

Paradigmes d'ordonnancement et de routage pour l'absence de contention dans un réseau dans un context 5G et au-delà

Maël Guiraud^{1,2}

¹David Laboratory, UVSQ

²Nokia Bell Labs France

27 septembre 2019

Introduction

Avec l'arrivée de la 5G et l'augmentation du nombre d'appareils connectés, les réseaux cellulaires sont confrontés à des enjeux tels que l'augmentation de la bande passante et la diminution de la latence.

Une des technologies privilégiées par les industriels des télécoms est le C-RAN (pour Cloud Radio Access Network). Il consiste à centraliser la partie calcul (Base Band Unit en anglais, que nous appellerons BBU) des différentes antennes radio d'un secteur dans un même datacenter, afin de faciliter la maintenance et le coût d'utilisation des équipements [2]. Le réseau d'accès entre le datacenter où se trouvent les BBU et la partie radio des antennes (Remote Radio Head, que nous appellerons RRH), s'appelle le fronthaul [3].

Le sujet de cette thèse nous a été proposé par Olivier Marcé et Brice Leclerc, chercheurs chez Nokia Bell Labs France. Nos travaux visent à maîtriser la latence dans le fronthaul. Nous nous basons dans un premier temps sur les contraintes de temps imposées par les protocoles utilisés pour la communication entre les antennes et les mobiles pour la 4G, pour lesquels la latence de bout en bout doit être de 3ms maximum [4]. Pour la 5G, la contrainte de latence de bout en bout pourrait descendre jusqu'à 1ms [1].

Nous proposons des algorithmes centralisés, qui correspondent très bien à la tendance actuelle en télécom, qui consiste à avoir un contrôleur SDN (Software Defined Network) [5], centralisant les informations et la gestion du réseau, ne laissant que peu de décision aux équipements sur la gestion de leur flux. Les flux que nous étudions sont envoyés aussi bien par les BBU que par les RRH de façon déterministe : nous considérons que les paquets envoyés ont tous la même taille, et sont envoyés périodiquement, il est donc naturel de les gérer d'une manière déterministe. L'idée de gérer de manière déterministe le réseau

pour réduire la latence est populaire et explique en partie l'engouement pour le DetNet (où Deterministic Networking) [6].

Encadrement et conditions de travail

Ma thèse CIFRE est une collaboration entre le laboratoire DAVID de l'Université de Versailles Saint-Quentin et Nokia Bell Labs France. Du côté académique, je suis encadré par Dominique Barth et Yann Strozecki. Olivier Marcé et Brice Leclerc sont mes référents au sein de l'entreprise. La répartition du temps de travail sur une semaine varie en fonction des objectifs et des travaux en cours, mais je ne rencontre aucune difficulté pour partager mon temps. Nous organisons des réunions mensuelles avec tous les encadrants afin de se coordonner sur les objectifs de recherche.

1 Modèle

1.1 Modélisation par graphe et dynamique des messages

Nous avons choisi de modéliser le réseau comme un graphe orienté $G = (V, A)$, dans lequel les sommets représentent les noeuds du réseau et les arcs (u, v) représentent les liens physiques entre ces noeuds. Ces arcs sont pondérés par un entier $\Omega(u, v)$ qui représente le temps physique pris par un message pour traverser le lien. Une **route** est un chemin simple, c'est à dire une suite de sommets adjacents u_0, \dots, u_l , avec $(u_i, u_{i+1}) \in A$. La **latence** d'un sommet u_i sur une route r est définie par $\lambda(u_i, r) = \sum_{0 \leq j < i} \Omega(u_j, u_{j+1})$. En particulier, la

latence du premier sommet d'une route est $\lambda(u_0, r) = 0$. La longueur d'une route r est égale à $\lambda(r) = \lambda(u_l, r)$. Le premier sommet d'une route représente une RRH et le dernier sommet représente la BBU qui y est associée. Un **réseau routé** (G, \mathcal{R}) est un graphe G et un ensemble de routes dans ce graphe.

Nous travaillons dans un modèle à temps discret. L'unité de temps est appelée le **tic**, et modélise le temps nécessaire pour transmettre un paquet de 64 octets (la plus petite taille d'un paquet en ethernet) sur un lien. La vitesse de propagation des messages dans les liens étant fixe, on peut donner le poids des arcs en tics, ce qui correspond au temps nécessaire à un message pour traverser le lien. Dans les applications que nous étudions, un message est envoyé sur chaque route périodiquement. La durée de la période est notée P et nous considérons que l'envoi de tout les messages ont la même périodicité. Considérons la route $r = (u_0, \dots, u_l)$. Un message envoyé à la date m de u_0 , le premier sommet de la route r arrivera au sommet u_i à la date $t(u_i, r) = m + \lambda(u_i, r)$. Étant donné que l'envoi des messages est périodique, un message envoyé dans un arc à la date $t \in [0, P - 1]$ utilisera le même arc à la date $t + kP$, pour tout k entier positif. Tout les calculs peuvent donc être fait modulo P .

En général, les messages utilisent plus d'un tic pour être émis. La taille d'un message est notée τ et représente le nombre de tics consécutifs nécessaire pour envoyer ce message. Nous considérons que tout les messages ont la même taille.

L'ensemble des tics utilisés par un message sur une route r dans l'arc (u, v) est noté $[t(u, r)]_{P, \tau}$ et vaut $\{t(u, r) + i \bmod P \mid 0 \leq i < \tau\}$. Considérons deux routes r_1 et r_2 , et deux messages m_1 et m_2 envoyés respectivement du premier sommet de r_1 et r_2 . Deux routes sont en **conflit** si elles ont arc (u, v) en commun et que $[t(u, r_1)]_{P, \tau} \cap [t(u, r_2)]_{P, \tau} \neq \emptyset$.

Dans notre cas d'application du C-RAN, quand un message est envoyé depuis la RRH à la BBU, cette dernière effectue un calcul et envoie une réponse à la RRH. Dans nos travaux, un réseau routé contient deux ensembles de routes : un ensemble de routes aller, de chaque RRH à chaque BBU et un ensemble de routes retour, de chaque BBU à chaque RRH. Il y a une bijection ρ entre les routes aller et les routes retour. En général, les deux routes utilisent les mêmes arcs avec une orientation différente.

Un (P, τ) -**periodic assignment** d'un réseau routé (G, \mathcal{R}) est une fonction qui à chaque route $r \in \mathcal{R}$ associe un **offset** m_r , c'est à dire une date à laquelle un message est émis depuis le premier sommet de la route r . Dans un (P, τ) -periodic assignment, **aucune paire de route n'est en conflit**.

Un assignment représente le processus suivant : un message envoyé par u , le premier sommet d'une route aller r à la date m_r . Ce message est reçu par v , le dernier sommet de r à la date $t(v, r)$. Ensuite, la réponse est envoyée par v à u à la date $m_{\rho(r)}$. Le temps écoulé entre l'arrivée d'un message en v et l'envoi de sa réponse est appelé **temps d'attente** et est défini par $w_r = m_{\rho(r)} - t(v, r)$. La figure 1 illustre le processus d'échange de message entre une RRH et sa BBU.

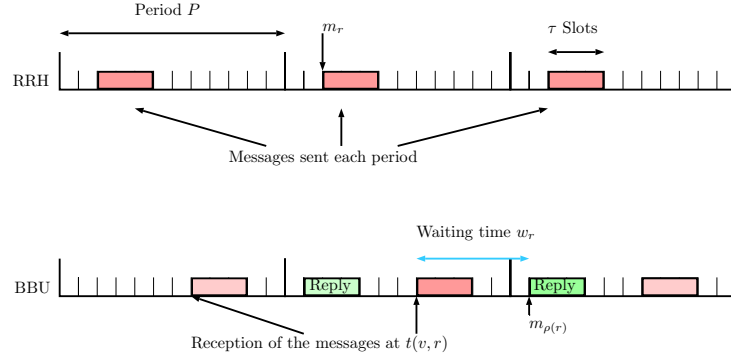


FIGURE 1 – L'envoi d'un message et de sa réponse.

Le temps de calcul avant l'envoi de la réponse n'est pas pris en compte. En effet, il est possible de considérer ce temps comme du temps de transport dans les arcs allant à la BBU afin de simplifier le problème. Le temps total d'aller retour d'un message envoyé sur une route r est $PT(r) = \lambda(r) + w_r + \lambda(r)$. Le temps total d'aller retour de chaque route doit respecter une **deadline**. Dans ce rapport, on considère la que deadline est la même pour toutes les routes et est notée T_{max} .

Nous définissons aussi le **temps de transmission** d'un réseau routé $TR(G, R) =$

$\max_{i \in \{0, \dots, n\}} t(\rho_i, u_i) - \min_{j \in \{0, \dots, n\}} m_j$. C'est le temps écoulé entre l'envoi du premier message à l'aller et la réception de la dernière réponse au retour.

1.2 Problèmes algorithmiques

Dans un premier temps, nous avons étudié un problème dans lequel les messages font un aller-retour et peuvent être mis en attente dans la BBU. Chaque message doit respecter un latence maximale. Cela correspond à des applications dans lesquelles les antennes n'ont pas besoin d'être synchrone pour envoyer et recevoir leur signaux, ce qui n'est pas possible avec les protocoles actuels. En revanche, on peut imaginer que les technologies à venir peuvent fonctionner de manière désynchronisées. De plus, ce cas de figure peut convenir à d'autres applications, comme un réseau de capteurs par exemple. Nous introduisons donc le problème suivant :

Periodic Assignment for Low Latency (PALL)

Entrées : un réseau routé (G, \mathcal{R}) , les entiers P , τ et T_{max} .

Problème de décision : Existe-t-il un (P, τ) -periodic assignment m de (G, \mathcal{R}) tel que pour toute route $r \in \mathcal{R}$, $PT(r) \leq T_{max}$?

Nous proposons aussi le problème d'optimisation suivant, dans lequel l'objectif est de minimiser la pire des latence des messages. En effet, réduire au plus possible ce temps permet d'accepter des applications de plus en plus exigeante en terme de latence.

Problème d'optimisation : Trouver un (P, τ) -periodic assignment qui minimise $\max_{r \in \mathcal{R}} PT(r)$.

Nous avons aussi étudié un autre problème, dans lequel les BBU n'ont pas le droit d'attendre après la réception d'un message avant d'envoyer la réponse, ce qui permet d'utiliser des technologies sans file d'attente.

Periodic Assignment for Zero Latency (PAZL)

Entrée : Un réseau routé (G, \mathcal{R}) , les entiers P et τ .

Problème de décision : Existe-t-il un assignment m de (G, \mathcal{R}) , dans lequel $w_r = 0$ pour toute route $r \in \mathcal{R}$.

Dans le cas d'application du C-RAN, les antennes envoient leurs messages de façon synchronisée, c'est à dire au même tic à chaque période. Ensuite, toutes les réponses doivent revenir aux antennes avant une deadline générale. Nous définissons donc le problème suivant :

Synchronized Periodic Assignment for Low Latency (SPALL)

Entrées : Un réseau routé (G, \mathcal{R}) , les entiers P , T_{max} .

Problème de décision : Existe-t-il un assignment m de (G, \mathcal{R}) tel que $TR(G, \mathcal{R}) \leq T_{max}$?

Problème d'optimisation : Trouver un (P, τ) -periodic assignment qui minimise $TR(G, \mathcal{R})$.

Dans le cas où les antennes sont synchronisées, il n'est pas pertinent de chercher des assignments avec zéro temps d'attente, car ça n'a pas d'impact sur la mesure que nous cherchons à optimiser, en l'occurrence $TR(G, \mathcal{R})$.

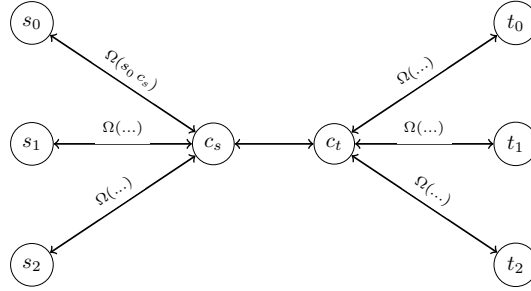


FIGURE 2 – Réseau sous forme d'étoile

Il apparaît que les problèmes exposés ressemblent à des problèmes d'ordonnancement. En revanche, la périodicité des messages nous empêche de réduire notre problème à un problème d'ordonnancement classique.

2 Topologies étudiées

Nous définissons la **profondeur de conflit** d'une route comme le nombre d'arc que cette route partage avec d'autres routes.

2.1 Réseau sous forme d'étoile : profondeur de conflit 1

Dans un premier temps, pendant mon stage de master et le début de ma thèse, nous avons étudié un réseau sous forme d'étoile. C'est un réseau réaliste qui correspond à des architectures Point-to-Multipoint (PtMP), assez courantes dans le C-RAN [7]. Le graphe G comprend deux ensembles de sommets $S = \{s_0, \dots, s_{n-1}\}$ et $T = \{t_0, \dots, t_{n-1}\}$ de taille n et deux sommets spéciaux : c_s (pour central sources) et c_t (pour central target). Il y a un arc entre c_s et c_t et pour chaque route i , il y a un arc entre s_i et c_s et entre t_i et c_t . Les routes sont les chemins $r_i = (s_i, c_s, c_t, t_i)$ à l'aller et $\rho(r_i) = (t_i, c_t, c_s, s_i)$ au retour. Ce genre de topologie peut paraître simple, mais nous avons pu prouver que le problème SPALL est NP-complet [8], même sur cette topologie, et nous suspectons que les problèmes PAZL et PALL le sont aussi car des variantes dans lequel le lien central est partagé le sont [9]. Dans la prochaine section nous verrons quelques algorithmes gloutons et FPT qui permettent d'obtenir de bonnes solutions pour ces trois problèmes sur cette topologie.

Nous avons proposé plusieurs solutions pour les problèmes PAZL et PALL pour cette topologie, les résultats ont été présentés à la conférence ICT 2018 [10], et une version longue a été soumise au journal TCS [11].

2.2 Anneau optique

J'ai travaillé avant le début de ma thèse pour le projet ANR NGREEN, pour lequel j'ai été amené à étudier le comportement du trafic C-RAN sur un anneau optique novateur [12, 13]. Après modélisation de l'anneau et de sa technologie et définition du problème algorithmique d'organisation des flux déterministes entre eux, nous nous sommes aperçu que notre problème était trivial en raison des paramètres spécifiques de l'architecture. Néanmoins, l'interface entre le réseau électronique et le réseau optique était une nouveauté qu'il a fallu modéliser correctement pour bien résoudre le problème. De plus, nous avons proposé des solutions pour organiser nos flux prioritaires de façon à impacter le moins possible la latence des flux dit "best effort" [11]. Suite à nos travaux, il s'avère que le fait d'organiser correctement les flux déterministes aide même à réduire la latence des flux best effort.

2.3 Topologie plus générale : profondeur de conflit 2

Après avoir beaucoup étudié la topologie de profondeur 1, nous nous sommes penchés sur une topologie plus générale, tout en restant sur des réseaux réalistes. Nous avons naturellement choisi de regarder des DAG de profondeur 2, c'est à dire dans lesquels une route n'a que deux arcs sur lesquels elle est en conflit avec les autres routes. La figure 3 représente le type de réseau que nous étudions.

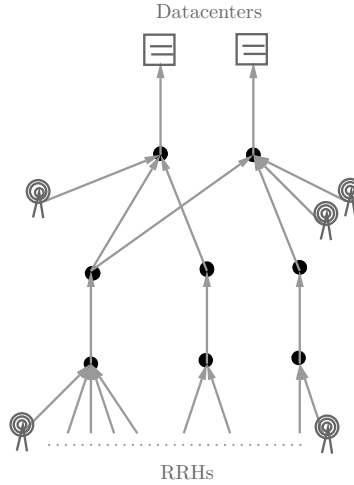


FIGURE 3 – Un réseau de profondeur de conflit 2.

3 Solutions algorithmiques pour l'étoile.

Nous définissons la **charge** du réseau comme le rapport entre le nombre de tics nécessaire pour faire passer les messages de toutes les routes et le nombre de tics d'une période.

3.1 PAZL

Nous avons dans un premier temps choisi d'étudier le problème PAZL, dans lequel toutes les solutions trouvées seraient optimales en latence.

Dans un premier temps nous proposons plusieurs algorithmes gloutons. Nous avons proposé un algorithme glouton qui nous assure de trouver une solution si la charge du réseau est inférieure à $1/3$. Cet algorithme fonctionne de la manière suivante : pour chaque route, on attribue le premier temps de départ disponible qui ne crée pas de conflit avec les routes déjà placées.

Nous avons aussi prouvé que si la différence entre la taille des routes est suffisamment petite par rapport à la période, il existe une solution au problème PAZL. Dans ce cas, on trie les routes par ordre croissant de taille et on envoie les messages collées dans l'arc central à l'aller, de celui qui utilise la route la plus courte à celui qui utilise la route la plus longue. Nous proposons également un algorithme de recherche exhaustive de complexité FPT en n le nombre de routes, ce qui est intéressant car le nombre de route est le paramètre le plus petit de nos instances. L'idée de l'algorithme est de décrire une solution sous une forme normale dans laquelle les messages sont collées entre eux soit à l'aller, soit au retour, et de tester tout les ordres possibles des messages à l'aller et au retour. Après avoir simulé le comportement de nos algorithmes sur des graphes avec des tailles de routes différentes :

- Dans le cas de routes courtes, (inférieure en nombre de tics à la taille des messages), il est facile d'organiser les messages.
- Dans le cas de routes longues, nous avons comparé nos algorithmes gloutons à l'algorithme FPT et nous nous sommes rendu compte que nos algorithmes ne trouvent que très peu de solutions au dessus de 0,7 de charge, alors que la recherche exhaustive montre qu'il existe des solutions jusqu'à 0,9 de charge.

La figure 4 représente les performances des algorithmes que nous venons de présenter, sur des instances avec des routes longues.

3.2 PALL

Après avoir vu les limites du problème PAZL, nous nous sommes intéressés au problème PALL, dans lequel nous pouvons ajouter du temps d'attente dans les BBU. Ce problème nous offre un plus grand degré de liberté, en effet, nous nous autorisons à dégrader la latence de certains messages pour pouvoir trouver des solutions dans lesquelles la condition de latence globale est satisfaite.

Nos algorithmes pour résoudre ce problème peuvent se décomposer en deux phases. D'abord on fixe la phase aller, de façon arbitraire, puis on essaye d'or-

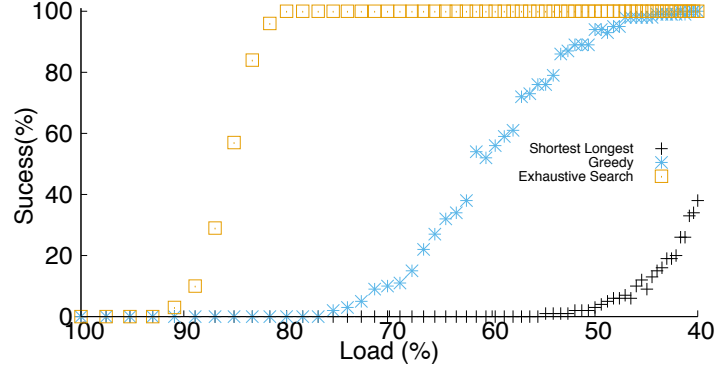


FIGURE 4 – Taux de réussite des algorithmes en fonction de la charge.

ganiser les messages au retour. Pour ce faire, nous avons étudié dans un premier temps des algorithmes gloutons, peu efficaces, qui renvoient les messages selon la taille des routes ou selon la deadline. Nous avons ensuite choisi d'utiliser le lien qui existe entre l'ordonnancement de la phase retour du problème et un problème classique d'ordonnancement. On a donc pour le retour des release times, des deadlines, et un lien en commun sur lequel faire passer tout les messages. Nous avons donc utilisé un algorithme d'ordonnancement qui trouve une solution dans laquelle le temps de complétion est minimal, en temps polynomial [14]. Cet algorithme ne donnait toujours pas de bonnes performances, car il ne prends pas la périodicité du problème en compte. Nous avons dans un premier temps essayé d'adapter au mieux possible les solutions données pour gérer ce problème de périodicité. Finalement, nous avons imaginée un algorithme FPT en le nombre de routes, qui réutilise l'algorithme d'ordonnancement mais qui travaille sur une seule période.

La figure 5 représente les performances de différents algorithmes que nous proposons. Un algorithme greedy adapté pour la périodicité (GD), l'algorithme d'ordonnancement qui ne gère pas du tout la périodicité (MLS), adapté pour la périodicité (PMLS), et l'algorithme FPT (FPT-PMLS). Pour chaque instance, nous fixons la latence maximal à deux fois la taille de la route la plus longue, puis nous avons fait varier le margin, qui correspond à du temps de latence supplémentaire donné aux messages.

Les deux algorithmes (PMLS et FPT PMLS) présentent d'excellentes performances, même sur un réseau très chargé. En pratique l'algorithme FPT ne trouve pas beaucoup plus de solutions que PMLS. En effet, il y a très peu d'instances pour lesquelles il est utile d'utiliser l'algorithme FPT. Il est donc recommandé d'utiliser l'algorithme PMLS, surtout quand il y a beaucoup de routes (calcul 15000 fois plus long avec l'algorithme FPT pour 24 routes).

Afin de fixer l'ordre et la date de départ des temps à l'aller, nous avons essayé plusieurs variantes. En ce qui concerne l'ordre, nous avons essayé d'envoyer les

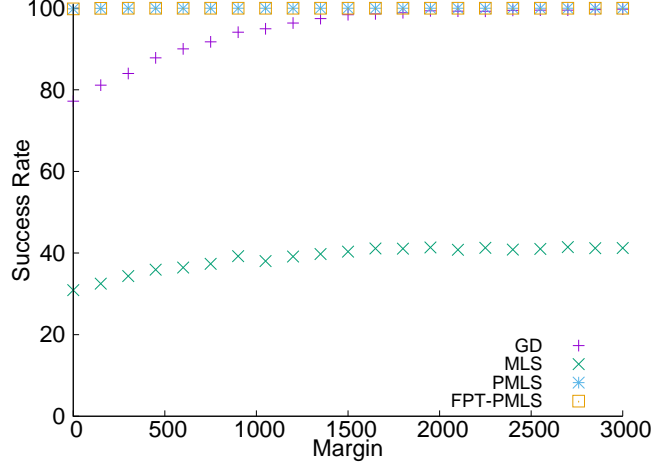


FIGURE 5 – Taux de réussite des algorithmes pour le retour, à 95% de charge.

messages sur les routes de la plus longue à la plus courte, où inversement, où de les envoyer dans un ordre aléatoire. En ce qui concerne les dates de départ, nous avons essayé de coller tout les messages (et donc de les envoyer le plus tôt possible), où encore d'espacer les messages de façon uniforme, ou aléatoire dans la période. Finalement, après des simulations, nous avons conclu que la meilleure façon de trouver une solution pour le problème PALL est de générer un grand nombre d'ordres aléatoires, de coller tout les messages et de chercher une solution pour le retour avec l'algorithme FPT présenté plus haut.

Nous avons aussi proposé un algorithme FPT pour le problème PALL (aller et retour). À une instance on associe un nombre exponentiel (qui dépend du nombre de routes) de systèmes linéaires tels que si l'un au moins d'entre eux est satisfait, il existe une solution. Nous ne l'avons pas implémenté car les performances de l'algorithme précédent sont déjà excellentes. Néanmoins, ayant trouvé des papiers récents en rapport avec nos travaux, nous avons contacté plusieurs équipes qui utilisent la programmation linéaire pour résoudre un problème très similaire au notre, afin de collaborer et de mettre en parallèle nos travaux respectifs [15, 16].

3.3 SPALL

Le problème SPALL est en fait une variante plus simple de PALL sur l'étoile. Étant donné que le temps est compté dès l'envoi du message depuis la RRH, il n'y a pas d'intérêt à retarder les messages dans la RRH. Pour choisir comment

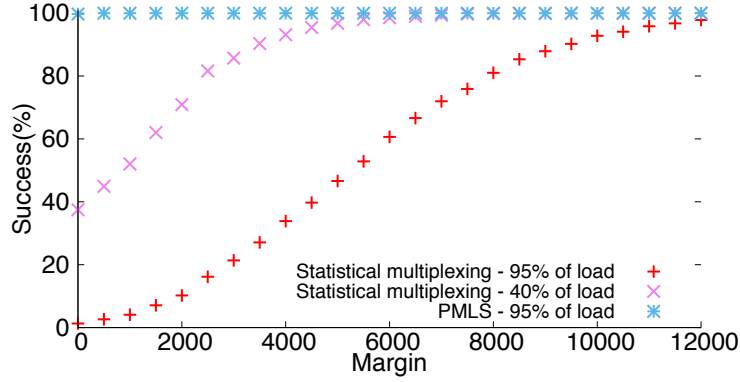


FIGURE 6 – Taux de réussite de PMLS comparé au multiplexage statistique.

envoyer les routes à l’aller, il suffit donc de choisir l’ordre d’envoi des routes. Ensuite, nous pouvons utiliser l’algorithme FPT utilisé pour PALL au retour. Nous avons donc un algorithme FPT pour résoudre le problème SPALL sur l’étoile.

3.4 Comparaison avec une gestion statistique des messages

Afin d’évaluer les performances de nos algorithmes d’organisation déterministe des flux dans les réseaux, nous avons choisis d’implémenter un simulateur du comportement actuel des réseaux ; tout les messages sont envoyées dès qu’ils sont disponible, et si ils doivent utiliser un même lien en même temps, alors l’un des deux est mis en attente dans le noeud de départ de ce lien en attendant qu’il puisse utiliser la ressource partagée. Nous proposons plusieurs politiques de gestions des messages dans les files d’attentes en cas de conflit. La première est le FIFO basique, et la seconde est d’envoyer le message qui à déjà passé le plus de temps dans le réseau. La figure 6 représente les performances de notre algorithme PMLS, comparé à une gestion statistique des flux.

4 Intégration du travail en entreprise

Dans le cadre de ma thèse CIFRE avec Nokia Bell Labs France, nous sommes amenés du côté industriel à justifier l’utilité et la faisabilité des solutions que nous proposons sur le plan technique. Pour cela, nous développons un prototype. L’idée est d’utiliser un contrôleur SDN, c’est à dire un serveur qui centralise l’intelligence du réseau. Il est directement connecté à tout les appareils du réseau et contrôle la gestion des flux dans chaque noeud.

Dans notre cas, nous avons choisis ONOS (Open Network Operating System) comme support logiciel pour gérer notre réseau. ONOS est, comme son nom

l'indique un système d'exploitation avec des fonctionnalités implémentée pour récupérer les informations sur le réseau et envoyer les instructions de gestion des flux aux noeuds.

Un des enjeu technique consiste à gérer les messages à la nanoseconde près, et donc à utiliser du materiel adéquat. La technologie utilisée pour gérer les flux de manière très fine est relativement nouvelle et il existe très peu de switch dit "TSN" (Time Sensitive Networking). De plus, ce genre de matériel est très cher. Nous avons donc choisi pour des raisons principalement économiques d'utiliser une machine standard avec un linux modifié qui nous permet d'utiliser la technologie TSN sur les cartes réseaux.

5 Objectifs pour la fin

Nous avons beaucoup de travaux en cours à terminer avant la fin de ma thèse. Du coté académique, nous prévoyons de finir les résultats théoriques que nous avons trouvés sur l'étoile. Nous avons imaginé un algorithme qui s'autorise à modifier les solutions partielles trouvées pour pouvoir faire passer les messages le plus contraints. Nous pensons qu'il est possible avec un tel algorithme de résoudre le problème PAZL sur l'étoile si la charge est inférieure à $2/3$. Nous avons aussi testé d'importer les algorithmes gloutons de l'étoile sur la topologie de profondeur de conflit 2, mais il s'avère que les performances ne sont pas très bonnes. De plus, les algorithmes FPT ne sont pas applicables à cette topologie car ils s'appuient sur le fait qu'il n'existe qu'un arc en commun entre toutes les routes. Nous avons néanmoins de bonnes idées, utilisant les techniques de voisinage pour résoudre le problème sur cette topologie plus générale. Nous prévoyons donc d'écrire deux papiers sur le sujet : Un papier plus théorique avec des résultats sur le problème PAZL, et un papier plus axé sur l'évaluation de performance de nos algorithmes sur la topologie de profondeur de conflit 2.

Du coté industriel, nous avons toutes les idées pour avancer sur la réalisation de notre prototype, mais nous allons surement nous heurter à des contraintes techniques au fur et a mesure que les choses avanceront. **TODO : enseignements**

Références

- [1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.
- [2] C. Mobile, "C-RAN : the road towards green RAN," *White Paper*, ver, vol. 2, 2011.
- [3] T.-S. N. T. G. of IEEE 802.1, "Time-sensitive networks for fronthaul," July 2016. IEEE P802.1/D0.4.
- [4] Y. Bouguen, E. Hardouin, A. Maloberti, and F.-X. Wolff, *LTE et les réseaux 4G*. Editions Eyrolles, 2012.

- [5] E. Haleplidis, K. Pentikousis, S. Denazis, J. H. Salim, D. Meyer, and O. Koufopavlou, “Software-defined networking (sdn) : Layers and architecture terminology,” tech. rep., 2015.
- [6] N. Finn and P. Thubert, “Deterministic Networking Architecture,” Internet-Draft draft-finn-detnet-architecture-08, Internet Engineering Task Force, Sept. 2016. Work in Progress.
- [7] Z. Tayq, L. A. Neto, B. Le Guyader, A. De Lannoy, M. Chouaref, C. Aupetit-Berthelemot, M. N. Anjanappa, S. Nguyen, K. Chowdhury, and P. Chanclou, “Real time demonstration of the transport of ethernet fronthaul based on vran in optical access networks,” in *Optical Fiber Communications Conference and Exhibition (OFC), 2017*, pp. 1–3, IEEE, 2017.
- [8] A. J. Orman and C. N. Potts, “On the complexity of coupled-task scheduling,” *Discrete Applied Mathematics*, vol. 72, no. 1-2, pp. 141–154, 1997.
- [9] W. Yu, H. Hoogeveen, and J. K. Lenstra, “Minimizing makespan in a two-machine flow shop with delays and unit-time operations is np-hard,” *Journal of Scheduling*, vol. 7, no. 5, pp. 333–348, 2004.
- [10] D. Barth, M. Guiraud, B. Leclerc, O. Marcé, and Y. Strobecki, “Deterministic scheduling of periodic messages for cloud RAN,” in *25th International Conference on Telecommunications, ICT 2018, Saint Malo, France, June 26-28, 2018*, pp. 405–410, 2018.
- [11] D. Barth, M. Guiraud, and Y. Strobecki, “Deterministic contention management for low latency cloud RAN over an optical ring,” *CoRR*, vol. abs/1902.03018, 2019.
- [12] D. Chiaroni, “Network energy : Problematic and solutions towards sustainable ICT,” *invited paper, International Commission of Optics (ICO-24)*, August 2017.
- [13] B. Uscumlic, D. Chiaroni, B. Leclerc, T. Zami, A. Gravey, P. Gravey, M. Morvan, D. Barth, and D. Amar, “Scalable deterministic scheduling for WDM slot switching xhaul with zero-jitter,” in *ONDM 2018*, IEEE.
- [14] B. Simons, “A fast algorithm for single processor scheduling,” in *Foundations of Computer Science, 1978., 19th Annual Symposium on*, pp. 246–252, IEEE, 1978.
- [15] W. Steiner, S. S. Craciunas, and R. S. Oliver, “Traffic planning for time-sensitive communication,” *IEEE Communications Standards Magazine*, vol. 2, no. 2, pp. 42–47, 2018.
- [16] S. S. Craciunas, R. S. Oliver, and W. Steiner, “Formal scheduling constraints for time-sensitive networks,” *arXiv preprint arXiv :1712.02246*, 2017.