nécessite uniquement la connaissance des deux premiers moments, ou de manière équivalente, de l'espérance et de l'écart type. Plus formellement, si X est une variable aléatoire d'espérance μ et d'écart type σ (de valeurs finies tous les deux), alors **l'inégalité de Chebyshev** indique que : Pour tout réel strictement positif α, on a :

$$Pr(|X - \mu| \geq \alpha.\sigma) \leq \frac{1}{\alpha^2}.$$

Autrement dit, pas plus de $\frac{1}{\alpha^2}$ des valeurs prises par la distribution sont éloignées de plus de α écart types à l'espérance. Par exemple, avec $\mu = 10$ et $\sigma = 3$ (et quelle que soit la forme exacte de la distribution de probabilité de X), en choisissant $\alpha = 2$, l'inégalité de Chebyshev indique que moins de 25% des valeurs seront situées à l'extérieur de l'intervalle $[4; 16]$.

L'approche paramétrique

La borne probabiliste fournie par l'inégalité de Chebyshev peut parfois être jugée trop large. Une autre façon alors d'exploiter la connaissance des deux premiers moments d'une variable aléatoire est de supposer une forme de distribution de probabilité pour la variable

aléatoire, et après calibrage de la distribution, d'en dériver des bornes probabilistes plus étroites. Cette approche par modèle paramétrique est très souvent utilisée lorsqu'on a une idée (même approchée), également appelé un "a priori", de la forme de la distribution de probabilité de la variable aléatoire. Par exemple, on peut supposer que la distribution de probabilité suit approximativement une distribution normale (par exemple par résultat de l'application du théorème central limite).
Par exemple, supposons que $\mu = 10$ et $\sigma = 3$ (comme dans l'exemple précédent) et que notre connaissance de la variable aléatoire nous permet également de considérer que la forme de sa distribution suit approximativement une distribution normale (c'est-à-dire avec un unique mode et symétrique autour de son espérance). En utilisant les propriétés de la loi normale paramétrée par $\mu = 10$ et $\sigma = 3$ (et après translation et dilatation pour se ramener à une loi normale centrée réduite), on obtient que moins de 25% des valeurs seront situées à l'extérieur de l'intervalle $[6.55; 13.45]$. Ce résultat est à comparer avec l'intervalle significativement plus large trouvé par l'inégalité de Chebyshev (qui ne faisait aucune hypothèse sur la forme de la distribution de probabilité de la variable aléatoire).

2.3 État de l'art

Après avoir expliqué l'intérêt de connaître les deux premiers moments d'un paramètre de performances comme le délai de bout-en-bout, nous présentons à présent la littérature associée.

2.3.1 Le système GPS

Le système GPS (Global Positioning System) [28], en plus de permettre la localisation d'une entité équipée d'un récepteur GPS, peut aussi permettre la synchronisation de cette entité sur l'horloge GPS. Les satellites GPS sont équipés d'horloges atomiques permettant d'avoir une très bonne précision du temps. Ces temps sont indiqués dans les messages envoyés par les satellites GPS. Ils sont alors utilisés pour localiser un récepteur GPS recevant des messages de plusieurs satellites, mais ils peuvent aussi être utilisés pour synchroniser le temps du récepteur GPS sur l'horloge du système GPS.
Si ce système permet d'avoir une très bonne synchronisation en temps, et est donc indirectement intéressant pour estimer le délai de bout-en-bout dans un réseau [10], il nécessite que les entités à synchroniser soient équipées de récepteur GPS et soient à portée des satellites GPS. Cette dernière caractéristique implique que les entités ne se trouvent pas à l'intérieur de bâtiments, ce qui est une contrainte difficile à respecter pour les réseaux filaires que nous considérons dans cette thèse. Le système GPS est en revanche plus naturel à utiliser dans les réseaux sans fil se trouvant en extérieur, comme dans [12] ou dans les réseaux véhiculaires [35].

2.3.2 Le protocole NTP

En préambule, rappelons qu'à moins d'utiliser du matériel spécialisé (carte GPS), les horloges des nœuds d'un réseau sont non synchronisées. D'une part parce que leur heure de référence diffère, et d'autre part parce que leur horloge dérive à des vitesses différentes. Pour remédier à ces problèmes, il faudrait donc disposer d'un mécanisme initialisant au même instant les horloges de tous les nœuds à une même valeur, et compensant les dérives naturelles de chacune. Le protocole NTP (Network Time Protocol), conçu en 1981 et régulièrement amélioré depuis, apporte justement une réponse à ces deux problèmes. NTP permet la synchronisation (approchée) entre les nœuds participants d'un réseau

grâce à la transmission de paquets NTP émis par un ou plusieurs serveur(s) NTP puis ré-émis à leur tour par certains nœuds [26]. Ainsi, les serveurs NTP ne synchronisent pas tous les nœuds participants. Le protocole NTP forme une hiérarchie entre les serveurs et les nœuds avec différents niveaux. La Figure 2.4 illustre ce principe avec trois niveaux de hiérarchie. Il est généralement accepté que NTP permet une précision de l'ordre de quelques dizaines de millisecondes sur un réseau d'opérateurs mais que les erreurs peuvent atteindre 100 ms ou plus dans des cas défavorables (conditions de charge importantes, routes asymétriques,...) [27]. Il faut aussi noter l'existence du protocole PTP (Precision Time Protocol) [1] qui offre des synchronisations dont l'ordre est inférieure à la micro-seconde [23, 42, 30] mais qui est défini pour des réseaux locaux.

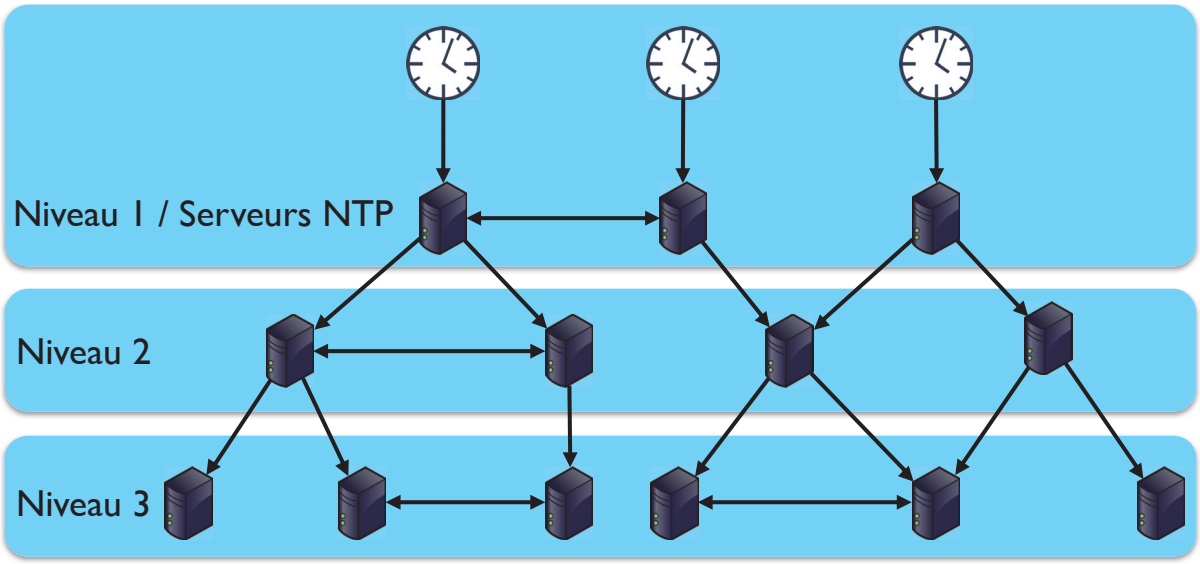


FIGURE 2.4 – Trois serveurs NTP propagent leur horloge à une multitude de nœuds participants.

2.3.3 La mesure directe

Une mesure directe du délai de bout-en-bout consiste ici à calculer les deux premiers moments du délai de bout-en-bout des paquets sur un chemin réseau en utilisant les estimateurs classiques de statistique [4].
Pour cela, il "suffirait" d'émettre des paquets de sonde ("probing packets") sur le chemin réseau considéré et de mesurer pour chacun de ces paquets son délai de bout-en-bout. Cette approche suppose donc des nœuds parfaitement synchronisés (ce qui n'est pas le cas sauf à utiliser du matériel spécialisé et onéreux de type cartes GPS). Une fois les échantillons obtenus, on peut estimer l'espérance et l'écart type de la variable aléatoire considérée (ici le délai de bout-en-bout) en utilisant simplement les estimateurs empiriques classiques.
Bien que conceptuellement simple, la mise en œuvre d'une telle approche dans un réseau IP se heurte à de nombreuses complexités. Nous présentons ci-dessous une multitude de raisons pouvant rendre complexe, voire impossible, une telle approche.

- La difficulté de quantifier l'écart entre l'horloge du nœud expéditeur (où est mesuré l'instant de départ du paquet) et celle du nœud destinataire (où est mesuré l'instant d'arrivée) [5, 29, 43].
- Les délais subis par les paquets NTP eux-mêmes lorsqu'un unique serveur NTP est utilisé pour synchroniser les horloges des nœuds du réseau [29].

– Les phénomènes de dérive d'horloge (chaque horloge dérivant à une vitesse différente) si le réseau repose sur plusieurs serveurs NTP distribués [43].
– Le prix onéreux d'équipements spécialisés tels que les cartes GPS qui permettent de synchroniser les nœuds à une horloge GPS avec une très haute précision [16, 9, 8, 20].
– L'utilisation de paquets de sonde dont la présence peut altérer le comportement du trafic régulier [43].
– L'impact du profil de trafic pour les paquets de sonde (par exemple, la taille des paquets ou leur taux d'émission) sur les résultats obtenus [20].
– La nécessité de différencier les paquets de sonde à leur départ et également de les distinguer à leur arrivée.
– La nécessité d'intervalles de mesures suffisamment longs (par exemple, 10 minutes) pour obtenir des mesures statistiques cohérentes [8].
– L'introduction d'erreurs et d'incertitudes liées à chaque méthode de mesures [5, 16, 43].
– L'impact des méthodes d'échantillonnage sur la qualité des résultats [43].

2.3.4 Les mesures locales (sur un nœud)

En plus des tâches de traitement, de commutation et d'émission, les nœuds d'un réseau collectent des mesures sur leur état. Généralement ces mesures sont des compteurs localisés dans des mémoires registres. Ces mesures permettent de suivre le nombre d'octets ou de paquets qui ont été traités par le nœud pendant un intervalle de temps ce qui correspond donc à une mesure de débit, de connaître la taille des paquets traités et de déterminer le nombre de paquets en attente d'être émis dans le tampon du nœud [36]. Notons que la loi de Little [4] permet d'évaluer le délai moyen d'attente des paquets dans un tampon si l'on connaît le nombre moyen de paquets en attente et le débit moyen auquel les paquets entrent ou quittent le tampon.
Notons que pour ces mesures locales, contrairement au délai de bout-en-bout, il n'existe pas de problèmes liés à la non synchronisation des horloges puisqu'une seule horloge (celle du nœud courant) est utilisée.
Enfin soulignons que si l'on dispose d'échantillons pour ces quantités locales, on peut alors simplement calculer leur espérance et variance (ou de façon équivalente écart type) en utilisant les estimateurs empiriques classiques.

2.3.5 Aller de mesures locales vers une mesure globale du délai

Nous avons vu précédemment pourquoi il est coûteux et difficile, voire impossible, de mesurer directement avec précision le délai de bout-en-bout d'un paquet sur un chemin réseau. À l'inverse, nous avons expliqué pourquoi il est simple et peu coûteux de mesurer localement certaines quantités locales sur les nœuds d'un réseau.
Partant de ce double constat, il apparait naturel de se demander s'il est possible de caractériser le délai de bout-en-bout d'un chemin en utilisant uniquement des connaissances locales aux nœuds traversés. Autrement dit, est-il possible de développer un algorithme qui découvre les deux premiers moments du délai de bout-en-bout d'un chemin sans faire appel à des échantillons de ce délai, mais uniquement à des échantillons collectés sur les nœuds intermédiaires ?
Pour le cas de la valeur moyenne du délai de bout-en-bout d'un chemin réseau, il est très simple de calculer sa valeur à partir des mesures locales de délais sur chacun des nœuds traversés. En effet, le délai de bout-en-bout étant par construction la somme de tous les délais de traitement (négligeables), d'attente, d'émission et de propagation, et par propriété de linéarité de l'opérateur espérance (moyenne), on obtient que :

L'espérance du délai de bout-en-bout est égale à la somme des espérances des délais d'attente, d'émission et de propagation.

Et comme chaque nœud connaît ou est capable d'estimer avec précision ces quantités-là, il est possible de dériver la valeur espérance du délai de bout-en-bout d'un chemin en se basant uniquement sur des quantités mesurées localement sur les nœuds. Malheureusement, pour l'écart type du délai de bout-en-bout, les choses sont plus complexes. Nous expliquons pourquoi dans la section suivante.

2.4 Des délais corrélés sur les nœuds d'un chemin

Nous nous penchons à présent sur l'estimation de l'écart type du délai de bout-en-bout. En supposant que l'ensemble des délais sur les nœuds d'un chemin réseau sont **indépendants**, on peut alors simplement, à l'instar du calcul de l'espérance, estimer l'écart type du délai de bout-en-bout à partir des écart types des délais locaux mesurés sur chaque nœud. En effet, pour des variables aléatoires X et Y indépendantes, on a pour la variance Var :

$$Var(X + Y) = Var(X) + Var(Y).$$

Et donc l'écart type de $X + Y$ s'obtient par :

$$\sigma_{X+Y} = \sqrt{Var(X) + Var(Y)}.$$

Malheureusement bien que certains de ces délais puissent être considérés indépendants entre eux, il en va autrement pour d'autres. Par exemple, nos expériences ont démontré que les délais d'émission et d'attente sur les nœuds sont généralement étroitement corrélés. Ainsi, il suit que :

La variance du délai de bout-en-bout ***ne peut pas*** *être correctement approchée par la somme des variances des délais d'attente, d'émission et de propagation.*

Notons que la variance des délais de propagation peut être considérée nulle (puisque les délais de propagation sont constants quel que soit le paquet) et celle des délais de traitement peut être négligée (car ses valeurs sont très faibles).
Lorsque des variables aléatoires X et Y ne sont pas indépendantes, le calcul de la variance de leur somme implique un terme supplémentaire de covariance qui reflète le lien existant entre (corrélant) ces deux variables. Ainsi, dans le cas général, on a :

$$Var(X + Y) = Var(X) + Var(Y) + 2\ Cov(X, Y).$$

Nous présentons un exemple permettant d'illustrer le besoin de considérer, même approximativement les termes de covariance, lorsqu'on cherche à estimer la variance (ou l'écart type) du délai de bout-en-bout. Pour cela, nous simulons le comportement d'un chemin réseau composé de six nœuds et dans lequel on injecte des paquets en rejouant une trace capturée sur un réseau. L'exemple n'est pas décrit entièrement mais il sera repris dans les chapitres suivants et à ce moment-là l'ensemble du scénario sera indiqué. Nous présentons la valeur de l'écart type obtenu en sommant respectivement les variances des délais d'émission et d'attente mesurées sur chacun des nœuds puis en en prenant la racine carrée. Dans le simulateur, il nous est possible de mesurer exactement les délais de bout-en-bout de chaque paquet, et nous pouvons donc calculer la valeur "exacte" de l'écart type du délai de bout-en-bout en appliquant l'estimateur empirique classique. La Table 2.1 présente les résultats obtenus. Nous observons qu'une erreur relative de plus de 40% est commise sur l'écart type si l'on néglige les termes de covariance. Notons que, comme attendu, les

valeurs obtenues par les deux méthodes pour la valeur espérance coïncident. Basé sur cet exemple et sur de nombreux autres tests, nous pouvons conclure qu'il n'est pas possible, à partir des mesures locales aux nœuds, d'obtenir un estimateur précis de l'écart type du délai de bout-en-bout si l'on néglige les termes de covariance.

TABLE 2.1 – Erreur commise dans l'estimation de l'écart type du délai de bout-en-bout par la non prise en compte des termes de covariance

	Sans terme de covariance	Exact
Espérance	10.347 ms	10.347 ms
Écart type	2.923 ms	4.952 ms

Pour aller un peu plus loin, nous avons représenté dans la Figure 2.5 la distribution du délai de bout-en-bout obtenue par convolution des délais d'émission et d'attente sur l'ensemble des nœuds du chemin et la distribution "exacte" du délai de bout-en-bout (obtenu par le simulateur). La figure illustre bien l'importance de prendre en compte les termes de covariance (du moins certains) entre les délais d'émission et d'attente afin d'obtenir un estimateur suffisamment précis du délai de bout-en-bout.

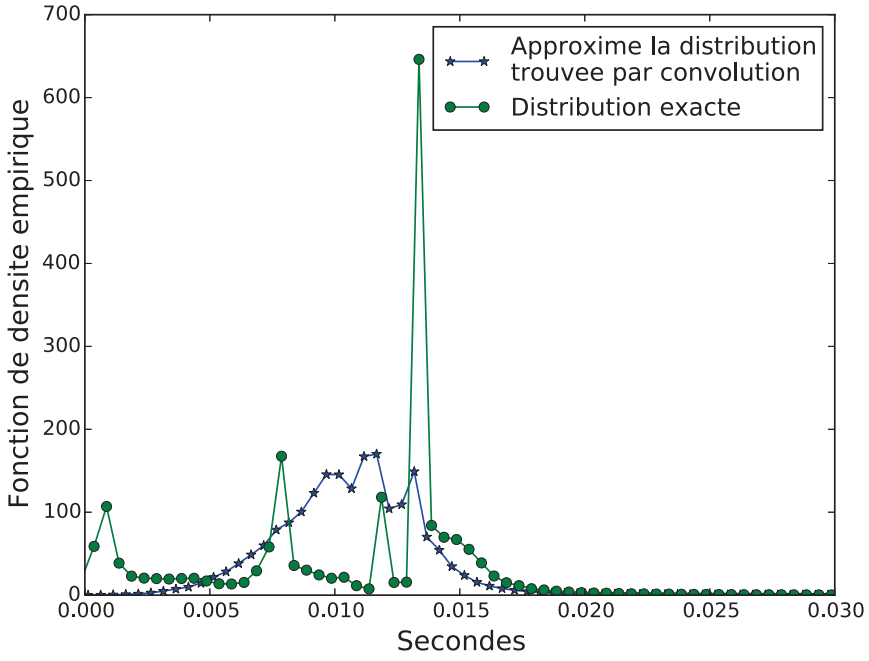


FIGURE 2.5 – Écart entre la distribution "exacte" du délai de bout-en-bout et celle obtenue par convolution des délais sur chaque nœud du chemin.

2.5 Conclusions

Après avoir présenté les principaux délais existants dans les réseaux, nous avons défini le délai de bout-en-bout d'un paquet sur un chemin réseau. Une connaissance, même partielle à travers ses deux premiers moments, du délai de bout-en-bout peut être très utile pour caractériser les performances des paquets sur ce chemin réseau. Nous avons

présenté deux façons (inégalité de Chebyshev et approche paramétrique) d'exploiter cette connaissance. Enfin, nous avons expliqué pourquoi l'écart type du délai de bout-en-bout est une quantité difficile à évaluer par la mesure à cause de la nécessité d'avoir des horloges hautement synchronisées sur les nœuds, ou par le calcul à partir des délais mesurés sur chaque nœud à cause des corrélations importantes existantes entre ces délais.
Ainsi, nous concluons qu'il serait utile pour une meilleure administration et gestion des ressources réseaux, en particulier dans le contexte d'un réseau SDN véhiculant du trafic non élastiques (type voix ou visioconférence), de disposer d'estimateurs permettant d'évaluer approximativement l'écart type du délai de bout-en-bout d'un chemin réseau à partir de mesures collectées localement sur les nœuds intermédiaires.
À notre connaissance, aucun travaux n'a présenté un tel estimateur. Dans les deux chapitres suivants, nous présentons deux façons différentes de le faire.

Chapitre 3

Méthode par décomposition

Contents

3.1 Introduction 19
3.2 Scénario, hypothèses et notations 19
3.3 L'espérance du délai de bout-en-bout 23
3.4 Méthode triviale pour l'estimation de la variance du délai de bout-en-bout 24
3.5 Méthode par récurrence 24
3.6 Méthode par décomposition 26
3.6.1 Décomposition de la variance 26
3.6.2 Description de la méthode par décomposition 27
3.7 Premiers résultats numériques 28
3.8 Conclusion 30

Résumé : *Ce chapitre présente une méthode pour estimer l'espérance et la variance sur le délai de bout-en-bout. Nous commençons par présenter le scénario, les hypothèses et les notations considérés. Puis nous décrivons une méthode très simple d'estimation de la variance du délai de bout-en-bout. Nous décrivons ensuite une approche par récurrence puis la méthode par décomposition. Cette méthode se base sur la décomposition de la variance du délai de bout-en-bout en plusieurs composantes et nous expliquons comment calculer chacune d'elles. La méthode par décomposition est l'ensemble de ces calculs. Ce chapitre se termine avec des premiers résultats numériques.*

3.1 Introduction

La mesure du délai de bout-en-bout est un paramètre important dans les réseaux. Le résultat obtenu sur la mesure du délai de bout-en-bout peut, par exemple, nous aider à déterminer la performance d'un réseau. On peut s'intéresser à différentes mesures du délai de bout-en-bout dans un réseau. Par exemple, on peut avoir besoin du délai de bout-en-bout maximal ou du délai de bout-en-bout minimal dans un réseau. La distribution du délai de bout-en-bout dans un réseau peut apporter beaucoup d'informations utiles sur ce réseau.
Comme nous avons discuté, certaines distributions se caractérisent grâce à ses deux premiers moments, comme, par exemple, la distribution gaussienne. Dans cette thèse, nous nous intéressons à l'espérance et à la variance du délai de bout-en-bout. La connaissance de ces deux premiers moments est une première étape vers une meilleure compréhension de la distribution du délai de bout-en-bout.
Comme nous allons le voir dans la suite de ce chapitre, en se basant sur des mesures de délai réalisées localement par chacun des nœuds du réseau, il est possible, grâce à l'architecture SDN, de déterminer l'espérance du délai de bout-en-bout. Il est, en revanche, plus compliqué de déterminer la variance du délai de bout-en-bout à partir de ces mesures locales, car les délais locaux ne sont pas des variables indépendantes.
Dans ce chapitre, nous allons tout d'abord présenter, dans la section 3.2, le scénario sur lequel notre étude repose, les hypothèses réalisées et les notations choisies. Ces notations sont aussi données dans la table 3.1. Nous présenterons ensuite, dans la section 3.4, une méthode très simple d'estimation de la variance du délai de bout-en-bout, appelée méthode triviale. Une méthode par récurrence sur le nœud 1 est donnée dans la section 3.5. Une autre méthode, plus précise et dénommée méthode par décomposition, sera présentée à la section 3.6. Des premiers résultats d'évaluation de ces méthodes sont donnés dans la section 3.7.

3.2 Scénario, hypothèses et notations

Il est tout d'abord important de caractériser le réseau que nous allons étudier. Nous remarquons que le délai de bout-en-bout est un paramètre particulier car il correspond très souvent au délai de bout-en-bout mesuré sur un chemin dans le réseau. Par conséquent, dans notre étude, les topologies considérées sont tout simplement des chemins et n'auront pas de structure plus complexe. Les chemins considérés sont orientés : nous nous intéressons au délai de bout-en-bout que les paquets vont subir lorsqu'ils traversent le chemin d'une extrémité à l'autre.
Nous supposons qu'il y a un seul tampon par lien de communication, tampon servant à stocker les paquets en attente de transmission sur ce lien. Comme nous nous intéressons au délai de bout-en-bout, seuls les tampons des liens traversés par les paquets sur le chemin étudié seront impliqués dans le délai de bout-bout et donc considérés. Il est possible que plusieurs flux traversent un même lien. Les paquets de ces différents flux vont alors se partager le tampon du lien traversé. Enfin, il est possible que des flux traversent des liens du chemin dans le sens opposé à celui qui nous intéresse. Ces flux n'auront pas d'impact sur le délai de bout-en-bout étudié car nous supposons que les liens sont full-duplex et que les tampons traversés par ces flux sont distincts (et indépendants) de ceux utilisés pour une transmission dans le sens inverse.
La figure 3.1 montre un exemple de simplification du réseau initial (les notations seront expliquées un peu plus loin dans ce chapitre). Si le but est d'étudier le délai de bout-en-bout du nœud 1 vers le nœud 4, en passant par les nœuds 2 et 3, alors les liens (1,3)

et (2,4) n'ont pas à être considérés pour notre étude car ils n'auront pas d'impact sur le délai de bout-en-bout du chemin étudié, de même que le flux $f_{2,1}$ qui traverse un lien du chemin dans le sens opposé à celui de l'étude.

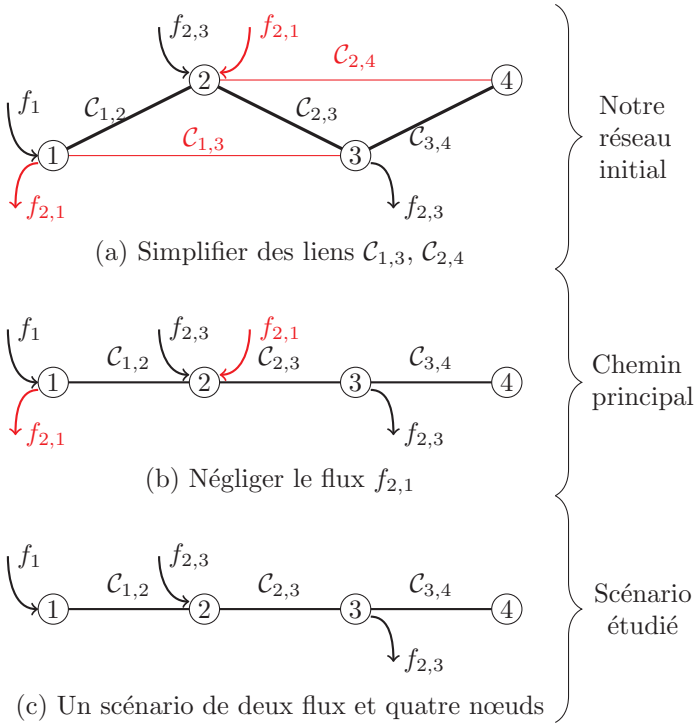


FIGURE 3.1 – Simplification du réseau réel en un scénario d'étude.

Le chemin considéré est constitué de $\mathcal{N}$ nœuds ($\mathcal{N} \geq 2$), indicés de 1 à $\mathcal{N}$. On s'intéresse au délai de bout-en-bout subit par les paquets traversant le chemin du nœud 1 vers le nœud $\mathcal{N}$. Les nœuds 1 et $\mathcal{N}$ sont donc les extrémités du chemin étudié. La figure 3.1 est un exemple de scénario avec $\mathcal{N} = 4$ nœuds.

Le lien entre le nœud k et le nœud $(k+1)$ est noté lien $(k, k+1)$. Nous utilisons la notation $\mathcal{C}_{k,k+1}$ pour indiquer la capacité du lien $(k, k+1)$. Elle est mesurée en bps (bits per second). Nous utilisons la notations $\mathcal{R}_{k,k+1}$ pour indiquer le temps de propagation sur le lien $(k, k+1)$.

Un flux, en général, est un ensemble des paquets qui partagent un même chemin. Comme dans notre modèle, nous nous intéressons seulement à un chemin, nous définissons donc : **a)** *Flux primaire* : ensemble des paquets qui traversent le chemin étudié du nœud 1 vers le nœud $\mathcal{N}$. On note f_1 le flux primaire. $\mathcal{M}$ indique le nombre total des paquets du flux primaire. **b)** *Flux secondaire* : ensemble de paquets qui croisent le chemin du flux primaire. On note $f_{k,l}$ le flux secondaire qui entre au nœud k et qui sort du nœud l. Nous écrirons simplement f_k quand le flux secondaire va jusqu'à bout ($l = \mathcal{N}$). Le flux secondaire $f_{k,l}$ partage donc le chemin du flux primaire sur les liens $(k, k+1)$, $(k+1, k+2)$, ..., $(l-1, l)$. Dans la figure 3.1, nous avons deux flux dans le scénario d'étude : un flux primaire f_1 et un flux secondaire $f_{2,3}$. Il est à noter que si plusieurs flux sont transmis sur un même chemin, alors ils sont considérés comme un seul flux dans notre étude.

Notation	Signification
$\mathcal{N}$	Nombre total de nœuds dans un scénario.
D	Variable aléatoire décrivant le délai de bout-en-bout.
$D_{k,k+1}$	Variable aléatoire décrivant le délai local d'un paquet traversant le nœud k vers le nœud $(k+1)$.
$\mathcal{K}$	Constante décrivant la taille du tampon associé à une interface du lien. Cette valeur est mesurée en octets. Elle est identique sur les nœuds.
$\mathcal{C}_{k,k+1}$	Constante décrivant la capacité du lien $(k, k+1)$. Cette valeur est mesurée en Mbps (ou Gbps).
$\mathcal{R}_{k,k+1}$	Constante décrivant le temps de propagation sur le lien entre le nœud k et le nœud $k+1$.
$\mathcal{S}_{k,k+1}$	Variable aléatoire décrivant le temps de transmission au nœud k pour aller au nœud $k+1$.
$s_{k,k+1}(m)$	Échantillon du temps de transmission au nœud k pour aller au nœud $k+1$ pour le paquet m.
$\mathcal{W}_{k,k+1}$	Variable aléatoire décrivant le temps d'attente au nœud k pour aller au nœud $k+1$.
$w_{k,k+1}(m)$	Échantillon du temps d'attente au nœud k pour aller au nœud $k+1$ pour le paquet m.
$\mathcal{W}^i_{k,k+1}$	Variable aléatoire résultat de la formule de récurrence appliquée par le nœud i pour estimer le temps d'attente au nœud k pour aller au nœud $k+1$.
$w^i_{k,k+1}(m)$	Résultat de la formule de récurrence appliquée par le nœud i pour estimer le temps d'attente au nœud k pour aller au nœud $k+1$ pour le paquet m.
$\widehat{\mathcal{W}^i_{k,k+1}}$	Variable aléatoire résultat de la formule de récurrence appliquée par le nœud i pour estimer le temps d'attente au nœud k pour aller au nœud $k+1$ mais en supposant que le nœud i connaît exactement les inter-arrivées au nœud k.
$\widehat{w}^i_{k,k+1}(m)$	Échantillon résultat pour le paquet m de la formule de récurrence appliquée par le nœud i pour estimer le temps d'attente au nœud k pour aller au nœud $k+1$ mais en supposant que le nœud i connaît exactement les inter-arrivées au nœud k.
I_k	Variable aléatoire décrivant les inter-arrivées au nœud k.
$i_k(m)$	Échantillon de l'inter-arrivée au nœud k du paquet m par rapport au paquet précédent.
$I^l_{k,j}$	Variable aléatoire décrivant les inter-arrivées au nœud k. Seuls les paquets passés par le nœud j sont considérés. $l = k$ ou j et indique lequel des nœuds (k ou j) calcule cette inter-arrivée.
$i^l_{k,j}(m)$	Inter-arrivée au nœud k du paquet m. Le paquet m est forcément un paquet passé par le nœud j. $l = k$ ou j et indique lequel des nœuds (k ou j) calcule cette inter-arrivée.
f_1	Flux primaire : arrive au nœud 1, traverse le chemin complet et sort au nœud $\mathcal{N}$.
f_k	Flux secondaire : arrive au nœud k et sort au nœud $\mathcal{N}$.
$f_{k,l}$	Flux secondaire : arrive au nœud k et sort au noeud l $(l > k)$.
P_{f_1}	Variable aléatoire décrivant la taille des paquets pour le flux f_1.
$p(m)$	Taille du paquet m pour le flux f_1.

TABLE 3.1 – Notations principales

Nœud

Cette partie discute du rôle d'un nœud dans notre modèle. Nous considérons qu'un nœud peut réaliser trois fonctions basiques : **a)** *Transmission* : un nœud peut recevoir des pa-

quets qui viennent de son voisin de gauche sur le chemin (quand il en a un) et il retransmet ces paquets à son voisin de droite sur le chemin (quand il en a un). **b)** *Tâche simple* : un nœud peut exécuter des tâches de calcul simples ou bien sauvegarder des variables locales. Une tâche peut être une estimation classique comme l'espérance ou la variance. **c)** *Communication* : Nous supposons qu'un nœud ne communique des informations de controle qu'avec le contrôleur pour lui transmettre de manière régulière les résultats de ses calculs par exemple. Il n'y a pas de signialisation entre les nœuds.
Un nœud peut distinguer les différents flux quand il y a plusieurs flux sur ce nœud. Selon la méthode d'estimation de la variance du délai de bout-en-bout proposée (et présentée par la suite), nous avons deux manières de distinguer des flux différents : **a)** *Dans la méthode par décomposition* : un nœud a juste besoin de distinguer le flux primaire des flux secondaires. **b)** *Dans la méthode conditionnelle* : un nœud doit distinguer deux types de flux : le flux primaire avec les flux secondaires, et les flux secondaires entre eux. La deuxième condition est très importante car ce sera une condition nécessaire pour créer l'histogramme.
Enfin, il est important de noter que les nœuds n'ont pas besoin d'être synchronisés pour les solutions que nous proposons.

Tampon

La discipline de gestion du tampon que nous utilisons dans ce manuscrit est la politique FCFS (First Come First Service). Nous supposons que la taille du tampon est identique sur chaque nœud. Nous notons $\mathcal{K}$ la taille du tampon en octets.

Les différents délais

Nous considérons que la taille des paquets est une variable aléatoire notée P. P_{f_1} correspond à la variable aléatoire décrivant la taille des paquets du flux primaire. Un paquet subit différents types de délai :

- *Délai de traitement* : c'est le temps nécessaire, sur un nœud, pour analyser l'entête du paquet, pour déterminer sur quel lien de sortie envoyer le paquet, pour mettre/enlever un paquet d'une file d'attente, etc. Ce délai dépend de la puissance du nœud. Nous supposons qu'il est stable (presque constant) et négligeable (moins grand qu'une microsecondes).
- *Le temps d'attente* : c'est le temps passé par le paquet dans le tampon associé au lien de transmission sur lequel le paquet doit être transmis. Comme nous considérons un discipline FCFS, tous les paquets arrivés avant le paquet considéré et non encore transmis doivent d'abord être traités avant que le paquet considéré soit servi. Nous notons $\mathcal{W}_{k,k+1}$ le temps d'attente sur le nœud k dans le tampon associé au lien de communication $(k, k+1)$.
- *Délai de transmission* : c'est le temps nécessaire pour la transmission du paquet sur le lien de communication à partir du moment où le paquet est servi. Nous notons $\mathcal{S}_{k,k+1}$ le délai de transmission sur le lien $(k, k+1)$. Ce temps est le rapport entre la taille du paquet et la capacité du lien : $\mathcal{S}_{k,k+1} = \frac{P}{\mathcal{C}_{k,k+1}}$.
- *Délai de propagation* : c'est le temps nécessaire pour qu'un bit traverse un lien. Ce temps dépend de la vitesse de propagation du signal et de la longueur du lien. Ce délai est constant sur un lien. Nous notons $\mathcal{R}_{k,k+1}$ le délai de propagation sur le lien $(k, k+1)$.

Le délai local sur nœud k est une variable aléatoire, noté $D_{k,k+1}$ et est défini comme étant la somme des différents délais locaux (nous rappelons que nous négligeons le délai de

traitement) :

$$D_{k,k+1} = \mathcal{W}_{k,k+1} + \mathcal{S}_{k,k+1} + \mathcal{R}_{k,k+1}. \tag{3.1}$$

Le délai de bout-en-bout, noté D, est la somme des délais locaux subis sur les différents nœuds du chemin emprunté par le flux primaire : $D = \sum_{k=1}^{\mathcal{N}-1} D_{k,k+1}$. Nous pouvons noter que, comme le nœud $\mathcal{N}$ est la destination de flux primaire, le délai local est nul sur ce nœud.

3.3 L'espérance du délai de bout-en-bout

Comme indiqué précédemment le délai de bout-en-bout peut être exprimé de la manière suivante : $D = \sum_{k=1}^{\mathcal{N}-1} D_{k,k+1}$. Nous avons :

$$\begin{aligned}\sum_{k=1}^{\mathcal{N}-1} D_{k,k+1} &= \sum_{k=1}^{\mathcal{N}-1} \Big(\mathcal{W}_{k,k+1} + \mathcal{S}_{k,k+1} + \mathcal{R}_{k,k+1} \Big) \\ &= \sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1} + \sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1} + \sum_{k=1}^{\mathcal{N}-1} \mathcal{R}_{k,k+1}. \end{aligned} \tag{3.2}$$

À cette étape, nous appliquons la propriété de linéarité de l'espérance à l'équation 3.2 :

$$\begin{aligned}\mathbb{E}[\sum_{k=1}^{\mathcal{N}-1} D_{k,k+1}] &= \sum_{k=1}^{\mathcal{N}-1} \mathbb{E}[D_{k,k+1}] \\ &= \sum_{k=1}^{\mathcal{N}-1} \mathbb{E}\Big[\mathcal{W}_{k,k+1} + \mathcal{S}_{k,k+1} + \mathcal{R}_{k,k+1} \Big] \\ &= \sum_{k=1}^{\mathcal{N}-1} \Big(\mathbb{E}[\mathcal{W}_{k,k+1}] + \mathbb{E}[\mathcal{S}_{k,k+1}] + \mathbb{E}[\mathcal{R}_{k,k+1}] \Big) \qquad (3.3) \\ &= \sum_{k=1}^{\mathcal{N}-1} \mathbb{E}[\mathcal{W}_{k,k+1}] + \sum_{k=1}^{\mathcal{N}-1} \mathbb{E}[\mathcal{S}_{k,k+1}] + \sum_{k=1}^{\mathcal{N}-1} \mathbb{E}[\mathcal{R}_{k,k+1}]. \qquad (3.4) \end{aligned}$$

L'équation (3.3) indique que l'espérance du délai de bout-en-bout peut être calculée par des quantités locales. De plus, l'équation (3.4) indique que l'espérance du délai de bout-en-bout est simplement la somme de l'espérance de tous délais locaux pour tous les nœuds du chemin principal. Elle peut donc être simplement obtenue.

3.4 Méthode triviale pour l'estimation de la variance du délai de bout-en-bout

Intéressons nous maintenant à la variance du délai de bout-en-bout. Supposons, dans cette section, que les nœuds sont indépendants. Nous avons donc :

$$\mathbb{V}[\sum_{k=1}^{\mathcal{N}-1} D_{k,k+1}] = \sum_{k=1}^{\mathcal{N}-1} \mathbb{V}[D_{k,k+1}] \tag{3.5}$$

$$= \sum_{k=1}^{\mathcal{N}-1} \mathbb{V}\Big[\mathcal{W}_{k,k+1} + \mathcal{S}_{k,k+1} + \mathcal{R}_{k,k+1}\Big] \tag{3.6}$$

$$= \sum_{k=1}^{\mathcal{N}-1} \mathbb{V}\Big[\mathcal{W}_{k,k+1} + \mathcal{S}_{k,k+1}\Big] \tag{3.7}$$

$$= \sum_{k=1}^{\mathcal{N}-1} \mathbb{V}\Big[\mathcal{W}_{k,k+1}\Big] + \sum_{k=1}^{\mathcal{N}-1} \mathbb{V}\Big[\mathcal{S}_{k,k+1}\Big]. \tag{3.8}$$

Comme le délai de propagation est constant, il n'a aucun impact sur la variance et peut donc être retiré.
En supposant que les nœuds sont indépendants, nous obtenons une méthode très simple pour estimer la variance du délai de bout-en-bout. Cette méthode, appelée méthode triviale, nécessite de calculer la variance sur les délais locaux obtenus sur chacun des nœuds et de sommer les variances locales calculées par tous les nœuds du chemin.

3.5 Méthode par récurrence

À partir des informations sur les temps de transmission et les inter-arrivés des paquets, il est possible d'estimer le temps d'attente pour un paquet sur un nœud suivant sur le chemin étudié.
Nous savons que le nœud k peut mesurer les informations suivantes sur le $(m-1)$-ième paquet et le m-ième paquet :

1. $w_{k,k+1}(m-1)$ le temps d'attente du $(m-1)$-ième paquet.
2. $s_{k,k+1}(m-1)$ le temps de transmission du $(m-1)$-ième paquet.
3. $i_k(m)$ l'inter-arrivé du m-ième paquet par rapport au $(m-1)$-ième paquet sur le nœud k.

Alors le nœud k peut estimer le temps d'attente $w^k_{k,k+1}(m)$ du m-ième paquet de la manière suivante :

$$w^k_{k,k+1}(m) = \max\{0, w^k_{k,k+1}(m-1) + s_{k,k+1}(m-1) - i_k(m)\}\ , m > 1. \tag{3.9}$$

Il faut noter que le nœud k est aussi capable de mesurer les temps d'attente que les paquets vont subir sur ce nœud. Par la suite, le nœud k utilisera les temps d'attente mesurés sur le nœud k et non ceux qui peuvent être estimés grâce à cette fonction par récurrence.
La fonction par récurrence permet au nœud k d'estimer le temps inter-arrivé du m-ième paquet par rapport au $(m-1)$-ième paquet sur le nœud $(k+1)$ de la manière suivante :

$$i^k_{k+1,k}(m) = s_{k,k+1}(m) + \max\{0, i_k(m) - w^k_{k,k+1}(m-1) - s_{k,k+1}(m-1)\}. \tag{3.10}$$

Le nœud k est capable d'estimer le temps inter-arrivée du m-ième paquet par rapport au $(m-1)$-ième paquet sur le nœud $k+1$ s'il connaît les temps de transmission des m- et

$(m-1)$-ième paquets sur le lien $(k, k+1)$, le temps d'attente du $(m-1)$-ième paquet sur le nœud k et le temps inter-arrivée du m-ième paquet par rapport au $(m-1)$-ième paquet sur le nœud k.
En combinant l'équation (3.9) et l'équation (3.10), nous obtenons :

$$\begin{cases} i^k_{k+1,k}(m) & = s_{k,k+1}(m) + \max\{0, i_k(m) - w^k_{k,k+1}(m-1) - s_{k,k+1}(m-1)\} \\ w^k_{k+1,k+2}(m) & = \max\{0, w^k_{k+1,k+2}(m-1) + s_{k+1,k+2}(m-1) - i^k_{k+1,k}(m)\} \end{cases}, k \geq 1. \tag{3.11}$$

Ainsi, grâce à cette fonction de récurrence, le nœud k est capable d'estimer le temps d'attente de chacun des paquets sur le nœud $k+1$. Cette récurrence peut ensuite être appliquée sur le lien $(k+2, k+3)$, puis sur le lien $(k+3, k+4)$, etc., ainsi jusqu'au lien $(\mathcal{N}-1, \mathcal{N})$. Et ainsi le nœud k est capable d'estimer le temps d'attente de chacun des paquets sur tous les nœuds j tel que $k < j$.

Il est possible d'appliquer la fonction par récurrence uniquement sur le nœud 1. Cette estimation est donnée par l'algorithme 1 qui calcule, directement, le délai de bout-en-bout pour chacun des paquets qui traverse le nœud 1. Cet algorithme est réalisé par le nœud 1.

Algorithm 1 Estimation, par le nœud 1, du délai de bout-en-bout de chacun des paquets

Inputs :
$i_1(m)$ ▷ Temps inter-arrivée du m-ième paquet par rapport au $(m-1)$-ième paquet au niveau du nœud 1.
$p(m-1), p(m)$ ▷ Tailles des m-ième and $(m$-1)-ième paquets.
Output :
$d(m)$ ▷ Délai de bout-en-bout du m-ième paquet.
Calcul pour le m-ième paquet :
if $m = 1$ **then**
    $\forall j,\ w^1_{j,j+1}(1) \leftarrow 0$
    $\forall j,\ s_{j,j+1}(1) \leftarrow \frac{p(1)}{\mathcal{C}_{j,j+1}}$
    $d(1) \leftarrow \sum_j s_{j,j+1}(1)$
else
    $d(m) \leftarrow 0$
    for j from 1 to $N-1$ **do**
        $s_{j,j+1}(m-1) \leftarrow \frac{p(m-1)}{\mathcal{C}_{j,j+1}}$
        $s_{j,j+1}(m) \leftarrow \frac{p(m)}{\mathcal{C}_{j,j+1}}$
        Fonction de récurrence :
        $w^1_{j,j+1}(m) \leftarrow \max\{0, w^1_{j,j+1}(m-1) + s_{j,j+1}(m-1) - i^1_{j,1}(m)\}$
        $i^1_{j+1,1}(m) \leftarrow s_{j,j+1}(m) + \max\{0, i^1_{j,1}(m) - w^1_{j,j+1}(m-1) - s_{j,j+1}(m-1)\}$
        Update $d(m)$ **:**
        $d(m) \leftarrow d(m) + w^1_{j,j+1}(m) + s_{j,j+1}(m) + \mathcal{R}_{j,j+1}$
    end for
end if

Une fois l'algorithme 1 appliqué, le nœud 1 a estimé les délais de bout-en-bout des différents paquets qu'il a vu passer. Il peut alors calculer la variance sur ces estimations de délai de bout-en-bout afin d'estimer la variance du délai de bout-en-bout.
Cette méthode est a priori bien adapté pour les scénarios ne transmettant que le flux primaire. En effet, le nœud 1 n'est pas en mesure, avec cette méthode, de prendre en compte l'impact des flux secondaires sur les temps d'attente et les temps d'inter-arrivée des paquets du flux primaire sur les autres nœuds. La méthode par décomposition proposée juste après prend en compte cet impact.

3.6 Méthode par décomposition

L'hypothèse sur l'indépendance entre les nœuds est pratique mais est trop réductive. L'utilisation de la méthode par récurrence au nœud 1 ne permet pas de prendre en compte l'impact des flux secondaires. La méthode par décomposition cherche à résoudre ces limitations.

3.6.1 Décomposition de la variance

La variance du délai de bout-en-bout peut être développée comme suit :

$$\mathbb{V}[\sum_{k=1}^{\mathcal{N}-1} D_{k,k+1}] = \mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \left(\mathcal{W}_{k,k+1} + \mathcal{S}_{k,k+1} + \mathcal{R}_{k,k+1}\right)\right] \tag{3.12}$$

$$= \mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \left(\mathcal{W}_{k,k+1} + \mathcal{S}_{k,k+1}\right)\right] \tag{3.13}$$

$$= \mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1} + \sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1}\right] \tag{3.14}$$

$$= \mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1}\right] + \mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1}\right] + \sum_{1\leq k,l\leq \mathcal{N}-1} \mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{S}_{l,l+1}). \tag{3.15}$$

Nous allons examiner les différents termes de l'équation (3.15).

Variance $\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1}\right]$

Comme nous nous intéressons au délai de bout-en-bout du flux primaire, nous avons $\mathcal{S}_{k,k+1} = \frac{P_{f_1}}{\mathcal{C}_{k,k+1}}$. Nous pouvons donc calculer le temps de transmission si nous connaissons la taille des paquet du flux primaire P_{f_1} et la capacité du lien $\mathcal{C}_{k,k+1}$.
Nous avons donc :

$$\sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1} = \frac{P_{f_1}}{\mathcal{C}_{1,2}} + \frac{P_{f_1}}{\mathcal{C}_{2,3}} + \ldots + \frac{P_{f_1}}{\mathcal{C}_{\mathcal{N}-1,\mathcal{N}}}$$

$$= \left(\frac{1}{\mathcal{C}_{1,2}} + \frac{1}{\mathcal{C}_{2,3}} + \ldots + \frac{1}{\mathcal{C}_{\mathcal{N}-1,\mathcal{N}}}\right) P_{f_1}. \tag{3.16}$$

Si nous considérons la taille des paquet du flux primaire P_{f_1} comme une variable aléatoire et la capacité des liens comme une constante, nous avons la formule de la variance du temps de transmission suivante :

$$\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1}\right] = \left(\frac{1}{\mathcal{C}_{1,2}} + \frac{1}{\mathcal{C}_{2,3}} + \ldots + \frac{1}{\mathcal{C}_{\mathcal{N}-1,\mathcal{N}}}\right)^2 \mathbb{V}[P_{f_1}]. \tag{3.17}$$

L'équation (3.17) contient deux parties, une somme des inverses de la capacité des liens et la variance de la taille des paquets. La première partie peut être calculée facilement en supposant que les capacités du chemin emprunté par le flux primaire sont connues. La deuxième partie peut aussi être calculée grâce à une observation locale des tailles de paquets du flux primaire au nœud $\mathcal{N}$. Nous remarquons ici que nous calculons la variance $\mathbb{V}[P_{f_1}]$ au nœud $\mathcal{N}$ afin de ne prendre en compte que les paquets qui arrivent bien au nœud $\mathcal{N}$ et qui ne se sont pas perdus en chemin.

Covariance $\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{S}_{l,l+1})$

Selon la définition de la covariance nous avons :

$$\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{S}_{l,l+1}) = \mathbb{E}[\mathcal{W}_{k,k+1} \times \mathcal{S}_{l,l+1}] - \mathbb{E}[\mathcal{W}_{k,k+1}]\mathbb{E}[\mathcal{S}_{l,l+1}]. \tag{3.18}$$

En remplaçant $\mathcal{S}_{k,k+1}$ par $\frac{P_{f_1}}{\mathcal{C}_{k,k+1}}$, le nœud k peut calculer ce terme de covariance de la manière suivante :

$$\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{S}_{l,l+1}) \approxeq \mathbb{E}\left[\mathcal{W}_{k,k+1} \times \frac{P_{f_1}}{\mathcal{C}_{l,l+1}}\right] - \mathbb{E}[\mathcal{W}_{k,k+1}]\mathbb{E}\left[\frac{P_{f_1}}{\mathcal{C}_{l,l+1}}\right] \tag{3.19}$$

$$= \left(\mathbb{E}[\mathcal{W}_{k,k+1} \times P_{f_1}] - \mathbb{E}[\mathcal{W}_{k,k+1}]\mathbb{E}[P_{f_1}]\right) \frac{1}{\mathcal{C}_{l,l+1}}. \tag{3.20}$$

Nous remarquons que l'équation (3.19) est simplement une approximation car il est possible qu'il y ait des pertes de paquets et que les paquets présents au nœud k ne soient pas identiques à ceux présents au nœud l.
Le terme de covariance $\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{S}_{l,l+1})$ peut donc être estimé localement par le nœud k. Pour cela, il suffit que le nœud k détermine les temps d'attente des paquets du flux primaire sur le lien $(k, k+1)$, la taille des paquets du flux primaire et connaisse la capacité du lien $(l, l+1)$.

Variance $\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1}\right]$

La variance du temps d'attente $\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1}\right]$ est plus difficile à estimer que les termes précédents. En utilisant la formule classique de la variance, nous avons :

$$\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1}\right] = \sum_{k=1}^{\mathcal{N}-1} \mathbb{V}[\mathcal{W}_{k,k+1}] + 2 \times \sum_{1 \leq k < l \leq \mathcal{N}-1} \mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{W}_{l,l+1}). \tag{3.21}$$

Les termes de variance $\mathbb{V}[\mathcal{W}_{k,k+1}]$ peuvent être facilement calculés par chaque nœud k. En revanche, la covariance $\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{W}_{l,l+1})$ est plus difficile à calculer localement au nœud k, car ce dernier ne connaît pas le temps d'attente $\mathcal{W}_{l,l+1}$ au nœud l.
Par conséquent, nous allons chercher à estimer $\mathcal{W}_{l,l+1}$ par le nœud k.

3.6.2 Description de la méthode par décomposition

Dans cette méthode, nous allons nous baser sur la décomposition de la variance discutée dans la section 3.6.1 pour estimer la variance du délai de bout-en-bout. Pour rappel :

$$\mathbb{V}[\sum_{k=1}^{\mathcal{N}-1} D_{k,k+1}] = \mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1}\right] + \mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1}\right] + \sum_{1 \leq k,l \leq \mathcal{N}-1} \mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{S}_{l,l+1}). \tag{3.22}$$

Le terme $\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{S}_{k,k+1}\right]$ peut être calculé par le nœud $\mathcal{N}$ comme indiqué dans la section 3.6.1. Les termes $\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{S}_{l,l+1})$ (avec $1 \leq l < \mathcal{N}$) peuvent être calculés par le nœud k, comme indiqué dans la section 3.6.1.

Regardons maintenant plus précisément le terme $\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1}\right]$. Comme indiqué dans la section 3.6.1, nous avons :

$$\mathbb{V}\left[\sum_{k=1}^{\mathcal{N}-1} \mathcal{W}_{k,k+1}\right] = \sum_{k=1}^{\mathcal{N}-1} \mathbb{V}[\mathcal{W}_{k,k+1}] + 2 \times \sum_{1 \leq k < l \leq \mathcal{N}-1} \mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{W}_{l,l+1}). \tag{3.23}$$

Le terme $\mathcal{W}_{k,k+1}$ peut facilement être calculé par chaque nœud k. Pour déterminer $\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{W}_{l,l+1})$, chaque nœud k va utiliser la fonction de récurrence (3.11) pour estimer, pour chaque paquet m, $w^k_{l,l+1}(m)$ (pour $k < l < \mathcal{N}$). Ce seront ces estimations qui seront utilisées pour estimer $\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{W}_{l,l+1})$,
Tous les nœuds envoient ensuite leurs différentes estimations au contrôleur qui peut ensuite sommer toutes ces valeurs afin d'obtenir la variance du délai de bout-en-bout.
Nous pouvons noter que cette méthode ne nécessite pas de synchronisation des horloges des différents nœuds car tous les nœuds peuvent calculer leurs estimations localement.

3.7 Premiers résultats numériques

Cette section a pour but de donner une première idée des performances de la solution par décomposition et de comparer les deux solutions proposées pour cette approche. Des résultats plus détaillés seront donnés dans le chapitre 5.
Nous utilisons le simulateur NS-3 pour effectuer cette évaluation. Tous les paramètres utilisés dans NS-3 sont donnés dans le chapitre 5. Le scénario étudié contient 4 nœuds et est donné dans la figure 3.2. Les capacités des liens sont les suivantes : $\mathcal{C}_{1,2} = 8, \mathcal{C}_{2,3} = 5, \mathcal{C}_{3,4} = 3$ (Mbps).

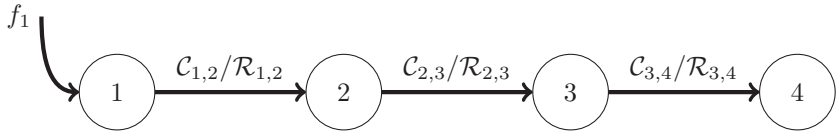


FIGURE 3.2 – Un scénario particulier avec un seul flux.

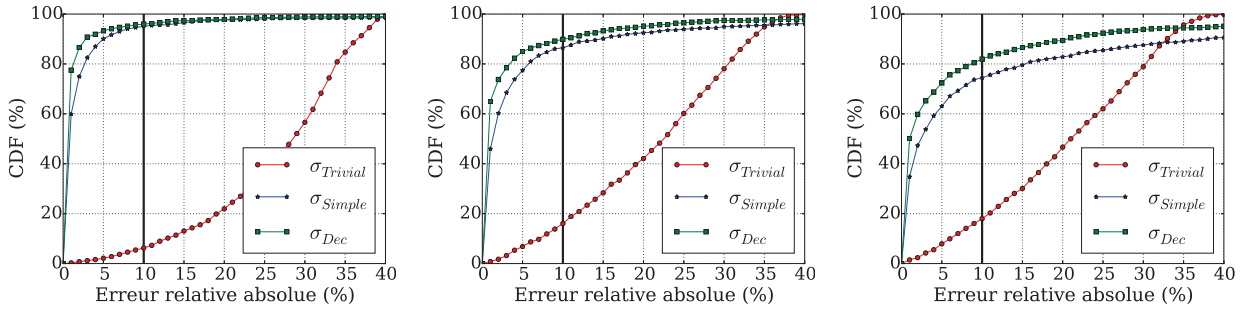
Le délai de propagation est identique pour tous les lien et est fixé à $\mathcal{R}_{1,2} = \mathcal{R}_{2,3} = \mathcal{R}_{3,4} = 2.10^{-3}$ secondes. La taille du tampon est aussi identique pour tous les nœuds et est égale à 15000 octets (ce qui est équivalent à 10 paquets de 1500 octets). La taille et l'inter-arrivé des paquets sont sélectionnées à partir de la trace réelle de l'université de Stuttgart. La trace comprend les paquets transmis dans une résidence de l'université de Stuttgart durant 4 heures (de 18h00 à 22h00). Les tailles des paquets varient entre 64 octets et 1500 octets comme indiqué dans la table 3.2.

TABLE 3.2 – Distribution de la taille des paquets dans la trace de l'université de Stuttgart.

Taille des paquets (octets)	< 150	150 - 1400	> 1400
Fréquence (%)	45.48	24.22	30.30

Dans les figures suivantes, nous présentons les résultats obtenus par les différentes solutions proposées dans cette section, à savoir la méthode triviale, appelée Trivial, et la méthode par décomposition. Pour cette dernière, l'approche développée pour un seul flux sera appelée Simple et l'approche développée pour plusieurs flux sera appelée Dec. Nous présentons les résultats obtenus sur l'écart-type σ ($\sigma = \sqrt{\mathbb{V}}$) plutôt que sur la variance.

Nous avons fait ce choix car l'unité de l'écart-type est le même que celui du délai de bout-en-bout. L'écart-type de la méthode triviale est noté σ_{Trivial}, celui de la méthode par récurrence est noté σ_{Simple} et celui de la méthode par décomposition σ_{Dec}. Grâce au simulateur NS-3, il est possible de capturer les délais de bout-en-bout exacts pour tous les paquets, et ainsi de calculer la variance (et aussi l'écart-type) exacte du délai de bout-en-bout. L'écart-type exact est noté σ_{Exact}.



(a) Débit de $f_1 = 1$ Mbps (b) Débit de $f_1 = 1.5$ Mbps (50% of bottleneck) (c) Débit de $f_1 = 2$ Mbps (66% of bottleneck)

FIGURE 3.3 – Fonction de répartition de l'erreur relative pour les trois méthodes proposées et pour trois débits pour le flux primaire.

Nous nous intéressons à l'erreur relative des solutions proposées. L'erreur relative correspond à $\frac{|\sigma_{\text{solution-testée}} - \sigma_{\text{Exact}}|}{\sigma_{\text{Exact}}}$. Sur la figure 3.3 nous présentons la fonction de répartition de l'erreur relative pour les différentes solutions proposées et pour pour 3 débits différents pour le flux primaire (1, 1,5 et 2 Mbps pour le débit de f_1).
Comme prévu, la méthode triviale ne donne pas de bons résultats. La cause principale provient de l'hypothèse de l'indépendance des délais sur les nœuds qui n'est pas valide dans les scénarios testés. On peut néanmoins noter que les résultats pour cette solution s'améliorent lorsque le débit de f_1 augmente. Ceci peut s'expliquer par le fait que quand le flux primaire a un débit plus élevé, le chemin devient plus chargé et les tampons sur les nœuds sont plus remplis. Ceci induit une certaine stabilité sur les temps d'attente dans les tampons qui a tendance à atténuer la dépendance entre les nœuds.
La méthode par récurrence et la méthode par décomposition présentent de très bons résultats quand le débit de f_1 est de 1 Mbps. La méthode par décomposition (σ_{Dec}) a des résultats sensiblement meilleurs. On voit que des résultats deviennent moins bons lorsque le débit du flux primaire augmente. Ceci peut s'expliquer par la perte des paquets qui augmente lorsque le débit de f_1 augmente. La méthode par décomposition (σ_{Dec}) présente toujours de meilleurs résultats que la méthode par récurrence (σ_{Simple}).
La table 3.3 fait une synthèse des résultats numériques obtenus. Selon la table, la méthode par décomposition est meilleure que la méthode triviale. De plus, la méthode par décomposition (σ_{Dec}) de la méthode par décomposition est toujours meilleure que la méthode par récurrence. C'est donc toujours cette approche que nous considérerons dans la suite de ce manuscrit.

TABLE 3.3 – Synthèse sur la fonction de répartition de l'erreur relative pour l'écart-type du délai de bout-en-bout.

	Débit de $f_1 = 1$ Mbps				Débit de $f_1 = 1.5$ Mbps				Débit de $f_1 = 2$ Mbps			
Erreur	$\leq$ 10%	$\leq$ 20%	$>$ 30%	Moyen	$\leq$ 10%	$\leq$ 20%	$>$ 30%	Moyen	$\leq$ 10%	$\leq$ 20%	$>$ 30%	Moyen
$\sigma_{Trivial}$	**6.**25	**21.**94	**43.**40	**26.**42	**16.**04	**42.**08	**21.**94	**21.**18	**17.**99	**46.**67	**21.**11	**20.**47
σ_{Simple}	**95.**07	**97.**57	**1.**67	**3.**05	**86.**46	**92.**50	**5.**07	**6.**56	**74.**44	**82.**85	**12.**43	**12.**22
σ_{Dec}	**95.**97	**97.**78	**1.**11	**1.**90	**89.**86	**95.**14	**2.**57	**4.**21	**81.**94	**89.**37	**6.**18	**7.**76

3.8 Conclusion

Dans ce chapitre, nous avons présenté des méthodes permettant le calcul des deux premiers moments du délai de bout-en-bout. L'espérance est très facile à calculer localement. Par contre la variance du délai de bout-en-bout est très compliquée. La cause principale vient de la corrélation entre les temps d'attente dans deux nœuds différents.
Nous avons présenté une approximation qui applique la fonction par récurrence au nœud 1. L'approximation fonctionne bien sauf dans le cas où il y a des paquets perdus. Nous avons aussi proposé une autre approximation, appelée méthode par décomposition.
Nous avons évalué numériquement ces deux méthodes dans un cas très simple. Selon ces résultats, nous trouvons que la méthode par décomposition fonctionne au moins dans un cas simple d'un seul flux. Au chapitre suivant, nous allons étudier un aspect de l'indépendance des flux, puis nous allons proposer à une autre approximation.

Chapitre 4

Méthode utilisant l'espérance conditionnelle (méthode conditionnelle)

Contents

4.1 Introduction 33
4.2 Problématique 33
4.3 Méthode conditionnelle 35
4.3.1 Bref rappel sur les espérances conditionnelles 35
4.3.2 Scénarios et notations 35
4.4 Méthode conditionnelle 37
4.4.1 Sans flux secondaire f_2 37
4.4.2 Avec flux secondaire f_2 38
4.4.3 Cas général : heuristique 38
4.5 Premiers résultats numériques 39
4.6 Conclusion 41

Résumé : *Ce chapitre présente une seconde méthode permettant d'approximer la variance du délai de bout-en-bout. Nous commençons par discuter des inconvénients de la méthode par décomposition, en particulier dans le cas où les flux sont dépendants. Nous introduisons l'idée théorique de cette méthode et un bref rappel sur les espérances conditionnelles. Puis, nous présentons l'algorithme permettant d'évaluer la covariance sur les temps d'attentes. Ce chapitre se termine par une discussion sur des résultats de simulations pour deux scénarios particuliers.*

4.1 Introduction

Dans le chapitre précédent nous avons introduit et étudié la variance du délai de bout-en-bout. La méthode triviale suppose que les délais d'attentes sur les différents nœuds sont indépendants. La variance du délai est alors la somme des variances locales à chaque nœud. Cette approximation néglige donc la corrélation entre les nœuds. Cette approximation est très pratique car elle ne demande pas beaucoup de calcul. Par contre, elle perd en précision comme nous avons pu le voir au chapitre précédent. L'autre approximation est la méthode par décomposition. Celle-ci utilise une décomposition de la variance du délai en plusieurs composantes. Nous avons montré que la plupart des composantes pouvait être calculée localement sur chaque nœud du chemin. Le seul terme qui ne peut pas être calculé localement est la covariance des temps d'attente :

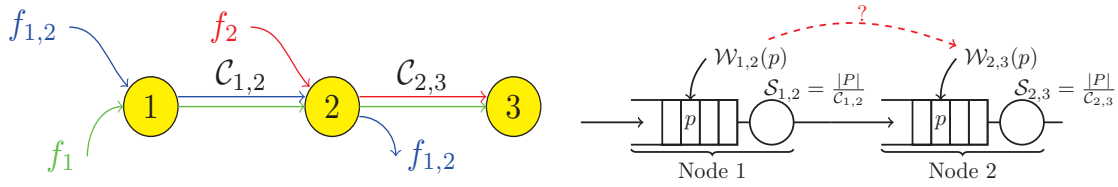
$$\mathrm{Cov}(\mathcal{W}_{k,k+1}, \mathcal{W}_{l,l+1}) = \mathbb{E}[\mathcal{W}_{k,k+1}\mathcal{W}_{l,l+1}] - \mathbb{E}[\mathcal{W}_{k,k+1}]\mathbb{E}[\mathcal{W}_{l,l+1}]. \tag{4.1}$$

où $\mathcal{W}_{k,k+1}, \mathcal{W}_{l,l+1}$ sont les temps d'attente au nœud k (vers le nœud $k+1$) et au nœud l (vers le nœud $l+1$).
La problématique se résume donc à l'évaluation de cette quantité. Nous nous concentrons exclusivement au calcul de celle-ci dans ce chapitre. Les deux variables aléatoires ($\mathcal{W}_{k,k+1}$ et $\mathcal{W}_{l,l+1}$) sont corrélées à cause de la taille des paquets (et donc de leur distribution) et du fait que la sortie de la file k (si $k < l$) est l'entrée de la file l plus loin dans le réseau.
La méthode triviale négligeait simplement cette covariance. Le principe de la méthode par décomposition est d'estimer, au nœud k, le temps d'attente sur la file l grâce à la formule de récurrence. La covariance $\mathbb{E}[\mathcal{W}_{k,k+1}\mathcal{W}_{l,l+1}]$ est donc approchée par $\mathbb{E}[\mathcal{W}_{k,k+1}\mathcal{W}^k_{l,l+1}]$ où $\mathcal{W}^k_{l,l+1}$ est donnée par la formule de récurrence. Mais la formule de récurrence ne prend pas en compte les délais introduits par les flux secondaires/externes qui arrivent sur les nœuds suivants (suivants au nœud k ici). La méthode s'appuie sur l'idée que, quand ces flux secondaires sont indépendants du flux primaire, les délais rajoutés par ces flux externes restent relativement décorrélés des délais sur les nœuds précédents. On peut alors négliger les flux externes dans le calcul de la covariance.
Nous avons donc chercher une méthode qui ne faisait pas cette hypothèse d'indépendance entre le flux primaire et les flux secondaires. La méthode proposée s'appelle *la méthode conditionnelle*. Elle a pour but de donner une approximation correcte de la covariance dans le cas général où les flux ne sont pas obligatoirement indépendants entre eux.
La suite de ce chapitre est organisé de la manière suivante. Tout d'abord nous présentons la problématique générale et l'idée de notre méthode. Deux rappels de résultats sur les espérances conditionnelles précèdent l'introduction du scénario, des notations et des hypothèses. Puis nous décrivons en détail la méthode. À la fin de ce chapitre, nous discutons des performances de la méthode conditionnelle au travers de quelques résultats de simulations.

4.2 Problématique

La difficulté dans l'estimation de la covariance donnée par l'équation (4.1) est liée au fait que les nœuds ne communiquent pas entre eux et ne peuvent pas s'échanger la valeur des échantillons. En effet, un estimateur simple et sans biais consisterait à rapprocher les échantillons un à un, c'est-à-dire que pour chaque paquet, un noeud k conserverait le temps d'attente local de celui-ci, récupérait le temps d'attente de ce paquet sur la file suivante (ou une autre file), et effectuerait la moyenne empirique du produit sur ces temps d'attente pour un ensemble de paquets. Ce principe est illustré figure 4.1(b). La complexité

protocolaire et l'*overhead* nécessaire pour la mise en œuvre d'une telle méthode la rend inapplicable.



(a) Un scénario de trois nœuds, trois flux.

(b) Le modèle de file d'attente sur les deux premiers nœuds. Le problème pour estimer la covariance est de rapprocher le temps d'attente des paquets sur les deux files.

FIGURE 4.1 – Scénario et problème de la corrélation des temps d'attente dans le cas de trois nœuds.

Nous ne pouvons donc pas faire directement le produit $\mathcal{W}_{k,k+1}\mathcal{W}_{l,l+1}$ pour chaque paquet. Plutôt, nous proposons d'utiliser un ensemble de conditions communes aux nœuds qui permettent de classifier les temps d'attente des paquets pour les regrouper ensuite. Appelons une de ces conditions $\mathcal{G}$. Nous utilisons la notation $\mathcal{W}_{k,k+1}|_{\mathcal{G}}$ pour décrire les temps d'attente au nœud k qui satisfont la condition $\mathcal{G}$. D'un point de vue pratique, cela décrit l'ensemble des paquets pour lequel les délais d'attentes vérifient la condition $\mathcal{G}$.
L'idée est la suivante. Nous définissons un ensemble de conditions qui permettent de classifier les échantillons des délais d'attentes de manière disjointe. Pour l'ensemble des paquets satisfaisant une des conditions $\mathcal{G}$, le nœud k calcule la moyenne empirique de l'espérance $\mathbb{E}[\mathcal{W}_{k,k+1}|_{\mathcal{G}}]$. Le nœud l ($l > k$) effectue le même calcul et obtient un estimateur de l'espérance $\mathbb{E}[\mathcal{W}_{l,l+1}|_{\mathcal{G}}]$. Le produit de ces deux espérances, pour la même condition, est calculé : $\mathbb{E}[\mathcal{W}_{k,k+1}|_{\mathcal{G}}]\mathbb{E}[\mathcal{W}_{l,l+1}|_{\mathcal{G}}]$. L'estimateur final est la somme de ces produits sur l'ensemble des conditions :

$$\sum_{\mathcal{G}} \mathbb{E}[\mathcal{W}_{k,k+1}|_{\mathcal{G}}]\mathbb{E}[\mathcal{W}_{l,l+1}|_{\mathcal{G}}]\mathbb{P}(\mathcal{G})$$

où $\mathbb{P}(\mathcal{G})$ est la probabilité que la condition $\mathcal{G}$ se produise. D'un point de vue protocolaire, les espérances conditionnelles sont calculées localement par chaque nœud, transmis au contrôleur qui effectue les produits et les sommes.
Le cœur de la méthode s'appuie donc sur l'ensemble des conditions choisies. Afin d'être le plus efficace possible, ces conditions devraient vérifier trois critères :

1. *Classification :* les conditions sur les délais d'attente des paquets doivent permettre de classifier les paquets dans des groupes disjoints. Autrement dit, ces conditions doivent amener à former une partition. Le nombre de conditions/classes doit être choisit de manière à offrir un bon compromis entre précision et complexité. Un nombre trop faible amènera à un résultat proche du produit $\mathbb{E}[\mathcal{W}_{k,k+1}]\mathbb{E}[\mathcal{W}_{l,l+1}]$ et sera en conséquence incapable d'estimer la corrélation. Un nombre trop important de conditions/classes rendra la méthode difficile à mettre en œuvre : générera un nombre de valeurs importants, un grand nombre de messages, ou des messages volumineux aux contrôleurs, des calculs importants, etc.
2. *Calculabilité :* la condition $\mathcal{G}$ doit être calculable sur l'ensemble des nœuds pour l'ensemble des échantillons de temps d'attentes. Elle doit assurer la mise en relation des temps d'attente $\mathcal{W}_{k,k+1}|_{\mathcal{G}}$ et $\mathcal{W}_{l,l+1}|_{\mathcal{G}}$ (pour deux files k et l ici). Autrement dit, l'attente d'un paquet donné à la file k qui vérifie la condition $\mathcal{G}$ doit aussi vérifier cette même condition à la file l.

3. *Complexité :* la classification, fonction de ces conditions, doit se faire au travers de calculs simples. Si les calculs sont trop complexes, les opérations à effectuer pour chaque paquet peuvent rendre la méthode rédhibitoire.

La discussion ci-dessus décrit l'idée principale de la méthode conditionnelle. Dans la suite, nous effectuons un rappel sur l'espérance conditionnelle avant de présenter scénarios, notations et finalement la méthode.

4.3 Méthode conditionnelle

4.3.1 Bref rappel sur les espérances conditionnelles

Avant de présenter la méthode conditionnelle, nous rappelons deux résultats de base des probabilités nécessaires à la compréhension de la méthode.
Nous démarrons par un théorème classique :

Théorème du changement de variable

Soit g une fonction de $\mathbb{R}$ dans $\mathbb{R}$, et X une variable aléatoire à valeur réelle admettant une densité de probabilité $f(.)$, alors $g(X)$ est également une variable aléatoire. L'espérance de $g(X)$ peut alors être calculée de la manière suivante :

$$\mathbb{E}[g(X)] = \int_{x\in\mathbb{R}} g(x)f(x)\mathtt{d}x. \tag{4.2}$$

Le deuxième théorème porte sur l'espérance conditionnelle. Il est à la base de la méthode proposée.

Théorème sur l'espérance conditionnelle

Soient X et Y deux variables aléatoires, alors

$$\mathbb{E}[X] = \mathbb{E}\big[\mathbb{E}[X|Y]\big]. \tag{4.3}$$

et par extension,

$$\mathbb{E}[XY] = \mathbb{E}\big[\mathbb{E}[X|Y]Y\big]. \tag{4.4}$$

Le côté gauche de l'équation (4.4) nous donne l'espérance du produit entre deux variables aléatoires $\mathbb{E}[XY]$ qui peut être difficile à calculer quand les variables aléatoires sont corrélées. Le côté droit nous permet de calculer séparément deux termes : X et $\mathbb{E}[Y|X]$, où le deuxième représente la moyenne de Y sachant X.
Cette décomposition a un intérêt pratique par rapport à notre problème. Dans $\mathbb{E}[XY]$ nous devons faire le produit XY pour chaque réalisation. Par contre dans $\mathbb{E}\big[X\mathbb{E}[Y|X]\big]$, nous pouvons regrouper les réalisations de Y pour des valeurs données de X, puis ensuite faire le produit $X\mathbb{E}[Y|X]$.

4.3.2 Scénarios et notations

Nous introduisons quelques notations supplémentaires par rapport au chapitre 3. Elles sont nécessaires à la présentation de la méthode. Nous nous focalisons sur une chaîne composée de 3 nœuds. Le cas général, avec un nombre de nœuds arbitraire, est simplement une généralisation de celui-ci.

Nous considérons trois nœuds et deux liens comme représenté sur la figure 4.1(a). Le flux primaire f_1 entre dans la chaîne au nœud 1. Le flux secondaire $f_{1,2}$ rentre également au nœud 1 mais sort de la chaîne avant le flux primaire. Le flux secondaire f_2 rentre au nœud 2.
Nous supposons que le nœud 2 est capable de distinguer trois types de paquet :

a) Les paquets du flux primaire f_1 : c'est une condition nécessaire puisque l'objectif de la méthode est l'estimation du délai de bout-en-bout pour ce flux particulier.

b) Les paquets des flux secondaires qui ne traversent pas le nœud 1 : le flux f_2 dans notre exemple.

c) Les paquets des flux secondaires qui traversent le nœud 1 : le flux $f_{1,2}$ dans notre exemple. Ces paquets sont connus du nœud précédent, nœud 1 ici.

Comme pour la méthode par décomposition, nous supposons qu'un nœud peut connaître le moment d'arrivé des paquets (dans ses propres files). Les informations mesurées sont donc les moments d'arrivées et les tailles des paquets.
Nous distinguons différentes variables aléatoires pour les inter-arrivées en fonction des flux :

a) I_2 désigne l'inter-arrivée d'un paquet sur le nœud 2 par rapport à l'arrivée du paquet précédent. Le paquet précédent peut appartenir au flux primaire ou non.

b) $I^2_{2,1}$ est l'inter-arrivée d'un paquet sur le nœud 2 (première indice) par rapport au paquet précédent provenant du nœud 1 (deuxième indice) donc appartenant à $f_{1,2}$ ou f_1 ici. Autrement dit, les arrivées du flux secondaire f_2 (arrivant directement sur le noeud 2) ne sont pas pris en compte. Cette inter-arrivée est calculée au nœud 2 (d'où l'exposant).

c) $I^1_{2,1}$ désigne l'inter-arrivée d'un paquet sur le nœud 2, mesurée ou calculée par le nœud 1. Le nœud 1 n'ayant pas connaissance du flux f_2 il le néglige. Le nœud 1 peut mesurer les temps de départs des paquets ou bien appliquer une formule de récurrence similaire à (3.11) pour l'inférer. Les deux inter-arrivées $I^2_{2,2}$ et $I^1_{2,1}$ coïncident excepté lorsqu'il y a des pertes de paquets.

Nous noterons $i_2(p)$, $i^2_{2,1}(p)$, et $i^1_{2,1}(p)$ les échantillons correspondants pour l'inter-arrivée d'un paquet particulier p.

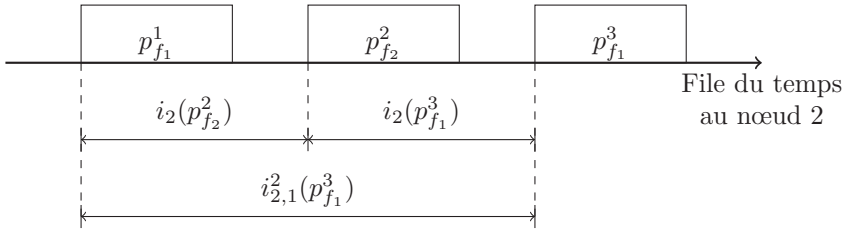


FIGURE 4.2 – Exemple illustrant les différentes notations liées aux inter-arrivées des paquets : $i_2(p^3_{f_1})$ et $i^2_{2,1}(p^3_{f_1})$.

La figure 4.2 présente un exemple illustrant les différences entre les différents types d'inter-arrivée. Soient $p^1_{f_1}$, $p^2_{f_2}$, et $p^3_{f_1}$ trois paquets consécutifs traversant le nœud 2 (l'indice indique le flux auquel il appartient). On observe que l'inter-arrivé $i^2_{2,1}(p^3_{f_1})$ néglige la présence du paquet $p^2_{f_2}$ appartenant au flux secondaire.
Nous introduisons également trois quantités :

a) $\mathcal{W}_{1,2}$ et $\mathcal{W}_{2,3}$ sont les variables aléatoires décrivant les "vrais" temps d'attente respectivement aux nœuds 1 et 2.

b) $\widehat{\mathcal{W}}^1_{2,3}$ est un estimateur de $\mathcal{W}_{2,3}$ par le nœud 1. Le nœud 1 approxime le temps d'attente sur le noeud 2 pour ce paquet à partir de la formule de récurrence. Cette formule de récurrence prend en compte les inter-arrivées "I_2" définit au-dessus, autrement dit les inter-arrivées du flux primaire et des flux secondaires. Le nœud 1 n'ayant pas connaissance des flux secondaires arrivant directement sur le nœud 2 (flux f_2), il lui est impossible de calculer $\widehat{\mathcal{W}}^1_{2,3}$ en pratique. Cependant, nous introduisons malgré tout cette notation pour discuter de la pertinence des approximations que nous ferons.

c) $\mathcal{W}^1_{2,3}$ et $\mathcal{W}^2_{2,3}$ sont les résultats de la fonction par récurrence calculée par le nœud 1, respectivement 2, en utilisant les inter-arrivées qui négligent le flux f_2, donc $I^1_{2,1}$ pour $\mathcal{W}^1_{2,3}$ et $I^2_{2,1}(p)$ pour $\mathcal{W}^2_{2,3}$. Les paquets des flux secondaires arrivant directement sur le nœud 2 ne sont donc pas pris en compte.

Comme précédemment, les réalisations de ces variables aléatoires pour un paquet particulier p sont notées en minuscules : $w_{1,2}(p)$, $\widehat{w}^1_{2,3}(p)$, $w^1_{2,3}(p)$ et $w^2_{2,3}(p)$.

4.4 Méthode conditionnelle

Cette section présente les détails de la méthode conditionnelle. Elle est séparée en deux parties. Les deux premières parties présentent l'aspect mathématique de la méthode. La deuxième traite des aspects pratiques et algorithmiques.
Afin de présenter la méthodologie utilisée, nous reprenons le scénario consistant en une chaîne de trois nœuds comme représenté dans la figure 4.1. Nous considérons la covariance entre les nœuds 1 et 2. Nous commençons par traiter deux cas (avec et sans flux secondaire f_2) puis nous généralisons.

4.4.1 Sans flux secondaire f_2

En utilisant les notations précédentes, nous avons :

$$\mathbb{E}[\mathcal{W}_{1,2}\mathcal{W}_{2,3}] = \mathbb{E}\big[\mathbb{E}[\mathcal{W}_{1,2} \mid \mathcal{W}_{2,3}]\mathcal{W}_{2,3}\big] \tag{4.5}$$

L'équation (4.5) est une application directe du théorème de l'équation (4.4), avec $X = \mathcal{W}_{1,2}$ et $Y = \mathcal{W}_{2,3}$. La variable $\mathbb{E}[\mathcal{W}_{1,2} \mid \mathcal{W}_{2,3}]$ est calculé par le nœud 1. Il doit être donc capable pour chacun de ces paquets, pour lequel il connaît le temps d'attente local $\mathcal{W}_{1,2}$, de calculer le temps d'attente correspondant sur le nœud suivant ($\mathcal{W}_{2,3}$). Le temps d'attente au nœud suivant va être approximé par la formule de récurrence. Pour le m-ième paquet du flux primaire noté $p^m_{f_1}$, la formule que le nœud 1 devrait appliquer pour obtenir une valeur exacte est :

$$\widehat{w}^1_{2,3}(p^m_{f_1}) = \max\{0, \widehat{w}^1_{2,3}(p^{m-1}_{f_1}) + s_{2,3}(p^{m-1}_{f_1}) - i_2(p^m_{f_1})\}\ , m > 1 \tag{4.6}$$

Mais le nœud 1 ne connaît que les paquets qui sortent de ses propres files (il ne connaît pas I_2). Il utilise donc plutôt les inter-arrivées $I^1_{2,1}$:

$$w^1_{2,3}(p^m_{f_1}) = \max\{0, w^1_{2,3}(p^{m-1}_{f_1}) + s_{2,3}(p^{m-1}_{f_1}) - i^1_{2,1}(p^m_{f_1})\}\ , m > 1 \tag{4.7}$$

A noter que dans le cas d'un seul flux, il n'y a pas de flux secondaires arrivant directement sur le nœud 2. Les quantités I_2 et $I^1_{2,1}$ coïncident donc, sauf si certains paquets sont perdus à l'arrivée sur le deuxième nœud.
L'estimateur de $\mathbb{E}\big[\mathbb{E}[\mathcal{W}_{1,2} \mid \mathcal{W}_{2,3}]\mathcal{W}_{2,3}\big]$ devient alors $\mathbb{E}\big[\mathbb{E}[\mathcal{W}_{1,2} \mid \mathcal{W}^1_{2,3}]\mathcal{W}_{2,3}\big]$.

4.4.2 Avec flux secondaire f_2

Contrairement au cas précédent, la présence des paquets du flux secondaire changent les temps d'attente au nœud 2. Ces temps d'attente supplémentaires ne peuvent pas être inférer par le nœud 1 puisqu'il ne voit pas passer les flux secondaires arrivant directement au nœud 2.
L'idée ici est, encore une fois, de regrouper les échantillons de temps d'attente par rapport aux résultats de la formule de récurrence. La quantité commune choisit sur les nœuds 1 et 2 pour classifier les paquets (la partie conditionnelle des espérances) est l'estimation du temps d'attente au nœud 2 par la formule de récurrence. Pour le nœud 1, la procédure est exactement la même que dans le cas "sans flux secondaire f_2". Il applique la formule donnée par l'équation (4.7). En conséquence, il obtient une estimation qui néglige complètement ce flux secondaire. Le nœud 2 classifie les temps d'attente locaux avec une formule de récurrence qui néglige aussi le flux f_2. Il connaît les inter-arrivées pour tous les flux. Il est donc capable de calculer la même fonction de récurrence que celle du nœud 1 (qui néglige le flux secondaire f_2). La formule appliquée par le nœud 2 est la suivante :

$$w^2_{2,3}(p^m_{f_1}) = \max\{0, w^2_{2,3}(p^{m-1}_{f_1}) + s_{2,3}(p^{m-1}_{f_1}) - i^2_{2,1}(p^m_{f_1})\}\ , m > 1 \tag{4.8}$$

D'un point de vue mathématique, l'estimateur du produit des délais locaux est calculée comme suit :

$$\mathbb{E}\big[\mathcal{W}_{1,2}\mathcal{W}_{2,3}\big] \approx \mathbb{E}\big[\mathbb{E}[\mathcal{W}_{1,2} \mid \mathcal{W}^1_{2,3}]\ \mathbb{E}[\mathcal{W}_{2,3} \mid \mathcal{W}^2_{2,3}]\big] \tag{4.9}$$

La formule est bien sûr une approximation. Les deux quantités/conditions $\mathcal{W}^1_{2,3}$ et $\mathcal{W}^2_{2,3}$ sont égales si il n'y a pas de pertes de paquets. Cela permet donc bien de regrouper les mêmes groupes de paquets sur les nœuds 1 et 2. Par contre, la valeur de ces deux estimateurs est différente du vraie temps d'attente sur le nœud 2, $W_{1,2}$, puisque le flux f_2 est complètement négligé. Nous donnons dans la section suivante, l'heuristique qui permet le calcul pratique de cette espérance.

4.4.3 Cas général : heuristique

Nous considérons désormais un scénario général avec $\mathcal{N}$ nœuds. Nous décrivons ici que le cas général avec la présence de flux secondaires. L'algorithme donné ici permet de calculer la covariance des temps d'attentes des nœuds k et $(k+1)$.
Nous partons donc de l'équation (4.9) et expliquons son calcul pratique. Afin de calculer l'espérance finale (qui regroupe les deux autres), les deux espérances conditionnelles sont approximées au travers de deux histogrammes. Le premier histogramme, calculé au niveau du nœud k, correspond à l'espérance $\mathbb{E}[\mathcal{W}_{k,k+1} \mid \mathcal{W}_{k+1,k+2}]$. L'abscisse de l'histogramme correspond aux intervalles des valeurs possibles de $\mathcal{W}^k_{k+1,k+2}$. Une des valeurs de l'histogramme, pour un intervalle donnée, correspond à la moyenne empirique des échantillons de $\mathcal{W}_{k,k+1}$ pour lesquels $\mathcal{W}^k_{k+1,k+2}$ tombent dans cet intervalle. L'autre histogramme est calculé par le nœud $(k+1)$ et correspond à l'espérance $\mathbb{E}[\mathcal{W}_{k+1,k+2} \mid \mathcal{W}^{k+1}_{k+1,k+2}]$.
Afin de permettre le rapprochement des deux histogrammes, le nombre et la taille des intervalles des abscisses de l'histogramme, correspondant à $\mathcal{W}^k_{k+1,k+2}$ et $\mathcal{W}^{k+1}_{k+1,k+2}$ ici, doivent être identiques. Soit L le nombre de ces intervalles (de colonnes) pour les deux histogrammes, nous notons Δ_i le i-ième intervalle $(1 \leq i \leq L)$.
Le calcul pratique de l'espérance conditionnelle se fait de la manière suivante.

Au nœud k. Le noeud k mesure le temps d'attente pour chaque paquet. Cette échantillon est noté $w_{k,k+1}(m)$ pour le m-ième paquet. Il calcule la quantité $w^k_{k+1,k+2}(m)$ correspondante en utilisant la fonction par récurrence (4.7).
Alors, l'histogramme pour le i-ième intervalle Δ_i est calculé comme :

$$\hat{\mathbb{E}}[\mathcal{W}_{k,k+1} \mid \mathcal{W}^k_{k+1,k+2}](i) \approx \frac{1}{c_i} \sum_{m=1}^{M} w_{k,k+1}(m) \mathbb{1}_{w^k_{k+1,k+2}(m) \in \Delta_i}. \tag{4.10}$$

où c_i est le nombre d'échantillons dans l'intervalle Δ_i, et M est le nombre total d'échantillons.

Au nœud $k+1$. Le nœud $(k+1)$ mesure le temps d'attente $w_{k+1,k+2}(m)$. Il calcule les quantités $w^{k+1}_{k+1,k+2}(m)$ en utilisant la fonction par récurrence (4.8) pour le m-ième paquet.
Alors, l'espérance conditionnelle pour le i-ième intervalle Δ_i est donnée par :

$$\hat{\mathbb{E}}[\mathcal{W}_{k+1,k+2} \mid \mathcal{W}^{k+1}_{k+1,k+2}](i) \approx \frac{1}{{c'}_i} \sum_{m=1}^{M'} w_{k+1,k+2}(m) \mathbb{1}_{w^{k+1}_{k+1,k+2}(m) \in \Delta_i}. \tag{4.11}$$

${c'}_i$ est le nombre d'échantillons dans l'intervalle Δ_i, et M' est le nombre total d'échantillons. Nous remarquons que la valeur ${c'}_i$ au nœud $(k+1)$ n'est pas forcement égal à c_i le nombre d'échantillons au nœud précédent.
Finalement, l'approximation finale consiste à regrouper les deux histogrammes intervalle par intervalle. L'équation obtenue est la suivante :

$$\begin{aligned} &\hat{\mathbb{E}}\big[\mathbb{E}[\mathcal{W}_{k,k+1} \mid \mathcal{W}^k_{k,k+1}]\ \mathbb{E}[\mathcal{W}_{k+1,k+2} \mid \mathcal{W}^{k+1}_{k+1,k+2}]\big] \\ &\qquad = \sum_{i=1}^{L} \frac{c_i}{M}\ \hat{\mathbb{E}}[\mathcal{W}_{k,k+1} \mid \mathcal{W}^k_{k+1,k+2}](i)\ \hat{\mathbb{E}}[\mathcal{W}_{k+1,k+2} \mid \mathcal{W}^{k+1}_{k+1,k+2}](i) \end{aligned} \tag{4.12}$$

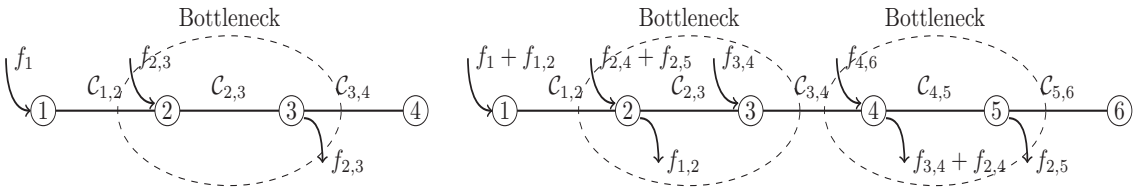
Notre méthode peut donc être résumée comme suit :

- **a)** Nous séparons le chemin emprunté par le flux primaire en groupes de deux nœuds consécutifs pour lesquels nous cherchons à calculer les covariances.
- **b)** Chaque paire de nœuds crée localement, idéalement pour exactement les mêmes groupes de paquets (grâce à la formule de récurrence qui permet de créer ces groupes), deux histogrammes de même tailles.
- **c)** Finalement, le contrôleur applique l'équation (4.12) pour rapprocher ces histogrammes et évaluer la covariance.

Dans la section suivante, nous effectuons un premier jeu de simulations permettant d'évaluer la précision numérique de cette méthode.

4.5 Premiers résultats numériques

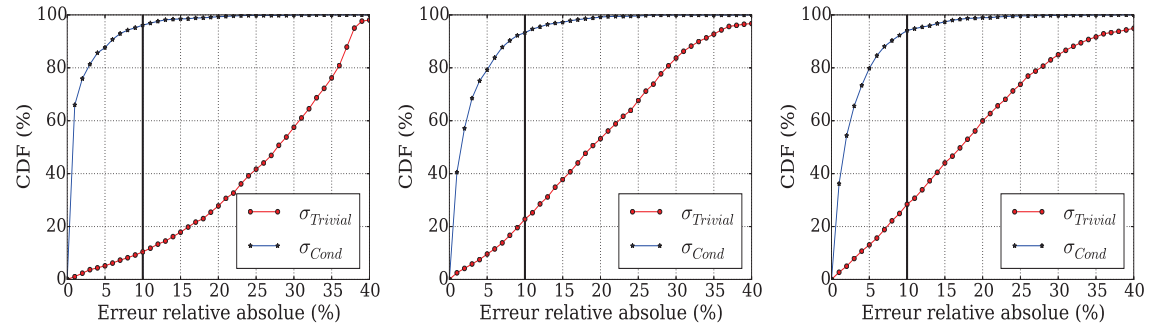
Comme pour le chapitre précédent, nous étudions les performances de la méthode en utilisant le simulateur NS-3. Nous réutilisons la trace provenant de la résidence de l'université de Stuttgart en Allemagne. Les détails de l'utilisation de cette trace pour générer les différents flux sont présentés au chapitre 5. Une estimation de la variance du délai de bout-en-bout est effectuée sur la base de 100 paquets. Les différentes courbes données ci-dessous (distribution de l'erreur, etc.) sont le résultat d'un grand nombre de ces



(a) Un scénario basic avec un goulot d'étranglement.

(b) Un scénario plus complexe avec deux goulots d'étranglements.

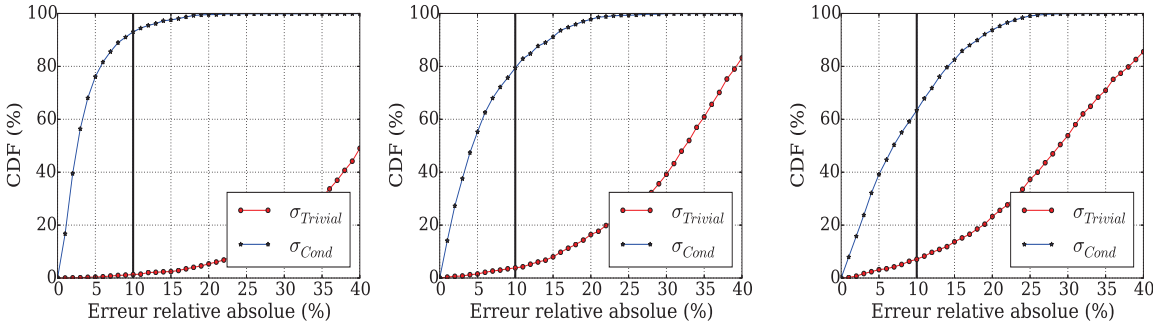
FIGURE 4.3 – Le flux primaire f_1 maintient un débit de 0.5, 1.0 et 1.5 Mbps. Dans la figure 4.3(a), les taux d'utilisations sur le premier goulot d'étranglement sont de 50%, 60% et 70%. Dans la figure 4.3(b), les taux d'utilisations sur le premier goulot d'étranglement sont de 37%, 50% et 63%, et de 32%, 50%, 66% sur le deuxième.



(a) Scénario A. Débit $f_1 = 0.5$ Mbps.

(b) Scénario A. Débit $f_1 = 1.0$ Mbps.

(c) Scénario A. Débit $f_1 = 1.5$ Mbps.



(d) Scénario B. Débit $f_1 = 0.5$ Mbps.

(e) Scénario B. Débit $f_1 = 1.0$ Mbps.

(f) Scénario B. Débit $f_1 = 1.5$ Mbps.

FIGURE 4.4 – Distribution des erreurs relative de l'écart-type du délai de bout-en-bout pour des charges différentes du flux primaire f_1.

estimations. Chaque charge du flux f_1 correspond à 1400 simulations différentes. Nous considérons les deux scénarios représentés sur la figure 4.3.

La comparaison est effectuée sur l'écart-type plutôt que sur la variance. La raison principale est que l'unité de l'écart-type est en secondes, donc identique à celle du délai de bout-en-bout contrairement à la variance exprimée en secondes carrées. Les valeurs sont donc plus facilement interprétables. Nous comparons la méthode conditionnelle aux deux autres méthodes.

La méthode exacte : Grâce à la simulation, nous n'avons pas de problème de synchronisation des horloges. En conséquence, le délai de bout-en-bout est simplement obtenu par la différence entre le temps de transmission d'un paquet par la source et le temps de son arrivée à la destination. Cette méthode offre donc un estimateur parfait et servira de

TABLE 4.1 – Distribution de l'erreur relative de l'écart-type du délai de bout-en-bout.

	Charge du f_1 : 0.5 Mbps				Charge du f_1 : 1 Mbps				Charge du f_1 : 1.5 Mbps			
Erreur	$\leq$ 10%	$\leq$ 20%	$>$ 30%	Moyen	$\leq$ 10%	$\leq$ 20%	$>$ 30%	Moyen	$\leq$ 10%	$\leq$ 20%	$>$ 30%	Moyen
$\sigma_{Cond-scen-A}$	**96.**11	**99.**31	**0.**07	**1.**84	**93.**26	**99.**17	**0.**07	**3.**05	**94.**01	**98.**96	**0.**21	**3.**19
$\sigma_{Cond-scen-B}$	**93.**06	**99.**58	**0.**14	**3.**80	**79.**40	**97.**77	**0.**07	**6.**02	**63.**37	**93.**75	**0.**07	**8.**47

référence.
La méthode triviale : Cette méthode est la méthode présentée au chapitre précédent. Elle néglige les corrélations.
Au niveau notations, nous utilisons $\sigma_{Trivial}$ pour indiquer la méthode triviale et σ_{Cond} pour la méthode conditionnelle.
Le premier scénario, scénario A, est représenté sur la figure 4.3(a). Les tailles de file d'attente sur tous les nœuds sont identiques et égalent à $\mathcal{K} = 15000$ octets (10 paquets de taille maximum). Dans ce scénario, les capacités des liens sont $C_{1,2} = 8, C_{2,3} = 5, C_{3,4} = 3$ Mbps. Le flux secondaire est fixé à 2 Mbps durant la simulation. Le flux primaire varie entre trois niveaux différents de charges : 0.5, 1.0, et 1.5 Mbps.
Les trois figures 4.4(a), 4.4(b), 4.4(c) montrent la distribution de l'erreur relative correspondant à la méthode triviale $\sigma_{Trivial}$ et la méthode conditionnelle σ_{Cond}.
Nous constatons clairement que l'approximation offerte par la méthode conditionnelle est meilleure que celle de la méthode triviale. Il apparaît également que la méthode conditionnelle fonctionne bien même dans le cas où le réseau est très saturé. Cependant, la précision est moins bonne lorsque le nombre de paquets perdus augmente (dû à la charge) comme illustré sur la figure 4.4(c).
Le deuxième scénario, scénario B, est beaucoup plus complexe. Il est représenté sur la figure 4.3(b). Les tailles des files d'attente sont restées identiques à celles du premier scénario $\mathcal{K} = 15000$. Les capacités des liens varient plus. Elles sont de $C_{1,2} = 6, C_{2,3} = 4, C_{3,4} = 7, C_{4,5} = 3, C_{5,6} = 5$ Mbps. Dans ce scénario, nous maintenons 5 flux secondaires de capacités $f_{1,2} = 1, f_{2,4} = 0.8, f_{2,5} = 0.2, f_{3,4} = 2$, et $f_{4,6} = 0.2$ Mbps. Le flux primaire est étudié à trois niveaux différents de charges : 0.5, 1.0, 1.5 Mbps.
Les résultats numériques dans ce scénario sont différents du premier scénario. Les résultats de la figure 4.4(d) ne montrent pas beaucoup de différences par rapport à ceux de la figure 4.4(a). Nous observons que la précision de la méthode conditionnelle de la figure 4.4(f) présente des résultats moins bons que celle de la figure 4.4(c). Il apparaît donc que les paquets perdus impactent plus les résultats si le chemin principal est long. Cette observation peut-être expliquée intuitivement : les erreurs d'approximation se cumulent lorsque le nombre de nœuds est plus élevé dans la chaîne.
La table 4.1 résume les résultats numériques. Nous pouvons retrouver les trois différents niveaux de charges du réseau. Nous donnons aussi la valeur moyenne de l'erreur relative en fonction des différents paramètres. Pour le scénario simple avec 4 nœuds et 3 liens, nous avons une erreur relative qui est en moyenne inférieure à 4%. Les résultats numériques pour le scénario B est moins bon mais reste néanmoins inférieure à 9%.

4.6 Conclusion

Dans ce chapitre, nous avons expliqué la difficulté d'évaluer la covariance des délais d'attente dans un contexte où les nœuds ne sont pas synchronisés et ne communiquent pas entre eux. Nous avons aussi proposé une méthode générale, où les temps d'attente sont

regroupés/classifiés au travers de conditions. Nous avons énumérés les propriétés que doit avoir cette classification : *classification*, *calculabilité* et *complexité*. Nous avons aussi proposé une classification particulière basée sur la formule de récurrence.
Grâce à certaines simplifications/hypothèses, nous obtenons une formulation de la récurrence qui permet à deux nœuds consécutifs de calculer et de classifier les temps d'attentes sur les mêmes critères. Ils en déduisent des espérances conditionnelles sous formes d'histogrammes qu'il est possible de rapprocher. Nous avons enfin présenté des résultats numériques pour deux scénarios particuliers. Les distributions de l'erreur et de l'erreur moyenne nous montrent que cette approximation fonctionne bien même lorsque le scénario est très compliqué/réaliste.
Il est maintenant nécessaire de comparer entre elles les deux méthodes proposées (méthode conditionnelle, et par décomposition), et de les évaluer au travers de scénarios plus riches.

Chapitre 5

Évaluation et Discussion

Contents

5.1 Introduction 45
5.2 Configuration 45
5.2.1 Simulateur NS-3 45
5.2.2 Scénario 46
5.3 Trace réseau réaliste 46
5.3.1 Taille des paquets 46
5.3.2 Arrivée des paquets 47
5.3.3 Sélection des flux 48
5.4 Impact du nombre des paquets 49
5.4.1 Un très grand nombre de paquets observés 49
5.4.2 Un petit nombre de paquets observés 49
5.4.3 Nombre variable de paquets observés 53
5.5 Impact de la taille du tampon 55
5.5.1 500 paquets observés 56
5.5.2 1000 paquets observés 57
5.6 Impact de l'intervalle de temps 59
5.6.1 Trace de l'Université de Stuttgart 59
5.6.2 Trace CAIDA 60
5.7 Conclusions 60

Résumé : *Ce chapitre présente une évaluation et une composition des différentes méthodes proposés. Après une présentation du simulateur et des scénarios, nous comparons les méthodes en faisant varier différentes quantités : **a)** le nombre des paquets traités **b)** la taille des tampons **c)** le nombre d'intervalle des histogrammes (méthode conditionnelle) **d)** ou l'indépendance des flux.*

.

5.1 Introduction

Comme nous l'avons vu dans les chapitres précédents, il est très difficile pour un nœud de connaitre localement le temps d'attente exact sur un autre nœud du réseau. Par conséquence, nous savons que la covariance des temps d'attente entre deux nœuds, indiquée ci-dessous, est très difficile à calculer :

$$\mathrm{Cov}(\mathcal{W}_l, \mathcal{W}_k) = \mathbb{E}[\mathcal{W}_l \mathcal{W}_k] - \mathbb{E}[\mathcal{W}_l]\mathbb{E}[\mathcal{W}_k], \text{ avec } l \neq k.$$

À cause de cette difficulté, nous avons proposé deux méthodes pour approcher ce paramètre de covariance. La première méthode est la méthode par décomposition qui utilise une fonction par récurrence pour estimer le temps d'attente aux nœuds suivants. La valeur calculée n'est pas obligatoirement égale à la valeur exacte mais est par contre une valeur corrélée avec le temps d'attente au nœud actuel. La deuxième méthode est une autre approche qui essaye d'apprécier des groupes de temps d'attente et puis fait un produit entre eux pour obtenir une approximation de la covariance.
La précision de ces deux solutions va évidemment dépendre de plusieurs facteurs, comme, par exemple, la nature des flux, le taux de perte des paquets, etc. Ce chapitre s'intéresse à la précision des solutions proposées dans la thèse. Les résultats numériques obtenus sur la précision nous permettent de soulever des discussions sur des aspects cachés qui n'apparaissaient pas dans les chapitres précédents, ainsi que de mettre en avant des points forts et des points faibles pour ces deux méthodes.

5.2 Configuration

5.2.1 Simulateur NS-3

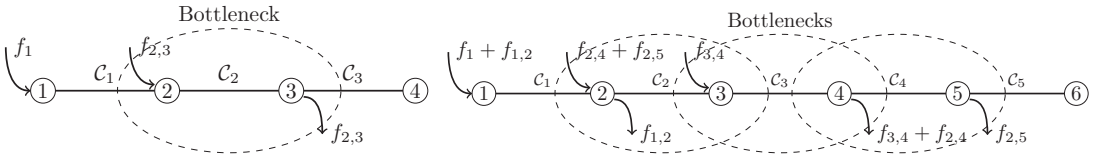
Pour examiner les performances des méthodes proposées dans le contexte de Software-Defined Networking, il est possible d'utiliser :

1. Soit un simulateur réseau comme NS-3 [19] (Network Simulator version 3) ou, éventuellement, NS-2 [21] (Network Simulator version 2).
2. Soit un émulateur reposant sur un réseau virtuel en utilisant, par exemple, Mininet [24] ou sur des machines virtuelles en utilisant, par exemple, VMware [39].
3. Soit une expérience avec de vrais routeurs.

Le premier choix est une option intéressante car les simulateurs, comme NS-2 et NS-3, sont des logiciels ouverts. Il est possible d'implanter directement nos méthodes dans ces simulateurs. De plus, NS-3 a la possibilité de simuler l'approche SDN. Le deuxième choix est aussi une possibilité mais nous devons payer pour certains logiciels, comme c'est le cas, par exemple, pour VMware. De plus, une machine virtuelle peut prendre beaucoup de ressources comme de la RAM (Random Access Memory) ou du CPU (Center Process Unit). Développer les solutions proposées dans un émulateur peut aussi prendre du temps. Enfin, le troisième choix peut être cher et très difficile à mettre en œuvre.
Après avoir réfléchi à ces différentes possibilités, le simulateur NS-3 nous a paru suffisant par rapport à nos contraintes. Le simulateur NS-3 est un logiciel ouvert, il a été développé avec le langage orienté objet C++, et il est relativement facile d'implanter nos méthodes dans ce simulateur. Il est aussi intéressant car il a déjà un module de Software-Defined Networking intégré.

5.2.2 Scénario

La figure 5.1 montre deux scénarios que nous avons considérés dans ce travail. Ce sont deux scénarios assez généraux si nous nous intéressons à un chemin dans un réseau. En



(a) Petit scénario : 2 flux et 4 nœuds. (b) Grand scénario : 5 flux et 6 nœuds.

FIGURE 5.1 – Scénarios considérés.

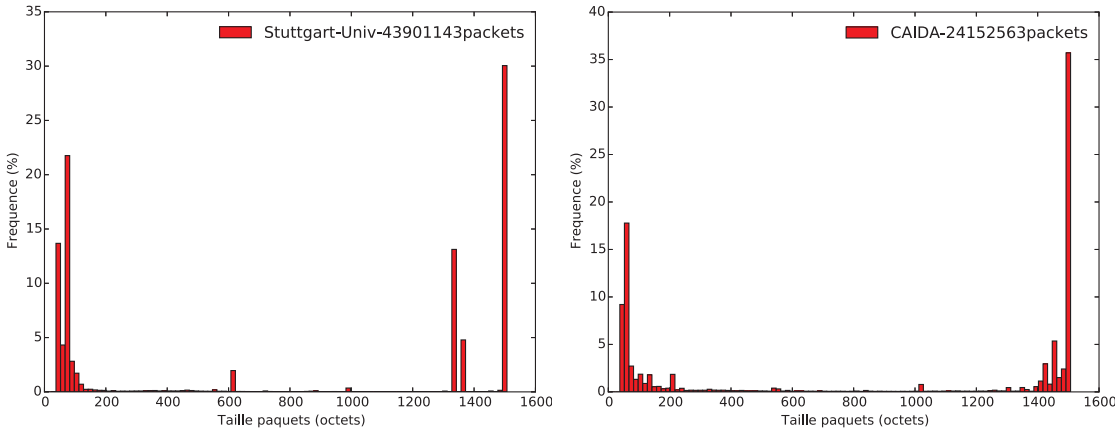
effet, un chemin est une topologie élémentaire dès qu'on s'intéresse à un flux de bout-en-bout dans un réseau. En fait, un réseau complexe est essentiellement composé de plusieurs sous-réseaux comme ceux de la figure 5.1.
La figure 5.1(a) présente un scénario composé de 4 nœuds, de 3 liens et de 2 flux. Le flux primaire, f_1, entre sur le nœud 1 et sort sur le nœud 4. Le flux secondaire, $f_{2,3}$, entre sur le nœud 2 et sort sur le nœud 3. Par la suite, nous appellerons le petit scénario. La figure 5.1(b) présente un scénario composé de 6 nœuds, de 5 liens et de 5 flux. Le flux primaire, f_1, entre sur le nœud 1 et sort sur le nœud 6. Les flux secondaires entrent et sortent sur différents nœuds de la chaîne. Par la suite, nous appellerons le grand scénario.
Au niveau de la capacité des liens, nous considérons deux cas : i) des capacités modérées exprimées en Mbps (Megabits per second) et des capacités élevées exprimées en Gbps (Gigabits per second). Un réseau avec des capacités modérées représente plutôt un réseau local, par exemple un réseau de résidence d'une université. Un réseau avec des capacités élevées représente, par exemple, un réseau de cœur ou un réseau d'une grande entreprise. Concernant la taille du tampon permettant de stocker les paquets en attente pour la transmission sur un lien de communication, nous choisissons une taille identique pour tous les nœuds dans un scénario donné. Sauf pour la section 5.5 où nous étudions l'impact de la taille du tampon sur la précision. Dans les autres sections, nous fixons la taille du tampon à 15000 octets (ceci est équivalent à 10 paquets de 1500 octets) pour un réseau avec des capacités modérées et 75000 octets (ceci est équivalent à 50 paquets de 1500 octets) pour un réseau avec des capacités élevées.

5.3 Trace réseau réaliste

Pour évaluer nos méthodes, nous utilisons deux traces réelles : une des traces provient d'une résidence de l'Université de Stuttgart en Allemagne [41], tandis que l'autre trace est disponible sur le site CAIDA [2, 3] (Center for Applied Internet Data Analysis). Ce dernier fournit une collection de plusieurs traces (136 fichiers de traces en 2014, 126 fichiers de traces en 2015). Pour nos simulations, nous avons choisi 5 fichiers parmi cette collection.

5.3.1 Taille des paquets

La figure 5.2 montre la distribution des tailles de paquets pour une sous-partie de chacune des deux traces. La figure 5.2(a) correspond à la trace de la résidence de l'Université de Stuttgart à partir de 18h00 et ce pendant quatre heures, en Octobre 2004. La figure 5.2(b) correspond à la trace CAIDA à partir de 13h33 et ce pendant une minute le 17 septembre



(a) Trace d'une résidence de l'Université de Stuttgart.

(b) Un morceau de trace CAIDA.

FIGURE 5.2 – Distribution des tailles de paquets.

2015. Nous voyons que la taille des paquets varie fortement puisque les valeurs sont comprises entre 28 octets et 1500 octets.
La précision de nos méthodes vont dépendre de la taille des paquets. Il est donc important de prendre, dans nos simulations, des tailles de paquets réalistes. C'est ce que nous faisons en utilisant des traces réelles pour simuler les flux qui vont traverser le réseau.

5.3.2 Arrivée des paquets

La trace brute (i.e. récupérée du site web concernée) est pré-traitée de telle sorte que seules les arrivées et les tailles des paquets sont fournies au simulateur NS-3.

Un exemple de trace fournie à NS-3

```
18 00 00.000115376 104
18 00 00.000127088 203
18 00 00.000146672 96
18 00 00.000327088 1500
18 00 00.000450480 104
```

Dans l'exemple ci-dessus, les trois premières colonnes montrent l'heure, les minutes et les secondes des arrivées, tandis que la dernière colonne montre la taille des paquets. Afin de pouvoir simuler des débits élevés, les secondes sont indiquées avec une précision de l'ordre de la nano-seconde.
Dans chaque simulation, nous considérons deux types de paquet :

1. Les paquets observés sur lesquels nos algorithmes se basent. Sur ces paquets, nous allons calculer la variance du délai de bout-en-bout en mesurant la taille de ces paquets, le temps d'attente, etc.
2. Les paquets de secours. À cause d'événements inévitables correspondant à des paquets perdus, certains nœuds auront moins de paquets à observer que d'autres. Pour faire en sorte que tous les nœuds aient un nombre de points de mesures égal, on considère un morceau de trace suffisamment grand. Cela permet d'intégrer des paquets de secours dans les mesures sur certains nœuds où des paquets observés sont manquants.

La figure 5.3 montre qu'un morceau de trace est alors découpé en un ensemble de paquets à observer pour nos solutions et un ensemble de paquets de secours. Très souvent, chaque morceau de trace est constitué de 100 paquets observés et de 1000 paquets de secours, sauf dans la section 5.4 où nous étudions l'impact du nombre de paquets observés.

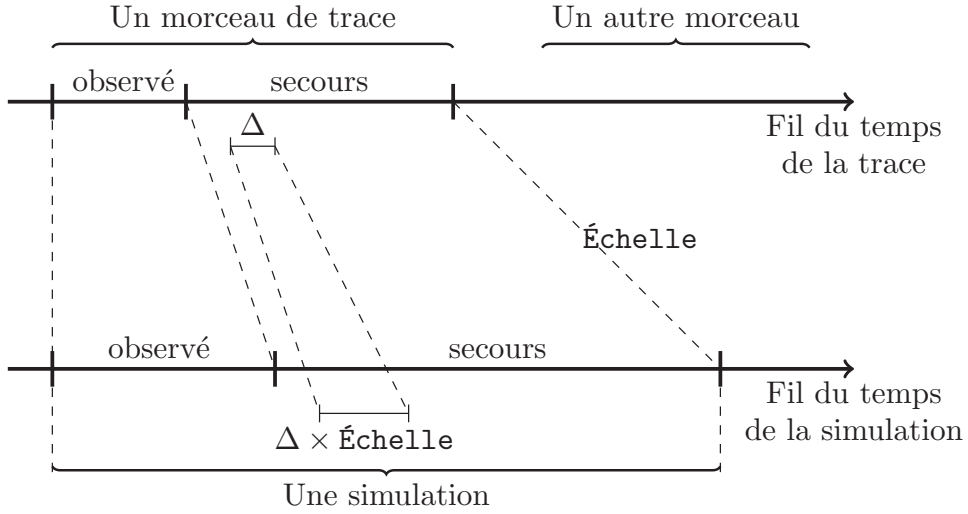


FIGURE 5.3 – Un morceau de trace et la simulation.

La trace brute va imposer un débit d'injection des paquets dans le réseau. Soit Y (exprimé en Mbps) le débit moyen de la trace. Pour pouvoir tester l'impact du débit des flux sur les solutions proposées, et donc tester les solutions sous n'importe quel débit donné, nous régulons le débit de la trace. Si X (exprimé en Mbps) est le débit moyen à fixer, alors il suffit de modifier les inter-arrivées Δ des paquets de la trace de la façon suivante : $\Delta \times \texttt{Échelle}$ avec $\texttt{Échelle} = \frac{\mathsf{X}}{\mathsf{Y}}$. La figure 5.3 montre comment la trace initiale est modifiée pour obtenir un débit spécifique d'injection des paquets.

5.3.3 Sélection des flux dans la trace

Dans une simulation nécessitant de multiples flux, nous assignons à chaque flux un morceau de trace différent. Le flux primaire démarre sur le début d'un morceau de trace et termine avec la fin du morceau de trace. Les flux secondaires sont choisis arbitrairement de telle sorte que les morceaux de trace choisis pour ces flux soient *assez* éloignés les uns des autres pour éviter la dépendance entre les flux.

Une sélection arbitraire des morceaux de traces : un exemple de code *python*

```
for hh in xrange (18, 22, 1) : # L'heure pour le flux primaire
  for mm in xrange (0, 60, 1) : # La minute pour le flux primaire
    for ss in xrange (0, 60, 10) : # La seconde pour le flux primaire
    # Le morceau de trace pour le deuxième flux est basé sur le flux primaire
    hh2 = 18
    mm2 = (mm + 20) % 60 # choix arbitraire
    ss2 = (ss + 10) % 60
    # Le morceau de trace pour le troisième flux est basé sur le deuxième flux
    hh3 = 19
    mm3 = (mm2 + 20) % 60
    ss3 = (ss2 + 10) % 60
```

Le code en *python* ci-dessus montre un exemple de sélection des flux.

5.4 Impact du nombre de paquets observés

La méthode par décomposition est basée sur une moyenne empirique. Plus nous avons de paquets observés, plus la variance calculée a des chances de s'approcher de la vraie variance du délai de bout-en-bout. En revanche, plus de paquets nécessaires à l'observation impliquent un temps de mesures plus long et donc un délai plus important dans l'obtention de la variance, ce qui peut être un problème dans les réseaux où la dynamique est importante au niveau des délais.
Par ailleurs, la méthode par l'espérance conditionnelle utilise la fonction par récurrence comme une condition. Cette condition divise les paquets en groupes. Les paquets dans un même groupe ont des temps d'attente proches. Nous remarquons que cette condition marche bien quand on a peu des paquets observés et devient pire quand le nombre des paquets observés augmente.

5.4.1 Un très grand nombre de paquets observés

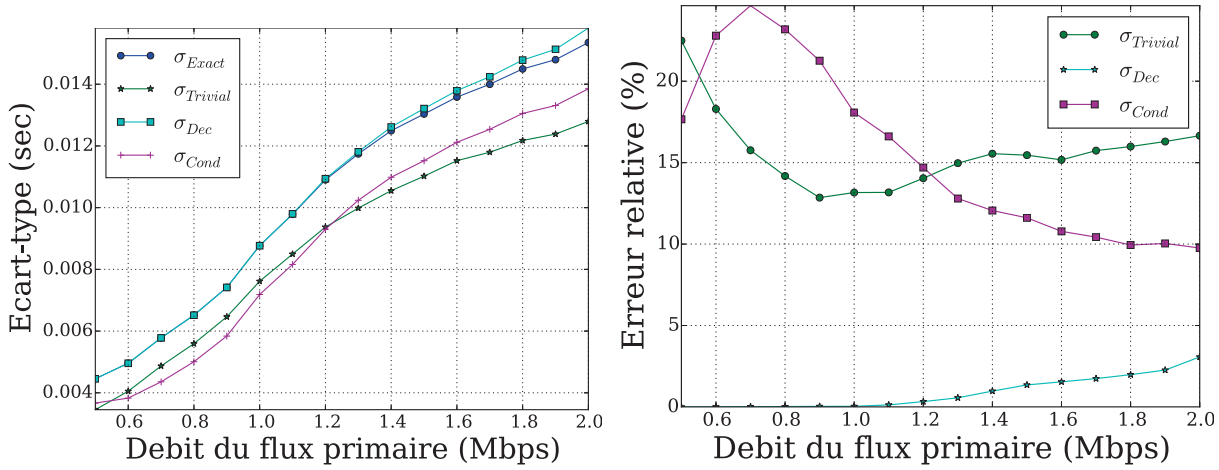
Les figures 5.4 et 5.5 montrent les résultats obtenus sur l'écart-type du délai de bout-en-bout pour le flux primaire et l'erreur relative en fonction du débit du flux primaire. L'erreur relative est calculée comme étant le ratio de la différence sur l'écart-type obtenu avec une des solutions et l'écart-type obtenu avec la méthode exacte sur l'écart-type de la méthode exacte ($\frac{|\sigma_{Solution}-\sigma_{Exact}|}{\sigma_{Exact}}$). σ_{Dec} correspond à la solution par décomposition tandis que σ_{Cond} correspond à la méthode par décomposition. $\sigma_{Trivial}$ correspond à la méthode triviale qui somme les écarts-types calculés localement sur chaque nœud (les termes de covariance sont donc négligés). σ_{Exact} correspond aux valeurs obtenues par la méthode exacte qui se base sur les délais de bout-en-bout exacts des paquets et qui peuvent être obtenus grâce à la simulation qui dispose d'une horloge globale. Chaque point correspond à une simulation avec 100100 paquets observés et 100000 paquets de secours. La figure 5.4 correspond aux résultats obtenus sur un petit scénario (figure 5.1(a)), tandis que la figure 5.5 correspond aux résultats obtenus sur un grand scénario (figure 5.1(b)).
Nous voyons que la méthode par décomposition est proche de la valeur exacte de l'écart-type, l'erreur relative étant inférieure à 5% pour le petit scénario et très souvent inférieure à 5% sur le grand scénario. A l'inverse, la méthode par l'espérance conditionnelle donne, très souvent, les plus mauvais résultats. Deux causes possibles sont des paquets perdus et l'absence de la covariance entre des nœuds plus loin qu'un saut.

5.4.2 Un petit nombre de paquets observés

Cette sous-section montre des résultats numériques avec 100 paquets observés. Comme les résultats varient plus fortement selon les morceaux de trace choisis qu'avec 100100 paquets observés, nous présentons différemment les résultats obtenus.

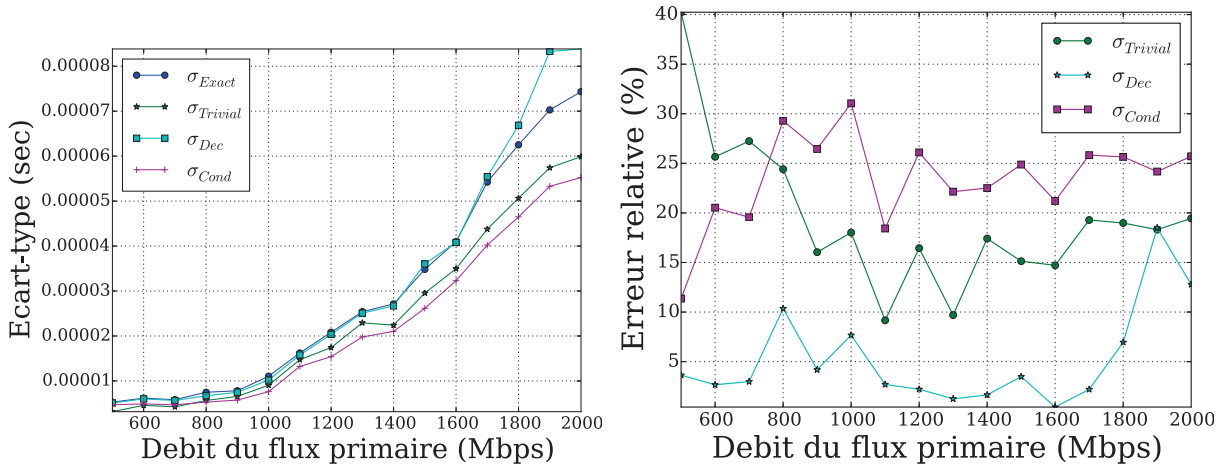
Trace allemande

La figure 5.6 montre, pour un petit scénario et avec la trace allemande, la moyenne sur l'écart-type avec un intervalle de confiance de 95% ainsi que la médiane sur l'écart-type. Les résultats ont été obtenus à partir 1400 simulations. Les résultats montrent que la méthode par l'espérance conditionnelle est plus résistante à la perte que la méthode par décomposition en ce qui concerne la moyenne sur l'écart-type. En revanche, la méthode



(a) Écart-type. (b) Erreur relative.

FIGURE 5.4 – L'écart-type et l'erreur relative sur un petit scénario 5.1(a) avec $C_{1,2} = 8, C_{2,3} = 5, C_{3,4} = 3$ (Mbps). La trace utilisée est la trace de l'Université de Stuttgart. Le débit du flux primaire varie de 0.5 Mbps à 2 Mbps alors que le débit du flux secondaire est fixé à $f_{2,3} = 2$ (Mbps).

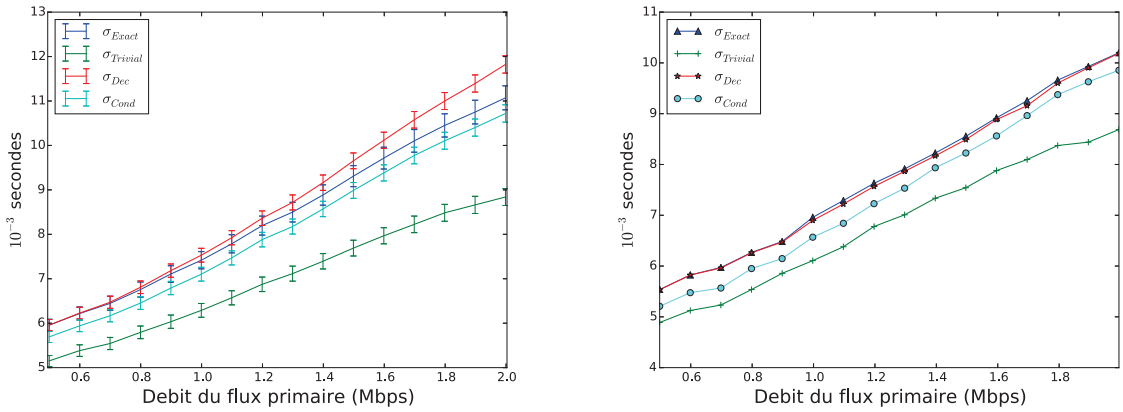


(a) Écart-type. (b) Erreur relative.

FIGURE 5.5 – L'écart-type et l'erreur relative sur un grand scénario 5.1(b) avec $C_{1,2} = 6, C_{2,3} = 4, C_{3,4} = 7, C_{4,5} = 3, C_{5,6} = 5$ (Gbps). La trace utilisée est la trace CAIDA. Le débit du flux primaire varie de 500 Mbps à 2 Gbps alors que les débits des flux secondaires sont fixés à $f_{1,2} = 1000, f_{2,5} = 200, f_{2,4} = 800, f_{3,4} = 1500$ (Mbps).

par décomposition est plus proche de la solution exacte quand on s'intéresse à la médiane. Ceci s'explique par le fait que certaines simulations donnent, pour la méthode par décomposition, de mauvais résultats qui impactent fortement la moyenne. Mais ces mauvais résultats étant en petit nombre, ils ont une influence beaucoup plus limitée sur la médiane.

La figure 5.7 donne la fonction de répartition de l'erreur relative sur l'écart-type pour le même scénario que la figure 5.6. Le débit du flux primaire est fixé à 1 Mbps ou à 1,5 Mbps tandis que le débit du flux secondaire est fixé à 2 Mbps. Les résultats montrent que la méthode par décomposition obtient de très bons résultats pour ces scénarios. Pour les deux valeurs de débit pour le flux primaire, l'erreur relative pour cette méthode est

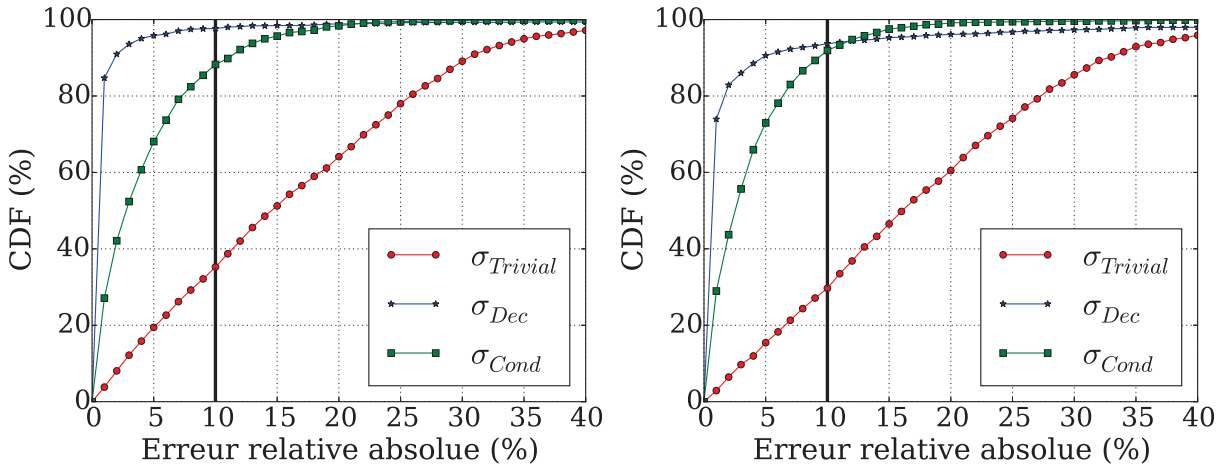


(a) Moyenne et intervalle de confiance.

(b) Médiane.

FIGURE 5.6 – Ecart-type sur un petit scénario 5.1(a) avec $C_{1,2} = 8, C_{2,3} = 5, C_{3,4} = 3$ (Mbps). Le débit du flux secondaire est fixé à 2 Mbps.

inférieure à 10% pour au moins 98% des cas. La méthode par l'espérance conditionnelle présente des résultats moins bons mais qui restent néanmoins très acceptables : l'erreur relative pour cette méthode est inférieure à 10% pour presque 90% des cas observés.



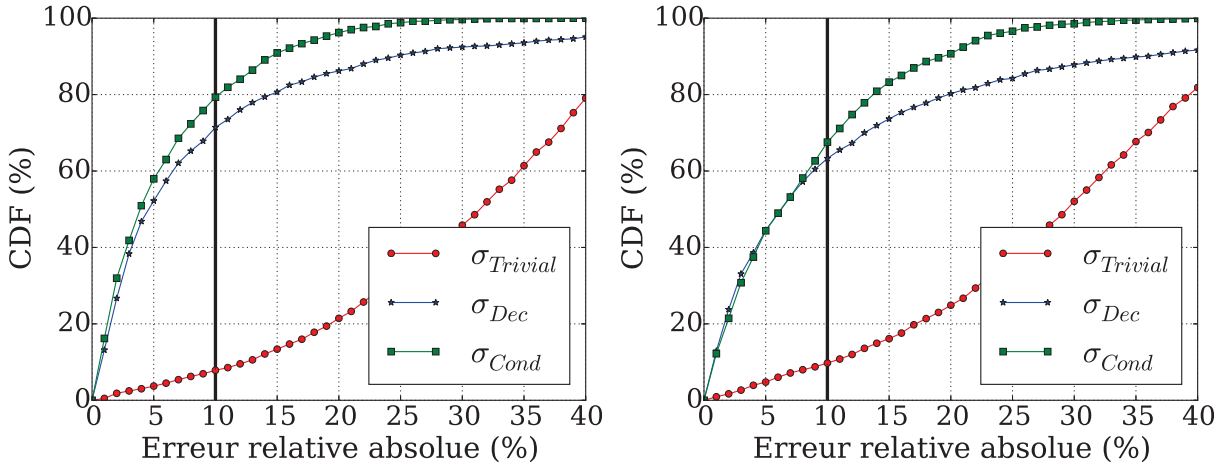
(a) $f_1 = 1.0$ (Mbps).

(b) $f_1 = 1.5$ (Mbps).

FIGURE 5.7 – La fonction de répartition de l'erreur relative pour un petit scénario 5.1(a) avec $C_{1,2} = 8, C_{2,3} = 5, C_{3,4} = 3$ (Mbps). Le débit du flux primaire est fixé à 1 Mbps ou à 1,5 Mbps tandis que le débit du flux secondaire est fixé à 2 Mbps.

La figure 5.8 présente la fonction de répartition de l'erreur relative sur un grand scénario. Les résultats obtenus sont un peu différents de ceux donnés dans la figure 5.7. Tout d'abord, les résultats sont un peu moins bons pour les deux méthodes proposées. Les deux méthodes ont aussi des résultats plus proches, bien que la méthode par décomposition soit un peu meilleure dans ce cas. Lorsque le débit du flux primaire est fixé à 1 Mbps, l'erreur relative de la méthode par décomposition est inférieure à 10% dans 80% des cas, tandis que c'est le cas dans 65% des cas lorsque le débit du flux primaire passe à 1,5 Mbps. Une des raisons possibles pourrait provenir du nombre de paquets observés qui ne sont peut-être pas suffisant pour un scénario complexe comme celui de la figure 5.1(b). Une autre raison possible pourrait provenir de la dépendance des flux qui existerait dans cette

trace allemande.

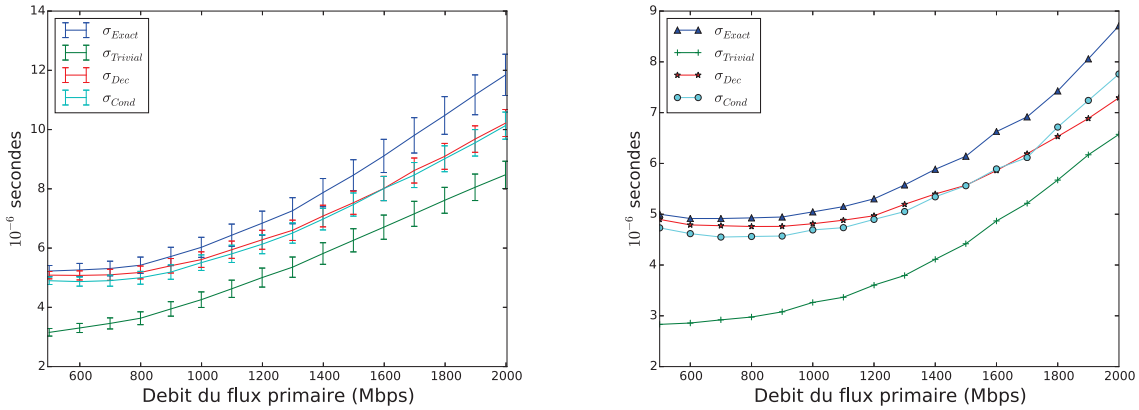


(a) $f_1 = 1.0$ (Mbps). (b) $f_1 = 1.5$ (Mbps).

FIGURE 5.8 – La fonction de répartition de l'erreur relative sur un grand scénario 5.1(b) avec $C_{1,2} = 6, C_{2,3} = 4, C_{3,4} = 7, C_{4,5} = 3, C_{5,6} = 5$ (Mbps). Le débit du flux primaire est fixé à 1 Mbps ou à 1,5 Mbps alors que les débits des flux secondaires sont fixés à : $f_{1,2} = 1; f_{2,5} = 0,2; f_{2,4} = 0,8; f_{3,4} = 1,5$ (Mbps).

Trace CAIDA

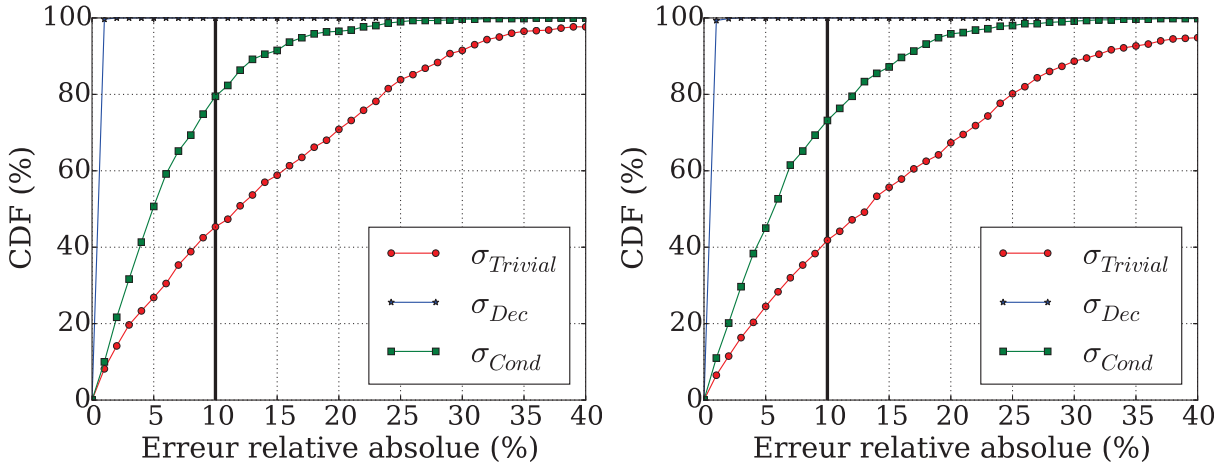
La figure 5.9 montre, pour un petit scénario et avec la trace CAIDA, la moyenne sur l'écart-type avec un intervalle de confiance de 95% ainsi que la médiane sur l'écart-type. Les résultats ont été obtenus à partir de 600 simulations. La figure montre que les deux méthodes obtiennent des résultats très proches pour la moyenne et pour la médiane, la méthode par décomposition étant un peu meilleure quand le débit du flux primaire n'est pas trop élevé. La trace CAIDA contient beaucoup plus de grands paquets que la trace allemande, c'est pourquoi la méthode par décomposition ne fonctionne pas dans le cas de haut débit (impliquant plus de pertes).



(a) Moyenne et intervalle de confiance. (b) Médiane.

FIGURE 5.9 – Ecart-type pour un petit scénario 5.1(a) avec $C_{1,2} = 8, C_{2,3} = 5, C_{3,4} = 3$ (Gbps). Le débit du flux secondaire est fixé à 2 Gbps.

La figure 5.10 donne la fonction de répartition de l'erreur relative sur l'écart-type pour le même scénario que la figure 5.9. Le débit du flux primaire est fixé à 1 Gbps ou à 1,5 Gbps tandis que le débit du flux secondaire est fixé à 2 Gbps. Comme pour la trace allemande, les résultats montrent que, sur un petit scénario, la méthode par décomposition obtient de très bons résultats. Pour les deux valeurs de débit pour le flux primaire, l'erreur relative pour cette méthode est inférieure à 10% pour presque 100% des cas. La méthode par l'espérance conditionnelle présente des résultats moins bons mais qui restent néanmoins très acceptables : l'erreur relative pour cette méthode est inférieure à 10% pour 85% des cas observés.



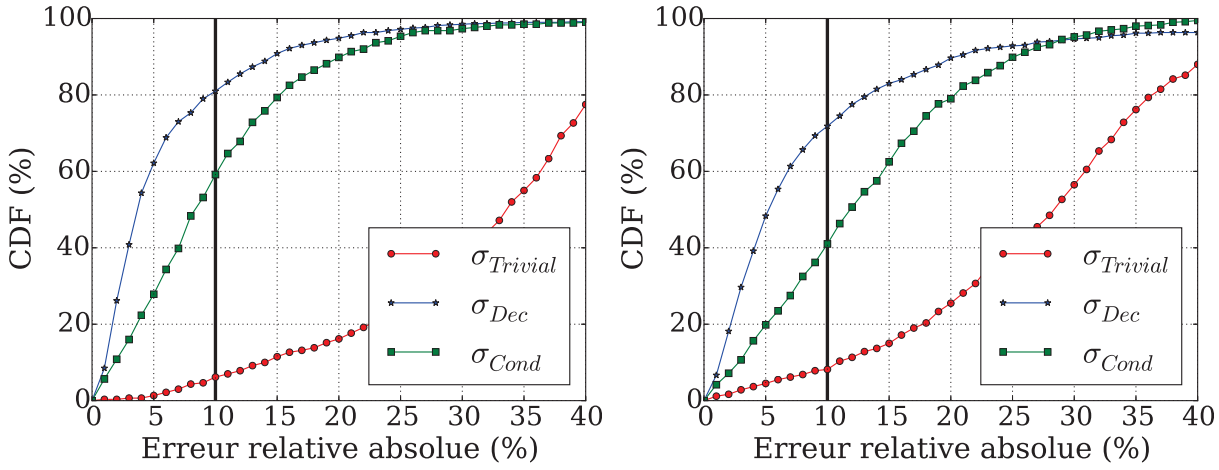
(a) $f_1 = 1.0$ (Gbps). (b) $f_1 = 1.5$ (Gbps).

FIGURE 5.10 – La fonction de répartition de l'erreur relative pour un petit scénario 5.1(a) avec $C_{1,2} = 8, C_{2,3} = 5, C_{3,4} = 3$ (Gbps). Le débit du flux primaire est fixé à 1 Gbps ou à 1,5 Gbps tandis que le débit du flux secondaire est fixé à 2 Gbps.

La figure 5.11 présente la fonction de répartition de l'erreur relative sur un grand scénario. Les résultats obtenus sont moins bons que ceux donnés dans la figure 5.10. Dans ce scénario, la méthode par décomposition reste meilleure que la méthode par l'espérance conditionnelle. Lorsque le débit du flux primaire est fixé à 1 Gbps, l'erreur relative de la méthode par décomposition est inférieure à 10% dans 80% des cas, tandis que c'est le cas dans 70% des cas lorsque le débit du flux primaire passe à 1,5 Gbps. La méthode par l'espérance conditionnelle donne des résultats beaucoup plus faibles puisque seulement 60% des cas dans la figure 5.11(a) et 40% des cas dans la figure 5.11(b) ont une erreur relative inférieure à 10%.

5.4.3 Nombre variable de paquets observés

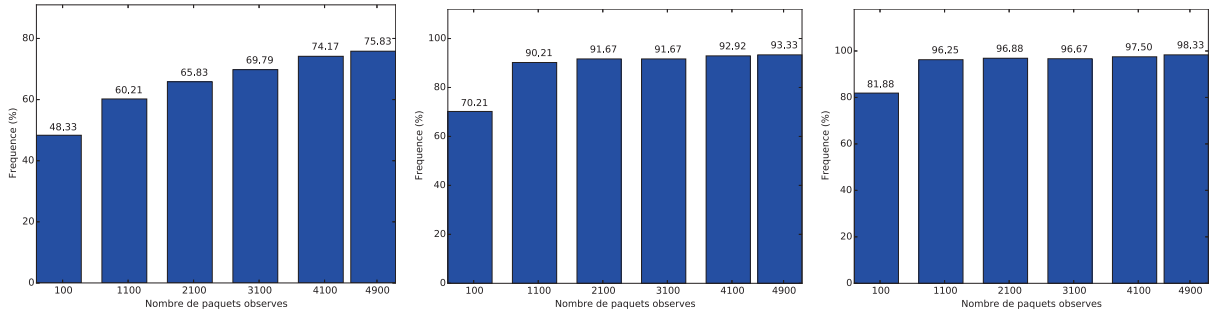
Dans cette sous-section, nous allons étudier l'impact du nombre des paquets observés en s'intéressant à des valeurs intermédiaires entre 100 et 100100 paquets observés. Ces simulations cherchent notamment à répondre à la question suivante : *Quel est le nombre nécessaire et suffisant de paquets observés pour bien estimer l'écart-type du délai de bout-en-bout ?*
La trace utilisée est la trace CAIDA. Elle est divisée en environ 500 morceaux différents. À chaque morceau de trace, nous allons exécuter des simulations avec des valeurs différentes sur le nombre de paquets observés. Le scénario testé est un grand scénario (figure 5.1(b)) avec des capacités de liens fixés à : $C_{1,2} = 6, C_{2,3} = 4, C_{3,4} = 7, C_{4,5} = 3, C_{5,6} = 5$ (Gbps). Les débits du flux primaire et des flux secondaires sont fixés à $f_1 = 1,5, f_{1,2} =$



(a) $f_1 = 1.0$ (Gbps). (b) $f_1 = 1.5$ (Gbps).

FIGURE 5.11 – La fonction de répartition de l'erreur relative sur un grand scénario 5.1(b) avec $C_{1,2} = 6, C_{2,3} = 4, C_{3,4} = 7, C_{4,5} = 3, C_{5,6} = 5$ (Gbps). Le débit du flux primaire est fixé à 1 Gbps et à 1,5 Gbps tandis que les débits des flux secondaires sont fixés à : $f_{1,2} = 1,0; f_{2,5} = 0,2; f_{2,4} = 0,8; f_{3,4} = 1,5$ (Gbps).

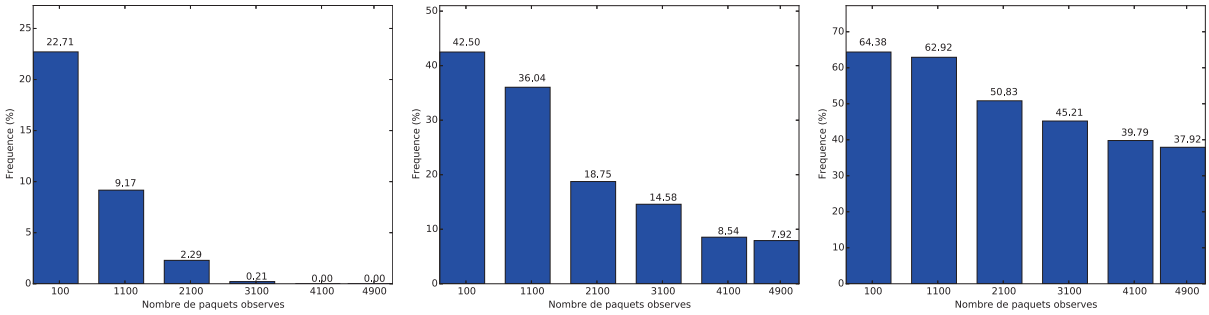
$1,0, f_{2,5} = 0,2, f_{2,4} = 0,8, f_{3,4} = 1,5$ (Gbps). La taille du tampon est fixée à 75000 octets (correspondant à environ 50 paquets en moyenne).



(a) 5 %. (b) 10 %. (c) 15 %.

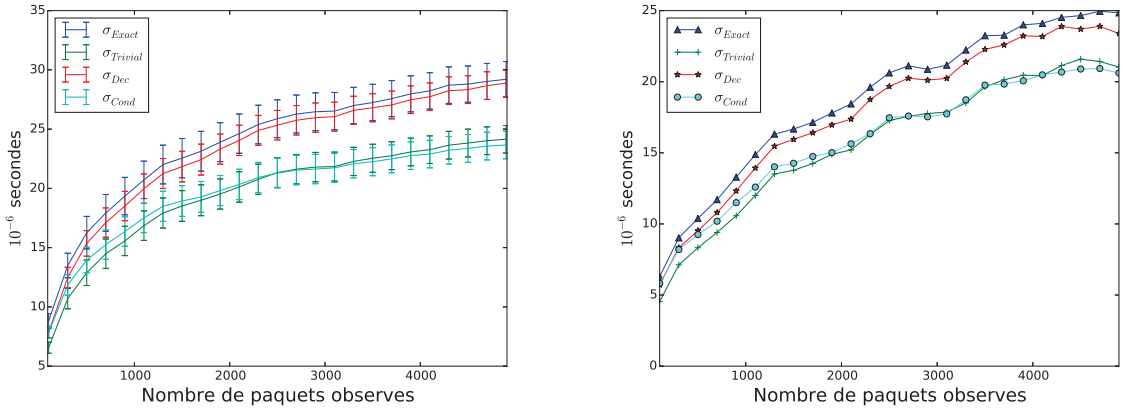
FIGURE 5.12 – Pourcentage de cas où l'erreur relative, pour la méthode par décomposition en fonction du nombre de paquets observés, est inférieure à : 5% (figure 5.12(a)), 10% (figure 5.12(b)) et 15% (figure 5.12(c)).

La figure 5.12 indique le pourcentage de cas présentant une erreur relative inférieure à un seuil donné pour la méthode par décomposition et en fonction du nombre de paquets observés. Comme attendu, cette méthode marche mieux si on a plus de paquets observés. On peut aussi observer que selon la précision attendue, il n'est pas nécessaire d'observer le même nombre de paquets. Par exemple un nombre de 5000 paquets ne suffit pas si nous avons besoin qu'au moins 80% des cas aient une erreur relative inférieure à 5%, tandis que 100 paquets sont suffisants si la précision attendue est seulement de 15%.
La figure 5.13 montre que la méthode par l'espérance conditionnelle fonctionne mieux si peu de paquets sont observés. La raison vient de la fonction par récurrence, qui peut donner de mauvaises prédictions quand nous avons plus de paquets à observer. Nous pouvons aussi noter que la méthode par l'espérance conditionnelle présente moins de cas favorables que la méthode par décomposition pour ce scénario.
Enfin, la figure 5.14 donne la moyenne de l'erreur relative avec un intervalle de confiance



(a) 5 %. (b) 10 %. (c) 15 %.

FIGURE 5.13 – Pourcentage de cas où l'erreur relative, pour la méthode par l'espérance conditionnelle en fonction du nombre de paquets observés, est inférieure à : 5% (figure 5.13(a)), 10% (figure 5.13(b)) et 15% (figure 5.13(c)).



(a) Moyenne intervalle de confiance. (b) Médiane.

FIGURE 5.14 – Moyenne et médiane de l'écart-type sur un grand scénario.

de 95%, ainsi que la médiane de l'écart-type en fonction du nombre de paquets observés. Le scénario est identique à celui des figures 5.12 et 5.13. On voit que, pour ce scénario, la méthode par décomposition donne clairement des résultats plus proches de la solution exacte que la méthode par l'espérance conditionnelle.

5.5 Impact de la taille du tampon

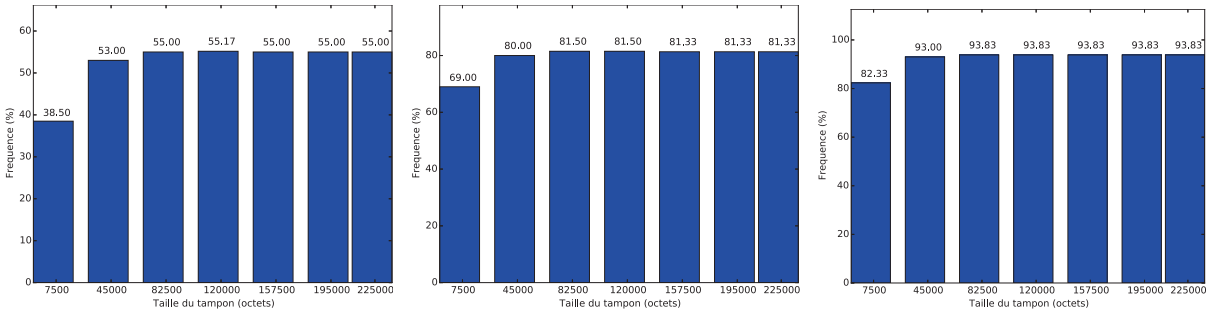
Il y a une relation évidente entre la taille du tampon et la perte des paquets. Or, cette perte de paquets a aussi un impact important sur la précision des méthodes proposées. Dans cette section, nous allons étudier, par simulation, l'impact de la taille du tampon sur les performances de nos estimations.

Nous allons exécuter les simulations avec la trace CAIDA, le scénario utilisé étant celui de la figure 5.1(b) (grand scénario). La trace utilisée pour le flux primaire est divisée en 600 morceaux différents. Chaque morceau sera simulé avec différentes valeurs de tampon, ce qui donne 600 simulations pour une taille de tampon fixée.

Intuitivement, il y a aussi un lien entre le nombre de paquets observés et l'impact de la perte des paquets. Par exemple, plus nous observons de paquets pour notre estimation, plus il y a de chances pour que des paquets soient perdus et plus l'impact de ces

pertes risque d'être important. Par conséquent, nous allons aussi considérer deux valeurs différentes pour le nombre de paquets observés : 500 paquets et 1000 paquets.

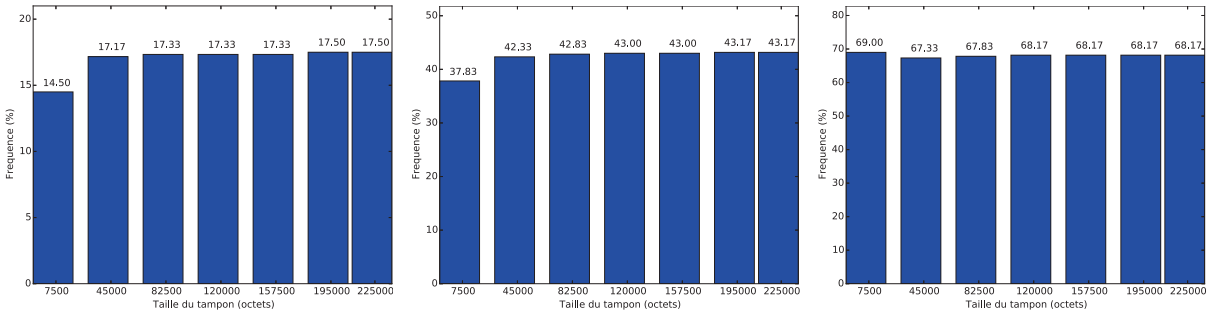
5.5.1 500 paquets observés



(a) 5 %. (b) 10 %. (c) 15 %.

FIGURE 5.15 – Pourcentage de cas où l'erreur relative, pour la méthode par décomposition en fonction de la taille du tampon, est inférieure à : 5% (figure 5.15(a)), 10% (figure 5.15(b)) et 15% (figure 5.15(c)).

Nous avons fait varier la taille du tampon sur chaque nœud entre 7500 et 225000 octets. La figure 5.15 présente le pourcentage de cas où l'erreur relative sur l'écart-type du délai de bout-en-bout, pour la méthode par décomposition, en fonction de la taille du tampon et pour 500 paquets observés, est inférieure à un seuil donné. On voit que qu'une taille de tampon inférieure à 82500 octets conduit à moins de cas présentant un bon écart-type du délai de bout-en-bout par rapport à la solution exacte. Ceci s'explique par le fait qu'un sous-dimensionnement du tampon induit des pertes de paquets qui ont un impact sur la précision de la méthode. Malgré tout, nous observons qu'avec un tampon de seulement 7500 octets, la méthode par décomposition permet d'assurer une erreur relative inférieure à 15% dans plus de 80% des cas.

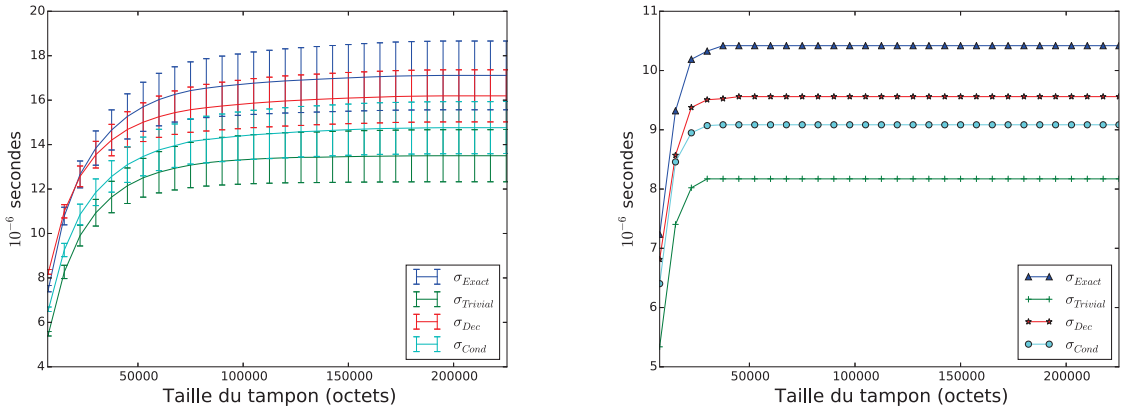


(a) 5 %. (b) 10 %. (c) 15 %.

FIGURE 5.16 – Pourcentage de cas où l'erreur relative, pour la méthode par l'espérance conditionnelle en fonction de la taille du tampon, est inférieure à : 5% (figure 5.16(a)), 10% (figure 5.16(b)) et 15% (figure 5.16(c)).

La figure 5.16 présente le pourcentage de cas où l'erreur relative sur l'écart-type du délai de bout-en-bout, pour la méthode par l'espérance conditionnelle, en fonction de la taille du tampon et pour 500 paquets observés, est inférieure à un seuil donné. La méthode par l'espérance conditionnelle se comporte différemment de la méthode par décomposition. Nous voyons que la taille du tampon a moins d'impact sur la précision avec la

méthode par l'espérance conditionnelle qu'avec la méthode par décomposition. Ceci peut s'expliquer par le fait que la méthode par l'espérance conditionnelle se base sur des temps d'attente mesurés, et non estimés, sur deux nœuds et que, par conséquent, elle est capable de reconnaître et prendre en compte les pertes dans son estimation. Les résultats sont néanmoins moins bons que ceux obtenus avec la méthode par décomposition.



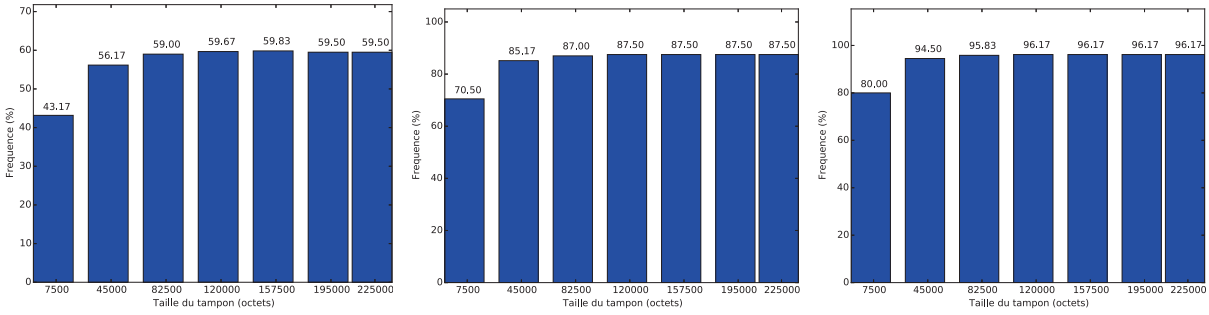
(a) Moyenne et intervalle de confiance.

(b) Médiane.

FIGURE 5.17 – Moyenne et médiane de l'écart-type pour les différentes méthodes.

La figure 5.17 présente la moyenne avec un intervalle de confiance de 95% et la médiane de l'écart-type du délai de bout-en-bout en fonction de la taille du tampon et pour 500 paquets observés. Dans ce contexte, c'est la méthode par décomposition qui obtient les résultats les plus proches de la solution exacte. Néanmoins, 500 paquets observés ne sont pas suffisants pour que les résultats soient très proches de ceux de la solution exacte.

5.5.2 1000 paquets observés



(a) 5 %.

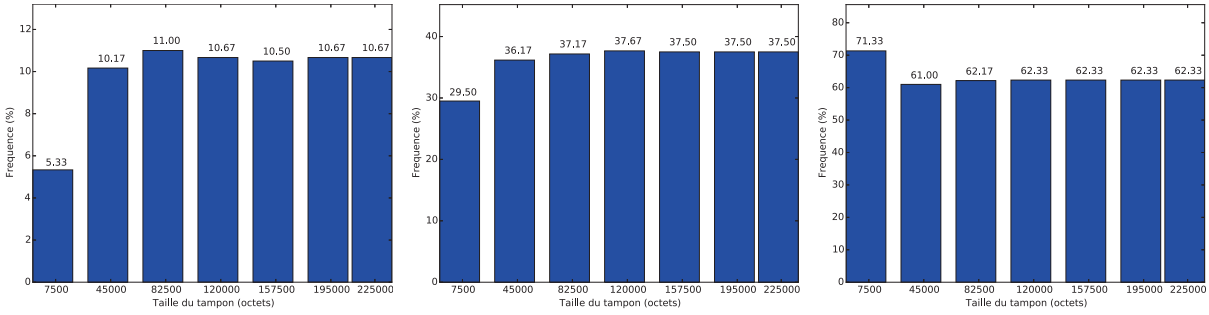
(b) 10 %.

(c) 15 %.

FIGURE 5.18 – Pourcentage de cas où l'erreur relative, pour la méthode par décomposition en fonction de la taille du tampon, est inférieure à : 5% (figure 5.18(a)), 10% (figure 5.18(b)) et 15% (figure 5.18(c)).

La taille du tampon sur chaque nœud varie là encore entre 7500 et 225000 octets. La figure 5.18 présente le pourcentage de cas où l'erreur relative sur l'écart-type du délai de bout-en-bout, pour la méthode par décomposition, en fonction de la taille du tampon et pour 1000 paquets observés, est inférieure à un seuil donné. Nous voyons que les résultats suivent la même évolution que celle observée sur la figure 5.15, mais qu'ils sont légèrement meilleurs.

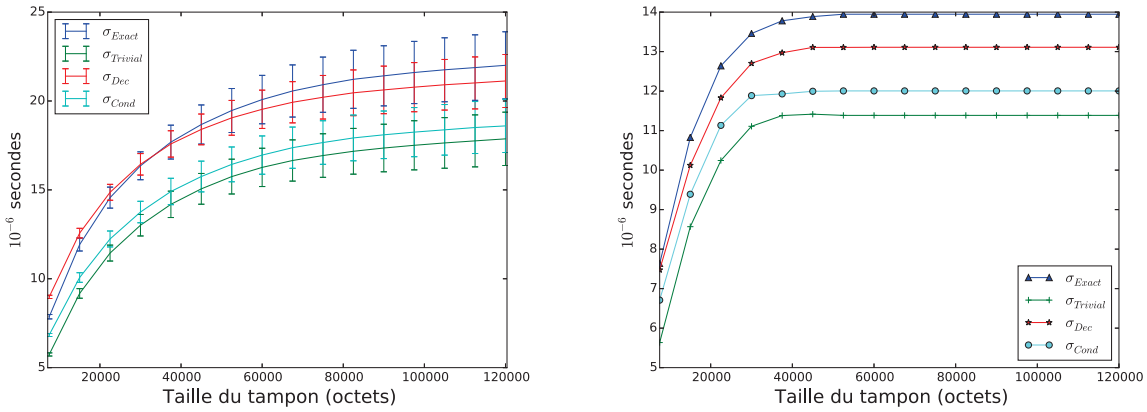
Ceci s'explique par le fait que deux fois plus de paquets sont observés par rapport à la section précédente. On peut noter qu'une taille de tampon de 45000 octets donne de bons résultats dès qu'on autorise une erreur relative sur l'écart-type inférieure à 10 ou 15%.
La figure 5.19 présente le pourcentage de cas où l'erreur relative sur l'écart-type du délai de bout-en-bout, pour la méthode par l'espérance conditionnelle, en fonction de la taille du tampon et pour 1000 paquets observés, est inférieure à un seuil donné. Là aussi, les résultats obtenus sont assez similaires à ceux donnés dans la figure 5.16, mais, à la différence de la méthode par décomposition, les résultats obtenus avec 1000 paquets observés sont moins bons qu'avec 500 paquets observés. Ceci est cohérent avec les résultats obtenus dans la section 5.4. On peut aussi noter que le comportement singulier obtenu avec un tampon de 7500 octets pour une erreur relative inférieure à 15% est accentué quand on passe de 500 paquets observés à 1000 paquets observés. Ce résultat peut peut-être s'expliquer par le fait que dans ce cas, le tampon est petit et il est donc toujours rempli. Par conséquent, le temps d'attente pour les paquets reçus devient plus stable et il est plus facile de prédire l'écart-type du délai de bout-en-bout.



(a) 5 %. (b) 10 %. (c) 15 %.

FIGURE 5.19 – Pourcentage de cas où l'erreur relative, pour la méthode par l'espérance conditionnelle en fonction de la taille du tampon, est inférieure à : 5% (figure 5.19(a)), 10% (figure 5.19(b)) et 15% (figure 5.19(c)).

Comme dans la sous-section précédente, la figure 5.20 présente la moyenne avec un intervalle de confiance de 95% et la médiane de l'écart-type du délai de bout-en-bout en fonction de la taille du tampon. Comme attendu, la méthode par décomposition est encore plus proche de la valeur exacte. On peut aussi noter que la méthode par décomposition semble sur-estimer l'écart-type quand la taille du tampon est petite. Avec les intervalles de confiance sur l'écart-type, on voit que la méthode par décomposition donne des valeurs très proches des valeurs exactes.



(a) Moyenne et intervalle de confiance.

(b) Médiane .

FIGURE 5.20 – Moyenne et médiane de l'écart-type pour les différentes méthodes.

5.6 Impact de l'intervalle de temps de la méthode par l'espérance conditionnelle

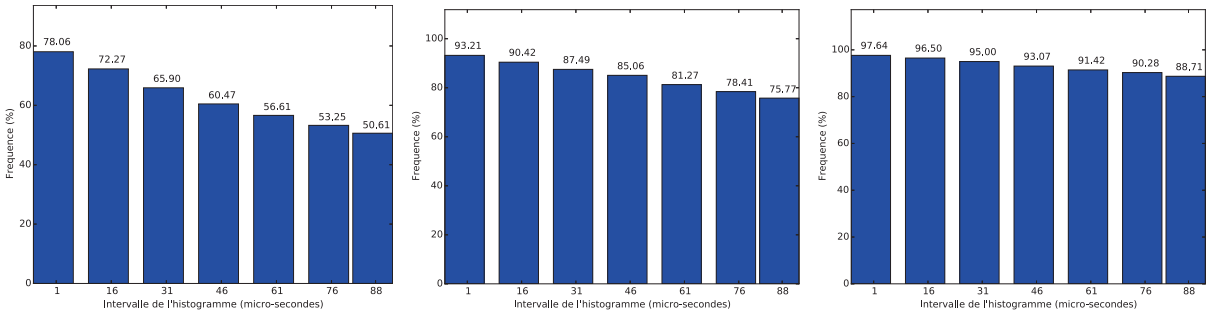
Cette section va étudier un paramètre important pour la méthode par l'espérance conditionnelle, à savoir l'intervalle de temps. Comme nous avons vu précédemment, la méthode par l'espérance conditionnelle regroupe des paquets selon le retour de la fonction par récurrence. Dans cette méthode, nous décidons de regrouper ou séparer des paquets selon une règle simple correspondant à un intervalle de temps. Nous considérons que deux paquets appartiennent à deux groupes différents si la métrique entre eux est plus grande que l'intervalle de temps. Intuitivement, un intervalle de temps plus petit donne un résultat plus proche de la valeur exacte. Dans cette section, nous évaluons l'impact de cet intervalle de temps sur la précision de la méthode.

Comme dans les sections précédentes, nous nous intéressons à trois niveaux pour l'erreur relative, à savoir inférieure à 5%, 10% et 15%.

5.6.1 Trace de l'Université de Stuttgart

Le scénario utilisé dans ce cas est celui de la figure 5.1(a) avec la trace de l'Université de Stuttgart. Le débit du flux primaire est fixé à $f_1 = 1,5$ (Mbps), le débit du flux secondaire est fixé à $f_{2,3} = 2$ (Mbps). Les trois liens ont les capacités suivantes : $\mathcal{C}_{1,2} = 8, \mathcal{C}_{1,2} = 5, \mathcal{C}_{1,2} = 3$ (Mbps). Chaque nœud a un tampon de 15000 octets. La trace de quatre heures est divisée en 1440 morceaux différents. Chaque morceau est utilisé dans plusieurs simulations ayant des valeurs différentes pour l'intervalle du temps. Le nombre de paquets observés dans ce scénario est 100 paquets.

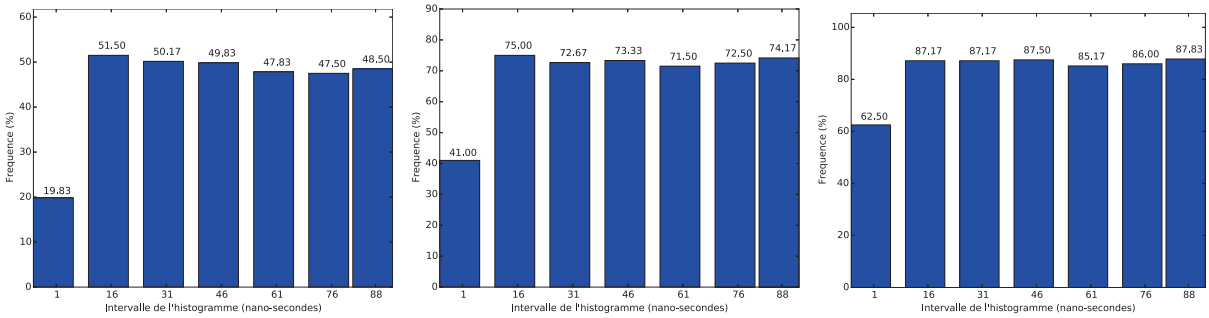
La figure 5.21 montre le pourcentage de cas où l'erreur relative est inférieure à un seuil donné pour la méthode par l'espérance conditionnelle et en fonction de la valeur de l'intervalle de temps. On observe que la courbe est décroissante : plus l'intervalle de temps est petite, meilleure est la précision de la méthode. Un petit intervalle de temps a néanmoins des inconvénients puisqu'il nécessite plus de calculs. Il faut donc trouver un compromis entre une bonne précision et un coût en calculs acceptable. On peut noter qu'une valeur de 0.000061 (secondes) permet d'obtenir plus de 80% de cas donnant une erreur relative inférieure à 10%.



(a) 5 %. (b) 10 %. (c) 15 %.

FIGURE 5.21 – Pourcentage de cas où l'erreur relative, pour la méthode par l'espérance conditionnelle en fonction de la valeur de l'intervalle de temps, est inférieure à : 5% (figure 5.21(a)), 10% (figure 5.21(b)) et 15% (figure 5.21(c)).

5.6.2 Trace CAIDA



(a) 5 %. (b) 10 %. (c) 15 %.

FIGURE 5.22 – Pourcentage de cas où l'erreur relative, pour la méthode par l'espérance conditionnelle en fonction de la valeur de l'intervalle de temps, est inférieure à : 5% (figure 5.22(a)), 10% (figure 5.22(b)) et 15% (figure 5.22(c)).

Le scénario utilisé dans cette sous-section est celui de la figure 5.1(a) avec la trace CAIDA. Le débit du flux primaire est fixé à $f_1 = 1,5$ (Gbps), les flux secondaires sont fixés à $f_1 = 1,5, f_{1,2} = 1,0, f_{2,5} = 0,2, f_{2,4} = 0,8, f_{3,4} = 1,5$ (Gbps). Les liens ont les capacités suivantes : $C_{1,2} = 6, C_{2,3} = 4, C_{3,4} = 7, C_{4,5} = 3, C_{5,6} = 5$ (Gbps). Chaque nœud a un tampon de 75000 octets (50 paquets en moyenne). La trace de quatre heures es divisée en 600 morceaux différents. Chaque morceau est utilisé dans plusieurs simulations ayant des valeurs différentes pour l'intervalle de temps. Le nombre des paquets observés dans ce scénario est 100 paquets.
Les résultats numériques obtenus sont plus surprenants. Les figures montrent clairement que le plus petit intervalle de temps ne mène pas au meilleur résultat. Au contraire, il semble que le premier cas avec un intervalle de 1ns (10^{-9} secondes) obtient le pire résultat.

5.7 Conclusions

Dans ce chapitre, nous avons évalué les performances des deux méthodes proposées en fonction de plusieurs paramètres.
– Nous avons d'abord étudié la structure de la trace réaliste dont la distribution de taille des paquets et l'inter-arrivé des paquets sont les deux paramètres principaux.

– Puis, nous avons examiné l'impact du nombre des paquets, l'impact de la taille du tampon et également l'impact de l'intervalle du temps (uniquement sur la méthode conditionnelle).
    1. Selon le nombre des paquets considérés, les performances de la méthode par décomposition s'améliore alors que la méthode conditionnelle se détériore en terme de précision.
    2. Le taux de perte a aussi un rôle important, c'est pourquoi les performances des deux méthodes devient meilleur quand la taille du tampon augmente.
    3. L'intervalle de temps considéré a un impact très important sur la précision de la méthode conditionnelle.

Les deux méthodes ont chacune des avantages et des inconvénients. Le choix d'utilisation d'une méthode dépend fortement du contexte : la taille du réseau, le niveau de l'indépendance des flux, la capacité d'un nœud (taille du tampon, capacité du lien, etc.), la fréquence des communications entre les nœuds et le contrôleur, etc.
Au chapitre suivant, nous résumons les contributions présentées dans ce manuscrit. Puis, nous allons aussi proposer quelques directions possibles pour les travaux futurs.

Chapitre 6

Conclusions et perspectives

Contents

6.1 Résumé des contributions 64
6.2 Travaux futurs 65

6.1 Résumé des contributions

Le délai de bout-en-bout : importance et difficultés

Dans le Chapitre 2, nous avons d'abord précisément défini le délai de bout-en-bout sur un chemin réseau et rappelé qu'il est le résultat de multiples facteurs (capacité des liens, type de trafic, délais de propagation,...). Le délai de bout-en-bout sur un chemin étant une quantité variable, il est naturel de le décrire par une caractérisation statistique. Comme souvent, une connaissance bien que partielle de ses deux premiers moments (équivalent à connaître sa moyenne et son écart type) fournit une bonne caractérisation. Ensuite nous avons expliqué l'importance de cette métrique de performances dans les réseaux IP actuels, et a fortiori dans les réseaux SDN. En effet le(s) contrôleur(s) SDN peut(vent) mettre à profit une connaissance de cette métrique dans leurs décisions de routage, d'équilibrage de charges et de garantie de qualité de service... Nous avons illustré deux façons possibles d'exploiter la connaissance des deux premiers moments du délai de bout-en-bout : grâce à l'inégalité de Chebyshev ou par une approche paramétrique.
Malheureusement la non-synchronisation des horloges sur les nœuds rend impossible la mesure directe du délai de bout-en-bout. En revanche il est possible de calculer ou de mesurer la plupart de ses composantes (délai d'attente, délai d'émission et délai de propagation). Ainsi l'objectif principal de ce manuscrit est de proposer des solutions permettant l'estimation des deux premiers moments du délai de bout-en-bout en s'appuyant uniquement sur des mesures localement mesurées sur chaque nœud. Toutefois, comme nous le montrons dans le Chapitre 2, certaines quantités mesurées localement peuvent être très corrélées entre elles, et ces corrélations doivent impérativement être prises en compte dans le calcul de l'écart type du délai de bout-en-bout. Notons que le cas (simple) du délai moyen de bout-en-bout d'un chemin réseau est traité en préambule du Chapitre 3.

Deux méthodes pour estimer son écart type

Les Chapitres 3 et 4 présentent deux méthodes différentes pour estimer l'écart type du délai de bout-en-bout : la méthode par décomposition et la méthode conditionnelle.

La méthode par décomposition

Le Chapitre 3 présente la méthode par décomposition. Elle repose sur la décomposition de l'écart type du délai de bout-en-bout en quantités locales à un nœud (ou à un lien) que sont les délais d'attente, délais d'émission et délais de propagation. Toutefois, apparaissent des termes de covariance entre des temps d'émission et des temps d'attente. Pour les estimer, nous faisons appel à une fonction dite de récurrence qui permet d'estimer les temps d'attente des paquets.

La méthode par l'espérance conditionnelle

Dans le Chapitre 4, nous présentons une toute autre méthode dite la méthode conditionnelle. Cette solution s'appuie sur la mise en correspondance des temps d'attente mesurés sur chacun des nœuds du chemin.

Étude empirique de leur précision

Le Chapitre 5 porte sur la précision des deux méthodes proposées. Pour cela nous considérons plusieurs exemples avec des topologies différentes, des tailles de paquets différentes,

du trafic différent (issu de traces réelles) et des niveaux de charge différents. Nous mettons également en évidence que la précision des méthodes dépend du nombre de paquets utilisés pour dériver les mesures locales. Nous expliquons pourquoi la méthode par décomposition voit sa précision baisser lorsque des pertes de paquets se produisent dans le réseau. Le principal inconvénient de la méthode par l'espérance conditionnelle est qu'elle doit faire parvenir régulièrement un ensemble de mesures au contrôleur SDN ce qui induit un trafic supplémentaire de données. Globalement, nos résultats numériques permettent de conclure que les deux méthodes se comportent généralement bien avec des erreurs relatives très souvent aux alentours de 10% et très rarement supérieur à 20%.

6.2 Travaux futurs

Les suites à donner à ce travail sont nombreuses. Nous en présentons brièvement quelques unes ci-dessous.

Les réseaux virtualisés

Les réseaux sont de plus en plus gérés comme une ressource dynamique et leurs resources peuvent être virtualisées et partagées entre plusieurs réseaux dits virtualisés (attention il ne s'agit pas de VLANs). Par exemple il est possible que dans un cloud plusieurs réseaux virtualisés soient déployés sur une seule et même infrastructure physique. Il est alors possible que le partage des ressources entre les différents réseaux virtualisés varie dans le temps (par exemple le nombre de CPU dans un commutateur qu'on alloue à un réseau donné peut changer selon le niveau de charge courant observé sur les autres réseaux co-existants). Si tel est le cas, des difficultés supplémentaires apparaîtront. Par exemple, pour estimer l'écart type du délai de bout-en-bout d'un chemin d'un réseau virtualisé, il faudra non seulement prendre en compte les inter-arrivées des paquets sur ce réseau virtualisé mais également celles des autres réseaux partageant le même nœud. Cela peut être compliqué notamment si une politique de sécurité veille à bien isoler les différents réseaux virtualisés. Le même type de problème peut se poser pour la capacité des liens partagée entre plusieurs réseaux virtualisés. Le partage de cette quantité peut être dynamiquement ré-ajustée entre les différents réseaux virtualisés. Le problème peut être rendu encore plus compliqué lorsque certains réseaux virtualisés sont prioritaires sur d'autres.

Les moments d'ordre supérieur

Dans ce manuscrit nous nous sommes concentrés exclusivement sur les deux premiers moments du délai de bout-en-bout. Toutefois une meilleure connaissance du délai de bout-en-bout passe notamment par la connaissance de ses autres moments d'ordre supérieur tels que le skewness et le kurtosis (ordre 3 et 4, respectivement).
Toutefois il est à vérifier, par exemple expérimentalement à travers des cas d'étude, quels seraient les gains liés à une connaissance plus fine (au delà de sa moyenne et de son écart type) du délai de bout-en-bout.

Découvrir des bornes théoriques

Dans notre travail, nous avons supposé que le trafic en entrée du réseau pouvait être quelconque. Cette hypothèse rend très difficile, sinon impossible, la découverte de bornes

supérieures sur la précision de nos approches. Un travail futur pourrait consister à considérer un trafic plus simple, par exemple un processus de Poisson pour les arrivées de paquets, et d'essayer de voir si, dans ce contexte plus restreint, il est possible de dériver des bornes probabilistes sur l'écart type du délai de bout-en-bout.

Bibliographie

[1] Ieee standard for a precision clock synchronization protocol for networked measurement and control systems. *IEEE Std 1588-2008 (Revision of IEEE Std 1588-2002)*, pages 1–269, July 2008.

[2] The CAIDA Anonymized Internet Traces 2014 Dataset, 2014. Trace avaiable : http ://www.caida.org/data/passive/passive_2014_dataset.xml.

[3] The CAIDA Anonymized Internet Traces 2015 Dataset, 2015. Trace avaiable : http ://www.caida.org/data/passive/passive_2015_dataset.xml.

[4] Arnold O. Allen. *Probablity, Statistics and Queueing Theory with Computer Science Applications, Second Edition*. Elsevier, 1990.

[5] G. Almes, S. Kalidindi, and M. Zekauskas. A One-way Delay Metric for IPPM. RFC 2679, IETF, 1999.

[6] Odlyzko Andrew. Networks are Lightly Utilized, and Will Stay that Way. *Review of Network Economics*, 2003.

[7] Roberto Bonafiglia, Ivano Cerrato, Francesco Ciaccia, Mario Nemirovsky, and Fulvio Risso. Assessing the performance of virtualization technologies for NFV : a preliminary benchmarking. In *IEEE Fourth European Workshop on Software Defined Networks*, 2015.

[8] Baek-Young Choi, Sue Moon, Rene Cruz, Zhi-Li Zhang, and Christophe Diot. Practical Delay Monitoring for ISPs. In *ACM CoNEXT*, 2005.

[9] Baek-Young Choi, Sue Moon, Zhi-Li Zhang, Konstantina Papagiannaki, and Christophe Diot. Analysis of Point-To-Point Packet Delay In an Operational Network. In *IEEE INFOCOM*, 2004.

[10] Gopal Dommety and Raj Jain. Potential networking applications of global positioning systems (gps). *arXiv preprint cs/9809079*, 1998.

[11] DPDK - Data Plane Development Kit. http ://dpdk.org/, 2016.

[12] Jeremy Elson, Lewis Girod, and Deborah Estrin. Fine-grained network time synchronization using reference broadcasts. *ACM SIGOPS Operating Systems Review*, 36(SI) :147–163, 2002.

[13] Nick Feamster, Jennifer Rexford, and Ellen Zegura. The road to sdn : an intellectual history of programmable networks. *ACM SIGCOMM Computer Communication Review*, 44(2) :87–98, 2014.

[14] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and S.C. Diot. Packet-level traffic measurements from the Sprint IP backbone. *IEEE Network*, 17(6), 2003.

[15] José Luis Garcia-Dorado, José Alberto Hernández, Javier Aracil, Jorge E. López de Vergara, and Sergio Lopez-Buedo. Characterization of the busy-hour traffic of IP networks based on their intrinsic features. *Computer Networks*, 55(9), 2011.

[16] Fotis Georgatos, Florian Gruber, Daniel Karrenberg, and Mark Santcroos. Providing Active Measurements as a Regular Service for ISP's. In *PAM*, 2001.

[17] A. Hassidim, M. Segalov, and A. Shaqed. Network utilization : The flow view. In *IEEE INFOCOM*, 2013.

[18] Hassan Hawilo, Abdallah Shami, Maysam Mirahmadi, and Rasool Asal. NFV : state of the art, challenges, and implementation in next generation mobile networks (vEPC). *IEEE Network*, 28(6) :18–26, 2014.

[19] Thomas R. Henderson et al. Network Simulations with the ns-3 Simulator. ACM SIGCOMM Demos, 2008. Code available : http ://www.nsnam.org/releases/.

[20] Ana Hernandez and Eduardo Magaña. One-way Delay Measurement and Characterization. In *IEEE ICNS*, 2007.

[21] Teerawat Issariyakul and Ekram Hossain. *Introduction to network simulator NS2*. Springer Science & Business Media, 2011. Code available : http ://networksimulator2.com/.

[22] Mary A Johnson and Michael R Taaffe. *The denseness of phase distributions*. School of Industrial Engineering, Purdue University, 1988.

[23] J. Kannisto, T. Vanhatupa, M. Hannikainen, and T. D. Hamalainen. Software and hardware prototypes of the ieee 1588 precision time protocol on wireless lan. In *2005 14th IEEE Workshop on Local Metropolitan Area Networks*, pages 6 pp.–6, Sept 2005.

[24] Bob Lantz, Nikhil Handigol, Brandon Heller, and Vimal Jeyakumar. Introduction to mininet. *Mininet Project,[En línea]. Available : https ://github.com/mininet/mininet/wiki/Introduction-to-Mininet.[Último acceso : 20 12 2015]*, 2014.

[25] Mark Leary. SDN, NFV, and open source : the operator's view, 2014.

[26] D Mills, J Martin, J Burbank, and W Kasch. RFC 5905 : Network time protocol version 4 : Protocol and algorithms specification. internet engineering task force (IETF), 2010. *tools. ietf. org/html/rfc5905*.

[27] David L Mills. Executive summary : computer network time synchronization, 2012.

[28] Pratap Misra and Per Enge. Global positioning system : Signals, measurements and performance second edition. *Massachusetts : Ganga-Jamuna Press*, 2006.

[29] Sue B. Moon. *Measurement and Analysis of End-to-end Delay and Loss in the Internet*. PhD thesis, University of Massachusetts at Amherst, 2000.

[30] C. Na, D. Obradovic, R. L. Scheiterer, G. Steindl, and F. J. Goetz. Synchronization performance of the precision time protocol. In *2007 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control and Communication*, pages 25–32, Oct 2007.

[31] A. Nucci, N. Taft, C. Barakat, and P. Thiran. Controlled use of excess backbone bandwidth for providing new services in IP-over-WDM networks. *IEEE JSAC*, 2004.

[32] Open vSwitch - An Open Virtual Switch. http ://www.openvswitch.org, 2016.

[33] OpenOnload. http ://www.openload.org, 2016. Solarflare.

[34] PacketShader. http ://shader.kaist.edu/packetshader/, February 2011.

[35] Sooksan Panichpapiboon and Wasan Pattara-Atikom. A review of information dissemination protocols for vehicular ad hoc networks. *IEEE Communications Surveys & Tutorials*, 14(3) :784–798, 2012.

[36] Konstantina Papagiannaki, Sue B. Moon, Chuck Fraleigh, Patrick Thiran, and Christophe Diot. Measurement and analysis of single-hop delay on an IP backbone network. *IEEE JSAC*, 21(6), 2003.

[37] Gergely Pongrácz, László Molnár, and Zoltán Lajos Kis. Removing roadblocks from SDN : openflow software switch performance on intel DPDK. In *2nd European Workshop on Software Defined Networks, EWSDN*, 2013.

[38] Luigi Rizzo. netmap : A novel framework for fast packet I/O. In *2012 USENIX Annual Technical Conference, June 13-15, 2012*, pages 101–112, 2012.

[39] Mendel Rosenblum. Vmware's virtual platform^{TM}. In *Proceedings of hot chips*, volume 1999, pages 185–196, 1999.

[40] Sandvine. Global Internet Phenomena Report, May 2013.

[41] D. Sass. The dormitory network "Selfnet" of the University of Stuttgart, October 2004. Trace avaiable : http ://www.ikr.uni-stuttgart.de/ sass/traces/Welcome.html#selfnet.

[42] R. L. Scheiterer, C. Na, D. Obradovic, and G. Steindl. Synchronization performance of the precision time protocol in industrial automation networks. *IEEE Transactions on Instrumentation and Measurement*, 58(6) :1849–1857, June 2009.

[43] J. Wang, M. Zhou, and Y. Li. Survey on the End-to-End internet Delay Measurements. In *HSNMC*, 2004.