

Expérimentations SAMU

Une fois le système réel analysé, il est alors possible de chercher à améliorer la situation courante et de réduire le temps d'attente des appels ainsi que d'améliorer le QoS comme décrit au chapitre 1.3.



Figure 5.1 – Partie de la méthodologie traitée dans ce chapitre

Ce chapitre correspond à la partie gauche (cf. figure 5.1) de notre cadre conceptuel présenté au chapitre 3. Dans cette partie, on cherche de nouveaux « traitements » pour améliorer la performance du SAMU (QoS et Temps d'attente maximum). Plutôt que de tester ces traitements sur le processus réel (long, ou impossible, ou coûteux, ou risqué, etc.), on décide de développer un double numérique, véritable réplique virtuelle du processus de prise en charge des appels, grâce auquel il sera possible de reproduire l'organisation existante pour mieux la comprendre et tester d'autres

schémas organisationnels, avant une implémentation éventuelle dans le monde réel. La question de recherche abordée dans ce chapitre est finalement :

Considérant l'hétérogénéité des SAMU, quel processus de décroché des appels et de gestion des ressources est le plus à même de soutenir leur profil d'arrivée en temps normal ainsi que lors d'une augmentation soudaine du nombre d'appels ?

5.1 Étape 3 : Diagnostiquer

5.1.1 Sous-étape 3.a : Paramétrer le double numérique ; Hypothèses et Modèle conceptuel

Nous avons développé notre modèle de manière à ce qu'il soit le plus proche possible de la réalité. Toutefois, comme dans tous les modèles, il est nécessaire de faire des hypothèses, car nous ne disposons pas de l'information, ou parce que nous ne pouvons pas prendre en compte de l'intégralité de l'information. Ainsi, notre modèle sera valable sous certaines conditions.

Sélectionner l'outil support au double numérique

La première étape est de choisir l'outil support à la mise en oeuvre de ce double numérique.

Il existe plusieurs types de simulation : on trouve notamment la simulation à base d'agents, la simulation à événements discrets ainsi que la dynamique des systèmes.

La simulation à événements discrets

La simulation à événements discrets (SED) concerne la modélisation d'un système qui évolue dans le temps en prenant différents états. Les changements d'état se produisent à des instants correspondant à des événements (Stewart Robinson, 2004). Lors des événements les variables qui définissent le système changent, ce qui peut amener le système à un nouvel état.

En effet, un événement peut ne pas changer l'état du système et peut simplement correspondre à une récupération d'information, ou encore à une planification d'un événement futur dans le système, ce qui n'a pas d'impact direct sur l'état du système.

La SED est une technique de recherche opérationnelle qui est utilisée dans des domaines variés allant de l'industrie, à la santé en passant par la défense (Law et al., 2000). Elle offre plusieurs avantages :

- Elle permet de modéliser des systèmes complexes ayant de nombreux paramètres.
- Les scénarios simulés peuvent être observés dynamiquement et en temps réel via la représentation graphique du modèle. Cela permet de décrire plus visuellement le comportement du système.

- Il est assez aisé d'affiner le modèle de manière incrémentale et de voir l'effet des modifications sur le système.

Pour prendre en compte les phénomènes aléatoires, on lance plusieurs répliques. Ces dernières se basent sur des suites de nombres pseudo aléatoire présentes dans le simulateur. Elles permettent de prendre en compte les phénomènes aléatoires qui régissent le modèle et d'appliquer cette variabilité dans les simulations afin d'avoir un comportement plus proche de la réalité. De ce fait, les résultats obtenus seront plus fiables.

La SED est une technique qui est largement utilisée dans l'industrie depuis les années 1950.

La simulation à base d'agents

La simulation à base d'agents est une méthode qui est plus récente, elle est développée au début des années 2000. Son essor est motivé par une volonté de la communauté à aller plus loin dans la modélisation de système complexe. Cela est rendu possible grâce aux avancées techniques et technologiques qui permettent d'avoir plus de puissance de calcul car la simulation à base d'agents nécessite des ressources plus importantes que les autres types de simulation (Borshchev, 2013a). Comme soulevé par Law et al., (Law et al., 2000) il est difficile de trouver une définition claire pour la simulation à base d'agents. Tout d'abord, qu'est-ce qu'un agent ? C'est une entité autonome qui peut interagir avec son environnement (incluant les autres agents), et utiliser les informations recueillies pour prendre des décisions. Les agents ont des attributs, et un panel de règles qui définissent leurs comportements. Ces comportements peuvent évoluer dans le temps et s'adapter à l'environnement de manière dynamique. Si on suit cette définition, la simulation à base d'agents est de ce fait, un cas particulier de la SED, où les comportements sont définis directement au niveau de l'entité et non pas au niveau des blocs, par lesquels l'entité passe.

La modélisation basée sur les agents se concentre sur les composants actifs individuels d'un système. Elle s'oppose à la fois à l'approche plus abstraite de la dynamique des systèmes et à la méthode des événements discrets axée sur le processus et les enchainements d'activités du système¹.

Avec la modélisation basée sur les agents, les entités actives, les agents donc, doivent être identifiées et leurs comportements définis. Il peut s'agir de personnes, de ménages, de véhicules, d'équipements, de produits ou d'entreprises, selon ce qui est pertinent pour le système. Des connexions entre elles sont établies, des variables environnementales sont définies et des simulations sont effectuées. La dynamique globale du système émerge alors des interactions entre les nombreux comportements individuels.

Ainsi, la simulation à base d'agents est intéressante dans plusieurs cas (Borshchev, 2013b) :

- Si les entités du système interagissent entre elles et avec l'environnement (des colis, des personnes dans un aéroport, des patients dans un hôpital...).

1. <https://www.anylogic.com/use-of-simulation/agent-based-modeling/>

- Lorsqu'il est important pour les entités d'apprendre et d'adapter leurs comportements.
- Lorsque les mouvements de l'entité dépendent de sa perception de la situation et de son environnement plutôt que de règles fixées au préalable.

Les systèmes dynamiques

La simulation utilisant les systèmes dynamiques est une méthode créée dans le milieu des années 1950 par Jay Forrester (Forrester, 1971). Elle permet d'étudier des systèmes dynamiques en se basant sur des lois issues de la physique (Borshchev, 2013a). Ces méthodes se basent sur plusieurs principes concernant les stocks, les flux et les boucles de retours (Sterman, 2010) :

- Les stocks sont les endroits dans le système où il y a des accumulations du flux, ils sont source de déséquilibre dans le système.
- Les boucles de retour sont au coeur des systèmes dynamiques. Dans le cas d'un centre d'appels, cela correspondrait aux appelants qui rappellent parce qu'ils n'ont pas obtenu de réponse.
- les événements et les décisions sont vus comme un flux continu d'information qui vient alimenter le modèle.

Là où les systèmes à base d'agents agissent au niveau de l'agent ou la SED au niveau des événements, les systèmes dynamiques utilisent un niveau d'abstraction plus élevé.

Un intérêt visuel et didactique pour les experts qui rend la simulation attrayante en santé

L'usage de la simulation dans le domaine de la santé tend à aller en augmentant au fil des années (Salleh et al., 2017). Salleh et al. proposent une revue de littérature de l'usage de la simulation dans le domaine de la santé. Leur but est de présenter une vue d'ensemble des revues de littérature sur l'utilisation de la simulation entre 1990 et 2017. Les auteurs identifient quatre principaux domaines d'applications en santé :

1. Opération de santé et conception de système : Utilisation de la simulation pour la gestion des ressources, leur dimensionnement, pour l'optimisation des flux et la prévision de la demande de soin.
2. Aide à la décision médicale : Évaluer les effets à court et long terme de certains programmes de soin.
3. Modélisation de maladies infectieuses : Applications d'épidémiologie pour prédire la diffusion des épidémies et pandémies ainsi que leurs conséquences économiques.
4. Études diverses : Simulations d'événements ponctuels (coupe du monde, attaques terroristes) et de leurs effets sur le système de santé. On trouve aussi des comparaisons de diverses techniques de simulation.

La grande majorité des revues de littérature étudiées par ces auteurs se focalisent uniquement sur la gestion des opérations en santé, ce qui est aussi le domaine dans lequel nous travaillons. On trouve des travaux de simulation pour évaluer les services des urgences (Aboueljinane et al., 2013 ; Gul et al., 2015 ; Paul et al., 2010), pour l'amélioration du parcours patient et la gestion des flux de patients

(Crema et al., 2013; Sobolev et al., 2011; Van Sambeek et al., 2010), pour gérer la chaîne d’approvisionnement logistique des hôpitaux (Beliën et al., 2012; Kammoun et al., 2014), la gestion des lits d’hôpitaux (Baru et al., 2015), la gestion et le dimensionnement des ressources (Isern et al., 2016), ou plus spécifiquement concernant la planification des salles d’opération (Cardoen et al., 2010).

La simulation dans le domaine de la santé est donc très utilisée depuis de nombreuses années car sa capacité à permettre de traduire des échanges et des interactions complexes est très intéressante. En effet, le domaine médical est constitué de nombreux facteurs humains qui viennent générer beaucoup d’aléatoires. Ce qui vient rendre très pertinent l’usage de la simulation pour traiter les problèmes. De plus, la simulation a un intérêt visuel non négligeable pour l’explication des résultats auprès des experts. Le fait de pouvoir rejouer les situations pour en expliquer les tenants et les aboutissants présente un avantage didactique et pédagogique.

Comme nous l’avons expliqué dans la section 3.2.4 nous avons choisi de travailler avec des techniques de simulation. Comme vu ci-dessus, il en existe plusieurs types et pour cette étude il nous a paru pertinent d’utiliser la SED. Notamment la simulation à base d’agents a été écartée car les appels ne sont pas des entités qui interagissent entre elles.

Des hypothèses pour la modélisation

Les simulations ont été réalisées sur un ordinateur équipé d’un processeur Intel® Core™ i5-6200U CPU à 2,3 GHz et d’une mémoire RAM de 8 Go. Le logiciel utilisé dans ce manuscrit est le logiciel Witness®, développé par Lanner² et qui fait partie des leaders sur le marché des moteurs de simulation à événements discrets. Un visuel du modèle de simulation est disponible dans la figure 5.2.

Dans le chapitre 2 nous avons détaillé les différents types d’appels qui sont reçus par les centres d’appels d’urgence. Pour rappel, on distingue les appels de la population, les appels provenant des équipes dépêchées sur le terrain, ainsi que les appels des autres acteurs de la chaîne des urgences. Les données du fichier de téléphonie nous permettent de savoir quel est le numéro composé pour joindre le centre d’appels. Toutefois nous n’avons pas les informations nécessaires pour différencier clairement les appels de la population des appels des intervenants. Comme précédemment décrit dans le chapitre 1.2, il y a plusieurs types d’appels. Pour la suite de cette étude, nous faisons l’hypothèse que tous les appels entrants sont considérés comme étant des appels de la population entrés par le 15 ou la PDSA. Dans un premier temps, nous ne différencions pas les appels bilans des appels de la population. Ce choix peut avoir un impact sur la précision du modèle, car les appels bilans n’ont pas forcément le même profil d’arrivée que les appels entrants. Les appels bilans dépendent directement du choix d’envoyer des moyens et des intervenants lors d’un primo appel.

Une deuxième hypothèse concerne les appels sortants. Dans la littérature scientifique relative aux centres d’appels d’urgences, il n’y a pas à notre connaissance de travaux

2. <https://www.lanner.com/fr-fr/technologie/witness-simulation-software.html>

dans lesquels les appels sortants sont considérés. Ceci constitue une faiblesse importante dans les études passées des centres d'appels urgents comme démontré en section 4.2.2. Nous proposons dans notre travail d'en tenir compte et de les inclure dans notre modélisation. Bien qu'il existe plusieurs types d'appels sortants, vers les intervenants, vers les établissements de santé, etc. (voir figure 1.9), nous ne distinguons néanmoins pas la nature de l'appel sortant ni son destinataire. Nous avons utilisé les données de téléphonie pour extraire le profil des appels sortants ainsi que la distribution de leur durée. Toutefois, nous n'avons pas l'information de la nature du destinataire, donc nous n'avons pas pu spécifier la génération ou la durée de l'appel sortant en fonction de son destinataire. C'est pourquoi nous avons considéré les appels sortants comme une grande famille d'appels sans définir de spécificités en son sein.

Concernant les appels internes, soit les appels qui sont générés entre les opérateurs du centre d'appels, nous avons uniquement considéré les appels de l'ARM vers le MR. D'après les données dont nous disposons, ces appels correspondent à 74% de la totalité des appels internes. 13% des appels internes sont de l'ARM vers le MR PDSA, nous les avons aussi considérés, et nous avons négligé les 13% restants qui concernent des appels d'un ARM vers un autre ARM ou d'un MR vers d'autres ressources.

Réplifications, Horizon, Explications

Définir le nombre de réplifications

Comme nous avons déjà pu le mentionner, les réplifications permettent de modifier les suites pseudo-aléatoires du simulateur et ainsi d'avoir de nouveaux tirages pour les valeurs aléatoires du système. Dans notre cas, nous avons de l'aléatoire en ce qui concerne l'arrivée des appels : le nombre d'appels par tranches de 30 minutes. Il y a aussi de l'aléatoire sur le niveau de gravité de l'appel, sur sa durée de traitement, sur la possibilité de générer un appel interne. Les réplifications nous permettent de couvrir tout le spectre des valeurs possibles lorsqu'il y a de l'aléatoire. L'objectif est d'avoir un nombre significatif de 100 réplifications afin de couvrir tous les types de journées « probables » d'un centre d'appels, les meilleurs comme les pires.

Définir l'horizon de simulation

Comme nous avons déjà pu l'expliquer dans le chapitre précédent, nous avons limité le périmètre d'études aux dimanches, car ce sont les journées les plus chargées avec une augmentation significative du nombre d'appels reçus. De ce fait, nous avons fixé l'horizon de simulation à 24h. Chaque simulation correspond à un dimanche dans un centre d'appels. Lorsqu'on lance une nouvelle réplification, cela revient à simuler un nouveau dimanche. Dans notre cas, nous avons travaillé sur une base de 100 réplifications, donc de 100 dimanches, soit environ deux ans de dimanches simulés.

5.1.2 Sous-étape 3.b : Évaluation de l'AS-IS : Un Modèle conceptuel pour guider la construction du modèle

La figure 5.3 présente le processus complet suivi par notre modèle de simulation : une version agrandie est disponible en annexe C. Nous décrivons ci-dessous, pour chacune des activités, les données utilisées et hypothèses considérées. Afin d'être le plus lisibles possible, nous décomposons ces explications par macro-activités. (cf. figure 5.4)

Macro-activité 1 : La gestion des appels entrants

Cette activité ne concerne que l'acteur ARM. Dans cette étude nous considérons deux types d'appels entrants, les appels provenant du numéro d'urgence nationale français, le 15, et les appels provenant de la permanence des soins ambulatoires (39 66 en Occitanie, 116 117 pour d'autres régions comme en Pays de la Loire, Normandie et en Corse). L'arrivée des appels est paramétrée en se basant sur le profil des données historiques comme expliqué dans le chapitre 4. Nous avons segmenté la journée en créneaux de 30 minutes. Ce choix de segmentation a été fait, car il permet d'avoir une vision assez précise et détaillée de l'évolution de la performance du centre d'appels au cours de la journée. Habituellement, les études mesurent les performances d'un centre d'appels à l'échelle de la journée au niveau global, toutefois la variabilité du nombre d'appels au cours de la journée est assez importante. C'est pourquoi il nous a paru intéressant de se baser sur des créneaux de 30 minutes pour étudier les performances du SAMU.

Sur chaque créneau nous avons regardé le nombre d'appels arrivés sur les 8 dimanches disponibles. Cet échantillon de 8 jours n'étant pas statistiquement significatif, il n'est pas possible ici d'identifier une distribution théorique satisfaisante pour l'arrivée des appels. Pour modéliser le caractère aléatoire du nombre d'appels sur chaque créneau horaire, nous avons choisi d'utiliser une loi de distribution uniforme s'étendant entre le nombre minimum et le nombre maximum d'appels reçus durant les 8 jours étudiés. Cela nous permet d'avoir des profils d'appels simulés qui sont cohérents par rapport aux données réelles. Chaque simulation ayant un horizon de 24 heures (1 journée pleine), cela va aussi permettre de lancer un nombre de répliques significatif, car l'arrivée des appels sur chaque créneau, même si elle est aléatoire, ne se produira qu'une seule fois au cours d'une même simulation. L'ensemble des répliques permettra donc de tirer plusieurs valeurs aléatoires d'arrivée d'appels sur chaque créneau et ainsi permettra de présenter les résultats sous la forme d'un intervalle de confiance. À titre d'exemple, la figure 5.5 montre le profil des appels 15 réels sur lesquels nous nous sommes basés. On peut remarquer deux pics d'appels dans une journée de dimanche autour de 9h et un autre autour de 16h.

Lorsqu'un appel arrive au SAMU, il arrive dans une file d'attente, l'appel va sonner au SAMU et pendant ce temps là pour l'appelant, le disque de bienvenue au SAMU déroule durant plusieurs secondes. Ainsi, il y a un temps d'attente minimal avant que l'appel puisse être décroché. Ce temps a été paramétré à 3 secondes en se basant sur

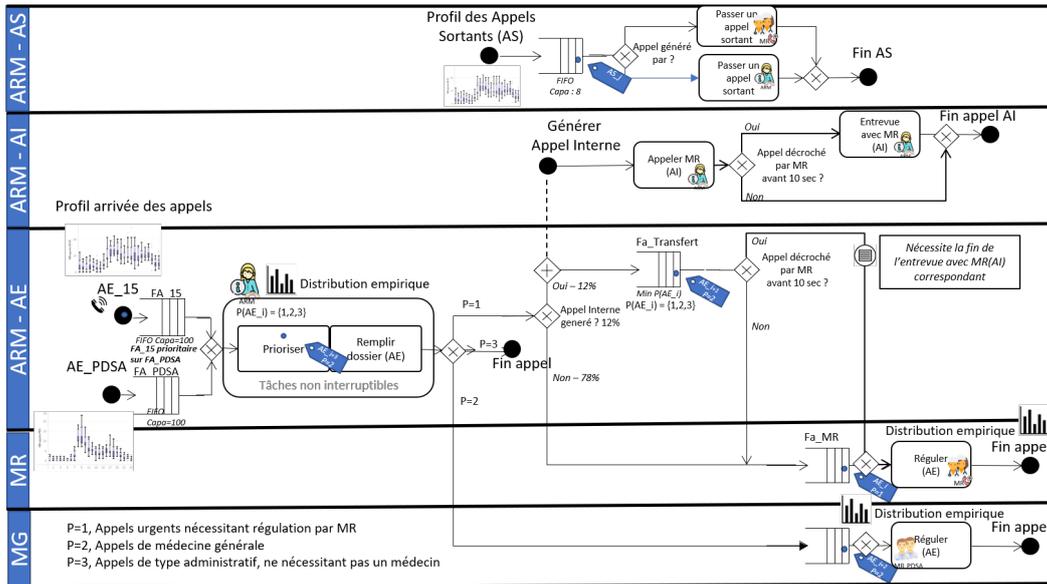


Figure 5.3 – Modèle conceptuel de la gestion des appels. (Voir version agrandie en annexe)

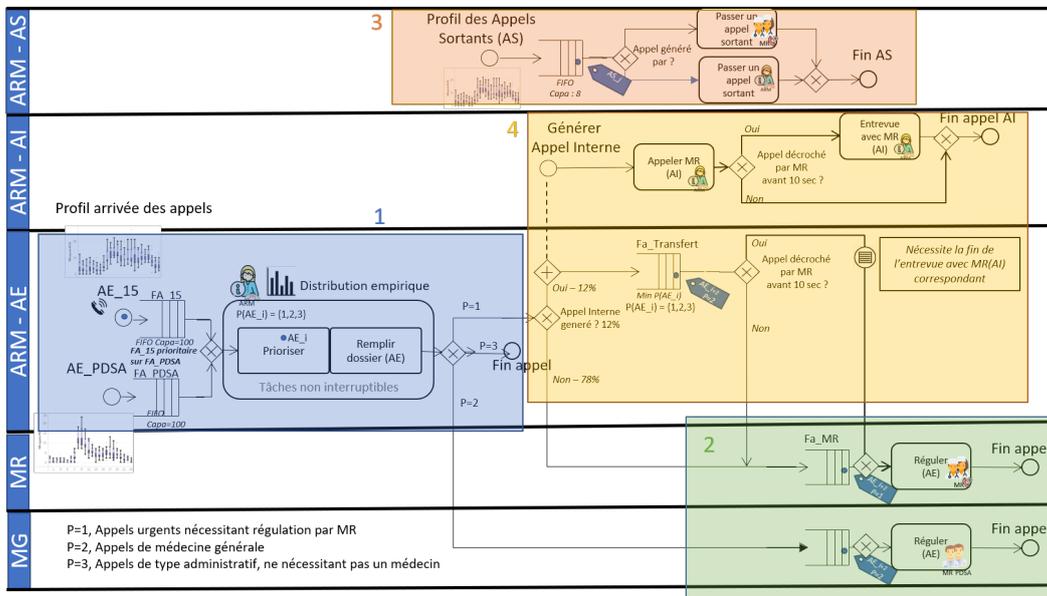


Figure 5.4 – Modèle conceptuel de la gestion des appels : Définition des macro-activités

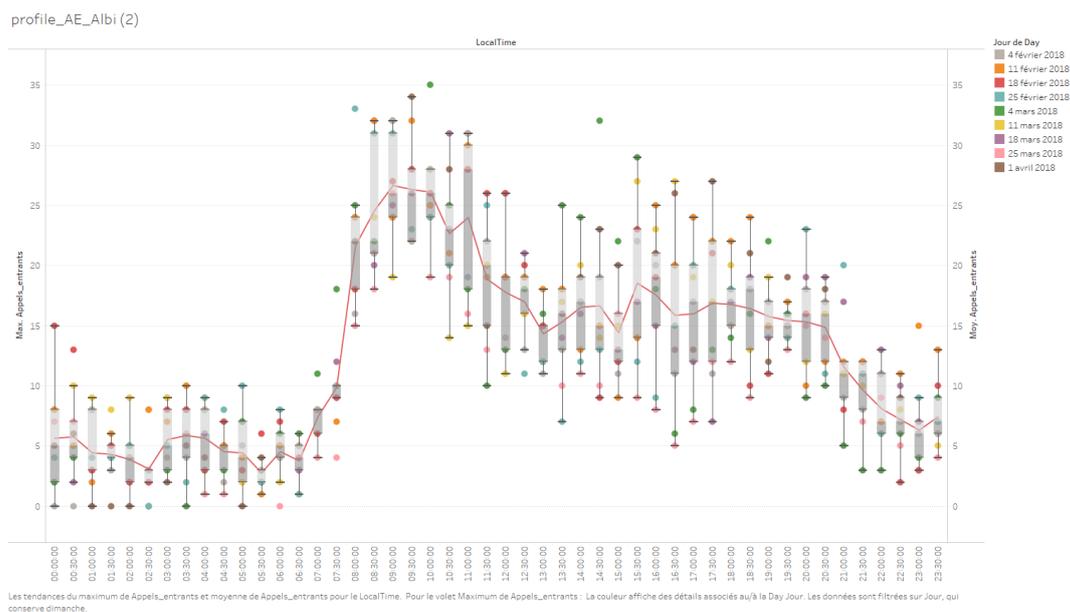


Figure 5.5 – Profil d'arrivée des appels entrants (un dimanche)

les données indiquées par les experts et notamment par les conclusions du rapport MARCUS (Moutard et al., 2019).

Concernant les files d'attente, il y a plusieurs questions qui se posent :

- Est-ce qu'il y a une file d'attente par canal d'entrée ? Dans ce cas, y a-t-il une file d'attente prioritaire par rapport à l'autre ?
- Est-ce qu'il y a une seule file d'attente gérée avec une loi de type « premier arrivé, premier servi » (FIFO) ?
- Est-ce qu'il y a une seule file d'attente avec une priorité pour les appels provenant du 15 par rapport à ceux provenant de la PDSA ?

Ces choix de modélisation sont à discuter avec les responsables des centres d'appels pour mettre en place ce qui correspond à l'usage dans le centre d'appels. Toutefois, il reste intéressant de s'interroger sur l'effet de ces dispositions de file d'attente sur les performances. Dans une étude préliminaire, nous avons testé différentes files d'attente (Petitdemange et al., 2018). Les résultats montrent que s'il peut paraître logique de prioriser les appels du 15 par rapport à ceux de la PDSA cela peut poser problème dans certains cas. En effet, bien qu'ils soient moins fréquents, des appels urgents qui nécessitent une régulation par un MR peuvent arriver par le canal de la PDSA. Ainsi la priorisation uniquement par la provenance de l'appel, n'est pas suffisante pour certains cas isolés.

Dans notre cas nous avons d'abord paramétré le modèle comme ayant deux files d'attente distinctes en fonction de la provenance de l'appel. Nous avons fait l'hypothèse qu'une priorité supérieure est appliquée pour la file d'attente des appels 15 (*FA_15*).

Une fois l'appel dans la file d'attente, il sera décroché dès lors qu'un ARM sera disponible. Durant ce premier temps de communication avec l'ARM, plusieurs tâches sont effectuées les unes à la suite des autres ces tâches ont déjà été présentée dans la section 1.2.5. Ces sous-tâches sont regroupées sous une macro-tâche qui correspond

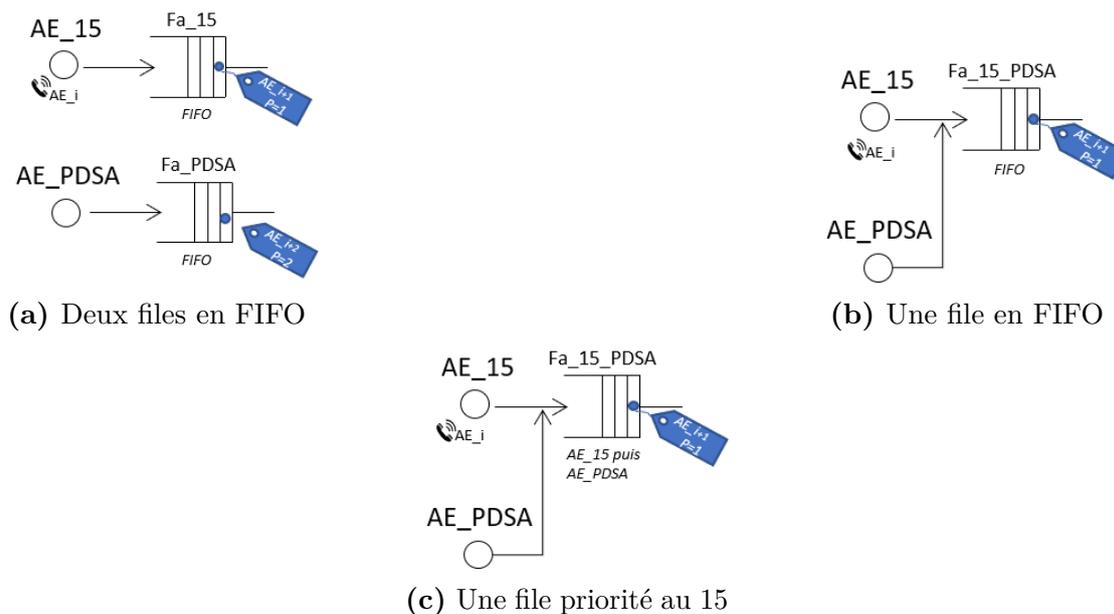


Figure 5.6 – différents types de file d’attentes

à ce premier temps de communication avec l’ARM. Dans nos données de téléphonie, nous n’avons pas d’indication pour identifier les temps de chaque sous-tâche. Cet enchaînement est bien souvent non interrompu, cela signifie que lorsqu’un ARM prend un appel, il effectue tout ou partie de la chaîne en fonction de la gravité de l’appel qui va donner lieu à sa priorisation (P1, P2 ou P3 comme présenté dans la section 1.2.5. C’est pourquoi nous avons modélisé le temps de communication de l’ARM d’un seul bloc.

Dans un premier temps, notre modèle est configuré pour que cet enchaînement d’activité corresponde à une seule macro-tâche qui ne peut être interrompue ou mise en pause. C’est-à-dire que lorsqu’un ARM décroche un appel, il le traite en entier avant d’être de nouveau disponible. C’est lors de cette étape que l’appel va être priorisé : L’ARM va lui attribuer un niveau de gravité qui va décider du devenir de l’appel. La Table 5.1 présente la répartition statistique de la gravité des appels par rapport à leur provenance. On constate que des appels de priorité 3 occupent 78% de la totalité des appels que ce soit pour les appels provenant de la PDSA ou du 15. Pour rappel, ce sont des appels de nature non médicale (adresse de la pharmacie de garde, coordonnées d’un médecin ou d’un dentiste, etc.), qui ne requièrent par de temps de communication avec un MR. Il serait alors utile d’assurer un traitement automatisé de ces appels, ce qui pourrait être un levier d’amélioration dans la gestion des appels d’urgence. Ensuite, d’après nos données, dans les appels provenant du 15, 18% sont des appels qui nécessitent une régulation auprès d’un MR et 2% auprès d’un généraliste. Les chiffres s’inversent pour les appels provenant de la PDSA. Cela peut paraître peu, toutefois parmi les appels entrants il y a aussi les appels bilans des équipes d’effecteurs (SMUR, pompiers, etc.) sur place qui parfois transmettent le bilan des interventions aux ARM plutôt qu’au MR. Ces données sont à nuancer : En effet, cela voudrait dire qu’1 appel sur 5 fait l’objet d’une régulation médicale, ce qui semble peu dans la mesure où tous les appels médicaux sont supposé faire l’objet

d'une régulation auprès du MR. Cette donnée est obtenue en analysant les données de téléphonie. Pour chaque identifiant de communication unique correspondant à l'appel d'un patient, nous identifions les ressources qui interviennent sur la gestion de cet appel. Toutefois, cela implique l'hypothèse suivante : tous les appels concernant un même dossier patient sont rassemblés sous le même identifiant de communication. Si ce n'est pas le cas, alors il se pourrait que la part des appels qui génère une communication avec le MR soit plus grande, mais pour vérifier cette hypothèse nous devons disposer du numéro de dossier patient et le relier aux données de téléphonies, ce qui à ce jour n'est pas le cas.

Tableau 5.1 – Répartition de la priorité des appels en fonction de la provenance des appels.

Priorité	P=1	P=2	P=3
	Régulation	Permanence des soins	Conseil
AE_15	18%	2%	62%
AE_PDSA	2%	22%	76%

Macro-activité 2 : La Régulation des appels

Cette activité de régulation concerne le MR pour les appels de priorité $P = 1$, et le MG pour les appels de priorité $P = 2$. Une fois les appels triés par l'ARM, ils sont envoyés dans la file d'attente idoine en attendant d'être régulés par un médecin. S'il n'y a pas de système de gravité codifié par l'ARM, les appels sont traités en FIFO dans leurs files d'attente respectives, à l'exception du MR lorsqu'il termine une communication en appel interne avec un ARM, il traite directement l'appel entrant en lien avec cet appel interne avant de continuer à gérer sa file d'attente.

Macro-activité 3 : La gestion des appels sortants externes

Les appels sortants sont des appels émis par les agents du centre d'appels vers l'extérieur. Ceux-ci peuvent correspondre à des appels vers les effecteurs pour envoyer une ambulance, le SMUR, une équipe des pompiers, ou encore un appel vers les établissements de santé pour savoir s'ils ont une place disponible pour un patient. Cela peut aussi être pour rappeler des patients (cette pratique est assez courante, notamment en ce qui concerne la PDSA). Plusieurs questions se posent en ce qui concerne les appels sortants :

- Qu'est-ce qui déclenche un appel sortant ?
- Est-ce que les appels sortants dépendent des appels entrants ?
- Est-ce que les appels sortants sont prioritaires par rapport aux appels entrants ?

Si on utilise uniquement les données de téléphonie, il est difficile d'identifier vers quels types de correspondants l'appel est passé. Pour cela il faudrait identifier les numéros récurrents, savoir à quels numéros correspondent les partenaires de l'urgence pour connaître la nature de l'appel sortant. Nous n'avons pas obtenu la liste des numéros connus des effecteurs, il nous a donc pas été possible de filtrer ces appels

dans la base de données. Une autre piste intéressante serait d'être capable de croiser les données d'intervention d'un appel entrant qui précède un appel sortant pour avoir plus d'information sur le moyen envoyé, savoir si cet appel sortant est relié au dossier de l'appel entrant. Ces informations permettraient d'estimer la probabilité qu'un appel entrant génère un appel sortant, mais aussi d'identifier les règles métiers qui font qu'un appel sortant est déclenché. Toutefois l'accès aux dossiers de régulation médicale est complexe et n'est pas possible en état pour des raisons de confidentialité. Ainsi, nous avons fait l'hypothèse que les appels sortants étaient générés en suivant leur profil d'émission historique. De la même manière que les appels entrants, nous avons analysé le nombre d'appels sortants émis sur chaque créneau de 30 minutes durant la journée. Afin d'avoir un profil type des émissions des appels sortants. La limite de cette technique est qu'elle ne prend pas en compte les interactions potentiellement existantes entre appels entrants et appels sortants. Néanmoins elle a l'avantage de considérer les appels sortants ce qui n'est pas le cas des études menées jusqu'à présent sur les centres d'appels d'urgence (cf section 5.1.1).

En conséquence, les appels sortants sont générés en suivant le profil présenté dans la figure 5.8. On constate que le profil des appels sortants a une forme similaire à celui des appels entrants, ce qui corrobore le fait que la génération des appels sortants dépend des appels entrants. Lorsqu'un appel est généré, il arrive dans une file d'attente jusqu'à ce qu'un agent puisse le composer. Le temps d'attente qui en découle est donc artificiel puisqu'il correspondrait au fait qu'un agent sait qu'il doit passer un appel, mais il gère d'autres activités au préalable. La durée des appels sortants est définie en suivant la distribution empirique issue des données historiques (cf. figure 5.7).

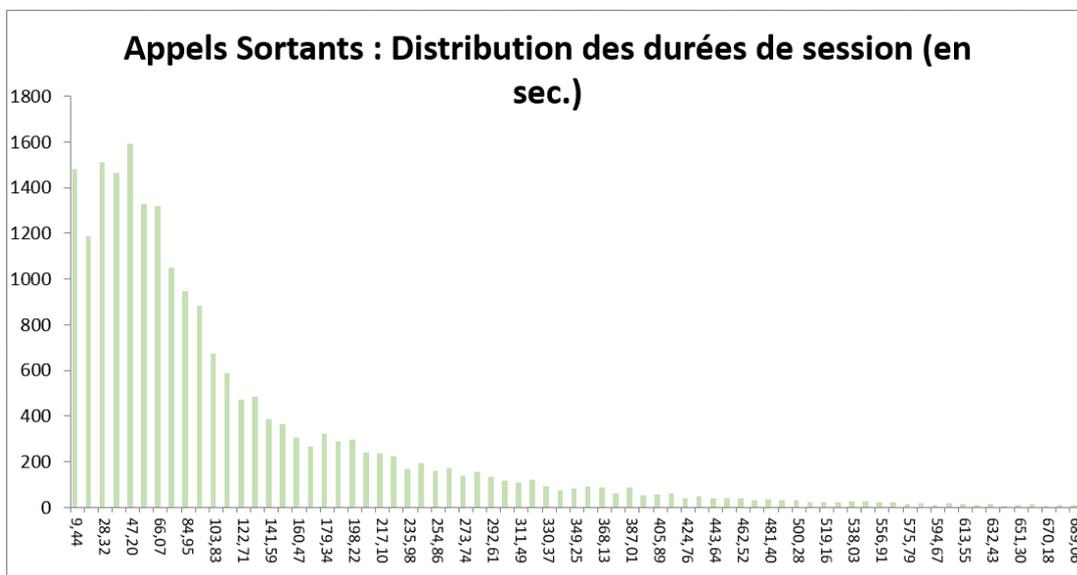


Figure 5.7 – répartition des durées de session pour les appels sortants

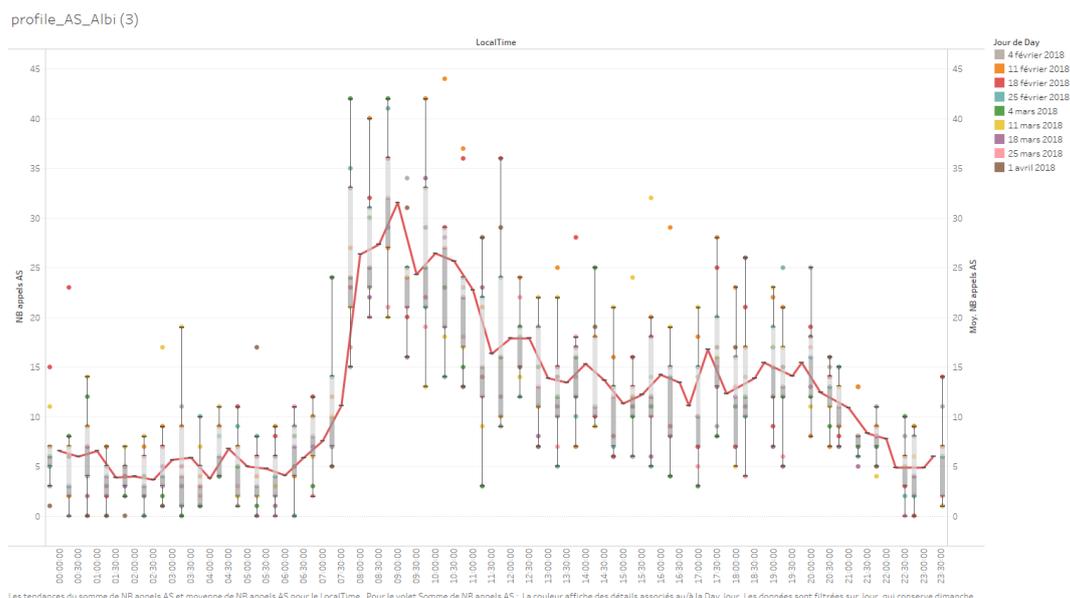


Figure 5.8 – Profil d'émission des appels sortants

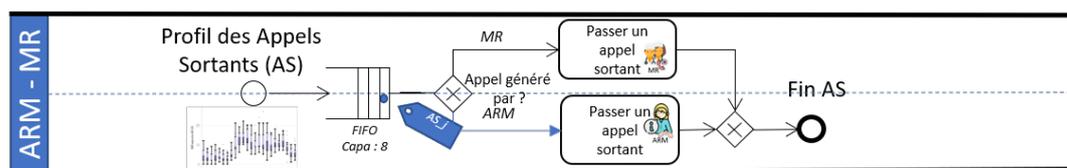


Figure 5.9 – Gestion des Appels sortants (AS)

Macro-activité 4 : La gestion des appels internes

Certains appels entrants peuvent donner lieu à un appel sortant interne (cf. figure 5.10) (une communication entre personnels du centre d'appels). Cela concerne 53% des appels de priorité P1, et 60% des appels de priorité P2. Dans ce cas, l'appel entrant est mis en attente (il est maintenu en ligne), l'ARM passe son appel interne. En accord avec les observations terrains réalisées, nous avons fait l'hypothèse que si le MR n'est pas disponible dans les 10 secondes, l'appel entrant est directement transféré dans la file d'attente du MR et l'appel interne est abandonné. L'ARM va alors soit passer un autre appel sortant s'il y a lieu, soit décrocher des appels.

Un temps de latence pour les agents

À travers nos observations sur le terrain et nos discussions avec les experts, nous avons constaté qu'après un appel, un agent peut prendre quelques secondes avant de décrocher directement un nouvel appel. Ce temps peut correspondre au temps nécessaire pour clôturer le dossier de régulation. Afin de reproduire cet effet dans notre modèle nous avons décidé d'ajouter un temps de latence après chaque appel dont le traitement vient de finir pour l'agent qui vient de raccrocher. Nous avons fixé ce temps en nous basant à la fois sur nos observations et sur l'avis des experts. Ce temps correspond à une loi uniforme qui s'étend entre 2 et 6 secondes. Ce temps

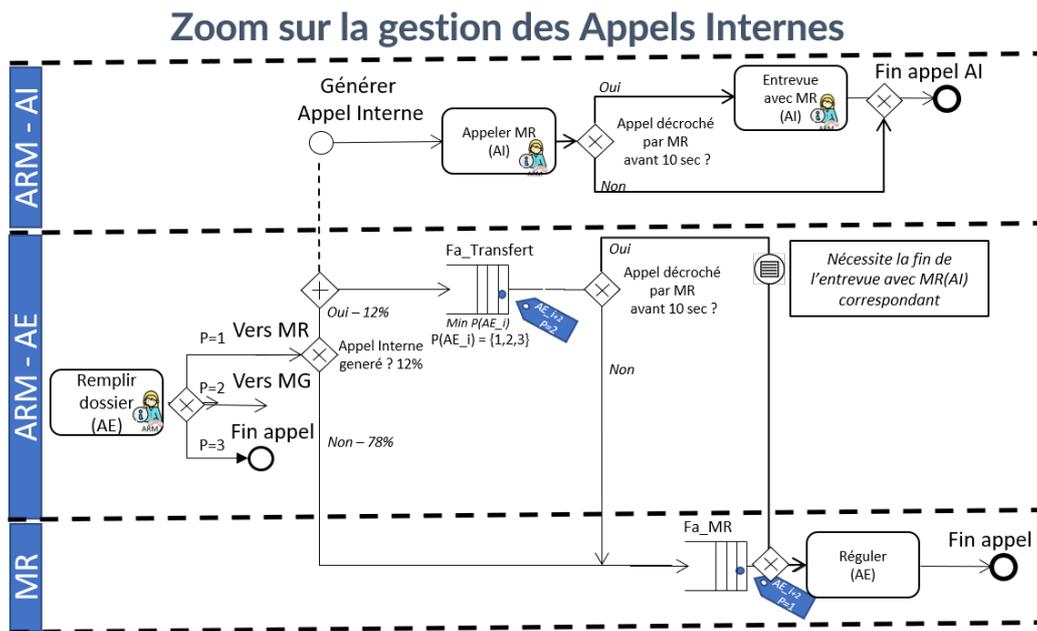


Figure 5.10 – Gestion des Appels internes (AI)

de latence est aussi appelé *wrap-up* il est d'ailleurs pris en compte dans les solutions de téléphonie avancée et ce temps est paramétrable. En effet l'agent peut prendre ce temps de latence ou l'annuler en fonction de son choix. Il serait toutefois intéressant de voir l'impact de la variation de ce temps de latence sur les performances du centre d'appels.

5.1.3 Sous-étape 3.c : Validation du modèle

Valider le profil d'arrivée des appels

Pour valider le profil d'arrivée, nous avons comparé nos données réelles à nos données simulées. Sur chaque créneau de 30 minutes, nous avons relevé le nombre d'appels entrants sur les 8 échantillons réels ainsi que sur chacune des 100 répliques. Pour chaque créneau horaire, nous avons mesuré l'écart sur le nombre minimum, la moyenne et le nombre maximum d'appels pour le profil réel et simulé. Les résultats sont présentés dans la figure 5.11. Au maximum on observe un écart de 7 appels sur le nombre d'appels maximum par jour. Nous pouvons donc conclure que le profil d'arrivée des appels simulés suit la même distribution que celui des appels réels.

Comparer les performances du modèle et de la réalité

Pour évaluer la justesse de notre modèle, nous avons comparé le temps d'attente des appels moyens, médians, et maximums ainsi que la déviation standard sur chacun des créneaux. Nous avons comparé les valeurs de notre échantillon de 8 dimanches avec celles de nos 100 répliques. On observe des écarts entre le réel et le simulé que nous allons expliquer et justifier. Les résultats sont présentés dans le tableau

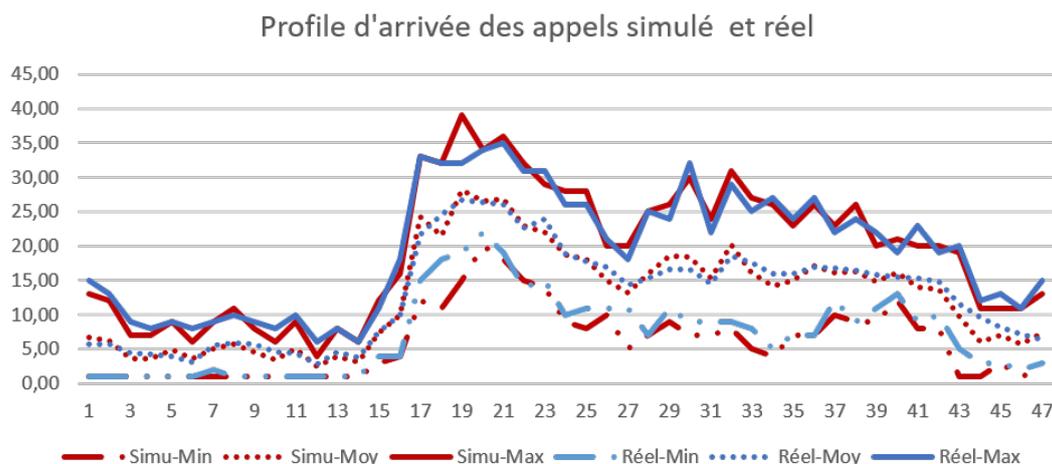


Figure 5.11 – Écarts entre les profils d'arrivées des appels réels et simulés

5.2. On note des écarts en moyenne qui varient entre 0,5 seconde et 6,7 secondes en fonction de l'heure de la journée. Ceci peut être problématique, car la moyenne de décrochée des appels réels varie entre 5 et 11,5 secondes, donc 6,7 secondes est un écart assez significatif. Les écarts au niveau de la médiane varient eux entre 0,5 et 1,5 seconde ce qui est tout à fait acceptable. C'est au niveau des temps d'attente maximums qu'on constate la plus grande variabilité. En effet, notamment durant les horaires de nuit profonde entre minuit et 7h00 les résultats de la simulation sont optimistes par rapport à la réalité. En effet le temps d'attente maximal est plus faible dans notre simulation qu'il n'a pu l'être en réalité sur les 8 dimanches étudiés. On obtient un temps maximal d'attente des appels simulé entre minuit et 07h00 qui est au maximum de 17 secondes. Dans les données réelles, les appels ayant attendu plus de 17 secondes correspondent à moins de 7% de la totalité des appels sur cet horizon. Notre modèle ne prend effectivement pas en compte ces appels qui doivent avoir une raison particulière d'avoir attendu parfois jusque 2 minutes 30. Typiquement, si on regarde le cas de l'appel qui a attendu 162 secondes, on constate que c'est un appel qui a appelé 2 fois avant d'être décroché par un ARM la troisième fois, or nous ne sommes pas en mesure d'expliquer pourquoi les ARM n'ont pas décrochés les deux premières fois, peut être ils n'étaient pas à leur poste de travail à ce moment-là. Nous pouvons encore améliorer la finesse de notre modèle afin qu'il soit en mesure de prendre en compte ces comportements particuliers qui sont à la marge du fonctionnement habituel du SAMU.

Les simplifications effectuées traduisent les comportements les plus fréquents d'un centre d'appels pour le traitement des appels entrants (voir figure 5.12). Il ne prend pas en compte les appels qui abandonnent, ou encore les personnes qui vont appeler, ne pas être décrochées, puis rappeler. Le fait de ne pas distinguer les appels entrants de la population et les appels entrants des équipes qui interviennent sur le terrain rajoute aussi de la variabilité. Ce sont pour toutes ces raisons notamment que l'on constate des écarts entre notre modèle et la réalité. Toutefois, bien que perfectible, ce modèle est une base de travail intéressante pour venir challenger le traitement des

appels entrants de la population. Pour la suite lors des évaluations des scénarios, les résultats seront comparés aux performances du modèle simulé.

Tableau 5.2 – Statistiques descriptives des performances du modèle comparées aux performances réelles

Créneau	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
Réel	Dev_std	18,0	6,1	5,0	8,3	12,1	4,2	8,3	19,8	18,2	11,0	14,6	16,5	12,8	17,8	14,3	9,9	15,1	15,6	14,4	10,6	27,0	26,6	18,0	12,5
	Max	152,0	34,0	31,0	76,0	90,0	17,0	56,0	163,0	220,0	84,0	181,0	248,0	115,0	151,0	112,0	96,0	134,0	119,0	128,0	57,0	249,0	245,0	102,0	109,0
	Moy	9,0	6,7	7,0	6,8	8,1	5,3	5,7	8,2	7,2	5,5	6,6	7,3	7,2	10,5	9,1	7,0	8,2	8,9	8,8	7,6	11,5	11,0	10,2	8,0
	Médiane	5,0	6,0	6,0	6,0	5,0	5,0	6,0	5,0	4,0	5,0	5,0	5,0	3,0	6,0	5,0	5,0	5,0	5,0	5,0	5,0	1,0	3,0	4,0	5,0
Simu	Dev_std	0,8	0,3	0,5	0,6	0,5	0,6	0,5	0,8	10,5	15,1	11,7	8,0	5,5	1,4	10,1	6,9	2,8	7,2	2,3	1,8	4,6	0,8	0,4	0,4
	Max	16,3	5,0	11,9	14,5	12,3	15,7	12,7	16,7	343,7	590,5	413,3	404,9	278,4	37,7	570,5	354,5	61,5	369,2	71,1	77,8	173,6	21,9	15,3	12,3
	Moy	4,6	4,5	4,5	4,5	4,5	4,5	4,5	4,6	5,8	6,5	6,1	5,1	4,9	4,7	5,1	5,0	4,8	4,9	4,7	4,6	4,8	4,6	4,5	4,5
	Médiane	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5
Écarts	Dev_std	17,2	5,8	4,4	7,7	11,6	3,6	7,9	19,0	7,7	-4,1	2,9	8,6	7,3	16,4	4,3	3,1	12,4	8,4	12,1	8,8	22,4	25,8	17,5	12,1
	Max	135,7	29,0	19,1	61,5	77,7	1,3	43,3	146,3	-123,7	-506,5	-232,3	-156,9	-163,4	113,3	-458,5	-258,5	72,5	-250,2	56,9	-20,8	75,4	223,1	86,7	96,7
	Moy	4,5	2,2	2,4	2,3	3,6	0,8	1,2	3,6	1,5	-1,0	0,5	2,2	2,3	5,9	4,1	2,0	3,4	4,0	4,1	3,0	6,7	6,5	5,7	3,5
	Médiane	0,5	1,5	1,5	1,5	0,5	0,5	1,5	0,5	0,5	-0,5	0,5	0,5	-1,5	1,5	0,5	0,5	0,5	0,5	0,5	0,5	-3,5	-1,5	-0,5	0,5

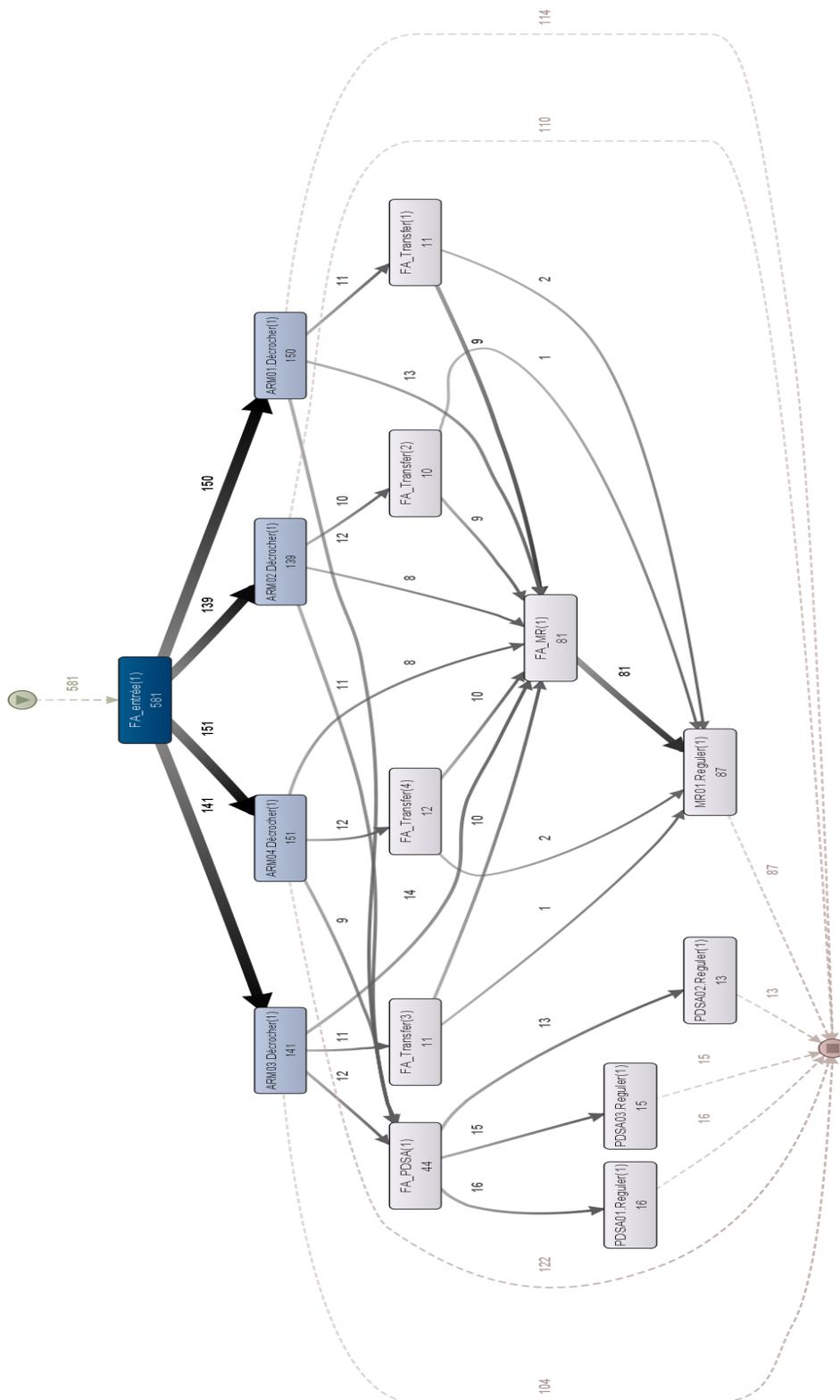


Figure 5.12 – Différents chemins pouvant être empruntés par un appel entrant dans le modèle de simulation

5.2 Étape 4 : Améliorer ; Scénarios de Simulation, Analyses et Résultats

5.2.1 Comment hiérarchiser les types d'appels, quel impact sur les performances ?

Nous avons identifié trois familles d'appels au sein d'un centre d'appels d'urgence : les appels entrants (AE), les appels sortants (AS) et les appels internes (AI). Nous nous sommes questionnés sur les règles de priorité qui peuvent exister entre ces appels. Comme il y a 3 types d'appels à trier et à prioriser, nous en avons donc 3!, soit 6 possibilités à envisager pour l'ARM et pour le MR, soit 12 cas possibles. Fixer ces priorités va permettre de définir les choix des ressources s'ils doivent choisir entre *Décrocher un appel entrant*, *Passer un appel interne* ou *Passer un appel sortant*. Concernant l'ARM 6 choix sont possibles :

- AS>AE>AI : Les appels sortants sont prioritaires sur les appels entrants, qui sont eux-mêmes prioritaires sur les appels internes. Cela implique qu'un ARM peut passer un appel interne uniquement s'il/elle n'a ni appels entrants à décrocher ni appels sortants à passer.
- AS>AI>AE : Les appels sortants sont prioritaires sur les appels internes, eux mêmes prioritaires sur les appels entrants. Cela implique qu'un ARM ne décroche un nouvel appel que s'il/elle a passé tous les appels sortants et appels internes en attente.
- AE>AI>AS : Les appels entrants sont prioritaires sur les appels internes, eux-mêmes prioritaires sur les appels sortants. Cela implique que le décroché des appels est prioritaire par rapport aux communications internes et externes. De plus si on ARM doit contacter un MR et passer un appel sortant, il/elle préférera d'abord contacter le MR.
- AE>AS>AI : Les appels entrants sont prioritaires sur les appels sortants, eux-mêmes prioritaires sur les appels internes. Cela implique que le décroché des appels est prioritaire par rapport aux communications internes et externes. De plus si on ARM doit contacter un MR et passer un appel sortant, il/elle préférera d'abord passer l'appel sortant.
- AI>AE>AS : Les appels internes sont prioritaires sur les appels entrants, eux mêmes prioritaires sur les appels sortants. Dans le cas où un ARM doit passer un appel interne et décrocher un nouvel appel, il/elle passera d'abord son appel interne avant de décrocher l'appel entrant et en dernier lieu il passera un appel sortant.
- AI>AS>AE : Les appels internes sont prioritaires sur les appels sortants eux-mêmes prioritaires sur les appels entrants.

5.2.2 Des propositions pour le traitement des appels : Le Triage et La Prépriorisation

Dans cette section nous proposons plusieurs modifications sur la manière de décrocher les appels. Nous présentons chacune des configurations, puis analysons les résultats

Scénario 0 : AS-IS

Samu Centre 15 référentiel et guide d'évaluation (2015)

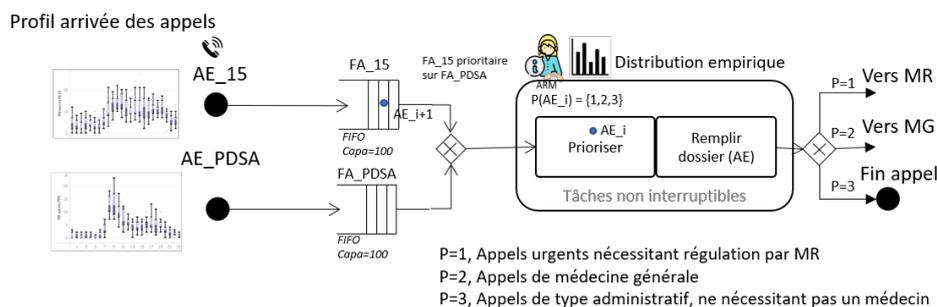


Figure 5.13 – Scénario AS-IS

donnés par chaque scénario. Le plan d'expériences suivi est présenté dans la figure 5.15. Deux nouvelles manières de tester les appels sont proposées. Le Triage et la Prépriorisation. Ces scénarios peuvent être combinés. Cela nous donne donc trois familles de scénarios différents. Pour les scénarios de Triage ainsi que pour le scénario combiné, nous proposons deux versions du scénario en faisant varier le paramètre de durée de la priorisation d'un appel. Cela nous fait donc un total de 6 scénarios à analyser. Afin de faciliter la lecture des résultats, nous proposons différents codes afin d'identifier les résultats :

- AS-IS : code pour un scénario qui correspond à la configuration initiale.
- PP : Scénario de Prépriorisation.
- T-1/3 : Scénario de Triage avec une durée de triage équivalente à un tiers du temps total de gestion de l'appel.
- T-2/3 : Scénario de Triage avec une durée de triage équivalente à deux tiers du temps total de gestion de l'appel.
- PP-T-1/3 : Scénario mixte de Prépriorisation et de Triage avec une durée de triage équivalente à un tiers du temps total de gestion de l'appel.
- PP-T-2/3 : Scénario mixte de Prépriorisation et de Triage avec une durée de triage équivalente à deux tiers du temps total de gestion de l'appel.

La Prépriorisation

En développant un scénario de prépriorisation, nous faisons l'hypothèse qu'à l'avenir il serait plausible d'avoir accès à de l'information sur la gravité de l'appel en amont. Certains pans de la recherche semblent très prometteurs à ce sujet : Que ce soit en utilisant un *Automatic Call Dispatcher* (ACD) pour permettre au patient de faire une première auto-évaluation de sa situation, ou encore du côté de la recherche sur l'analyse vocale des émotions à l'aide de l'intelligence artificielle (Blomberg et al., 2019). Blomberg et al, proposent par exemple d'utiliser un modèle de *Machine Learning* pour analyser la voix des patients afin d'identifier plus facilement les arrêts cardiaques. Leurs résultats montrent que l'algorithme de *Machine Learning* était en mesure de détecter les arrêts cardiaques plus rapidement que les opérateurs téléphoniques même s'il est encore nécessaire de l'améliorer pour augmenter sa précision. Nous pouvons espérer que le développement de ces techniques et de ces

technologies permettra aux centres d'appels de collecter plus de données sur l'appel au plus tôt dans le processus de gestion de l'urgence. Parmi ces nouvelles technologies on peut citer les objets connectés de santé qui peuvent aussi permettre de transmettre des informations permettant d'évaluer la gravité de l'état du patient (rythme cardiaque, hypoxie, pression artérielle, température, ECG etc.) sans oublier l'intégration de la télémédecine qui rajoutera la dimension visuelle au diagnostic.

Dans ce scénario, nous faisons donc l'hypothèse que l'information sur la gravité de l'appel est disponible lorsqu'il entre en file d'attente. De ce fait, l'ARM décrochera d'abord les appels selon leur priorité (cf. figure 5.14). Nous avons gardé une activité de priorisation de l'ARM suite au décroché de l'appel afin de vérifier et de valider la priorisation effectuée en amont.

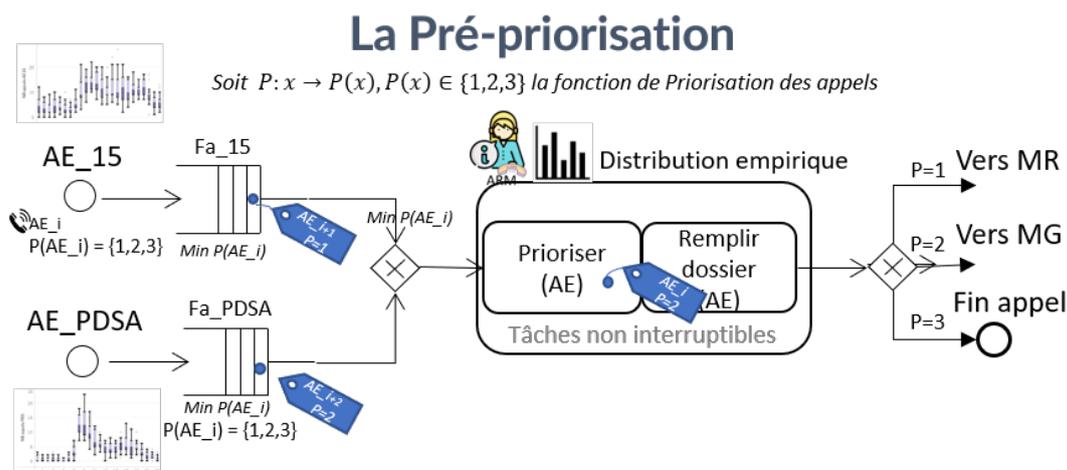


Figure 5.14 – Scénario de prépriorisation des appels

Le Triage

Une pratique qui va prochainement voir le jour suite au pacte de refondation des urgences avec la mise en place du SAS³ est le triage dans les services d'urgences physiques. Cela permet d'évaluer rapidement la gravité des patients pour les orientés du mieux possible et réduire l'engorgement (Storm-Versloot et al., 2011). Dans la continuité des travaux de Lamine et al. (Lamine et al., 2015). Nous proposons d'appliquer le même principe aux centres d'appels d'urgence en mettant la priorité sur le décroché des appels et l'évaluation de leur gravité avant de continuer dans le traitement de l'appel. Pour ce faire, nous scindons la macro activité 1 "Décrocher et prioriser" en deux étapes correspondant à *Prioriser* et *Remplir le dossier* (cf. figure 5.16). La première activité consiste à établir la gravité de l'appel, la deuxième concerne le reste du traitement, avec la création du dossier de régulation médicale et sa complétion. Nous ne disposons pas d'information précise sur la durée de l'activité de priorisation. Ainsi nous proposons de tester deux scénarios, un premier plutôt optimiste où on considère que le temps de priorisation d'un appel correspond à un tiers du temps total de la communication avec l'ARM. Et un deuxième scénario où ce temps de priorisation correspondant à deux tiers du temps total de la communication. À l'issue de cette première étape, l'ARM a plusieurs choix :

- Soit il n'y a pas d'appels dans la file d'attente d'arrivée des appels, alors on peut enchaîner sur la complétion du dossier de l'appel en cours.
- S'il y a un appel dans la file d'attente alors on regarde la priorité de l'appel en cours.
- Si l'appel en cours est de priorité 1 et est donc urgent, alors on continue de traiter cet appel-là.
- Si l'appel en cours est d'une priorité 2 ou 3 alors on le place dans une file d'attente intermédiaire et on priorise le nouvel appel.
- Si le nouvel appel est d'une priorité supérieure à l'appel mis en attente, on traite d'abord cet appel avant de finir par traiter l'appel mis en attente.

3. <https://www.gouvernement.fr/pacte-de-refondation-des-urgences-754-millions-et-des-mesures-perennes-pour-soutenir-les>

- Si le nouvel appel est d'une priorité inférieure ou égale à l'appel mis en attente, alors on met ce nouvel appel en attente et on reprend le traitement de l'appel précédent.

5.2.3 1ère Expérimentation : Quels sont les effets de la prépriorisation et du triage ?

Ainsi, nous avons fait une première expérimentation afin d'évaluer l'effet de la prépriorisation, du triage, des deux combinés par rapport au cas initial. Nous appellerons ce jeu de résultat : la configuration C1. La simulation est faite sur un horizon d'une journée, sur une base de 100 réplifications.

Effet sur la qualité de service

Pour chacun des scénarios nous avons calculé le QoS_{20} et le QoS_{60} moyenné sur les 100 réplifications sur chaque créneau de 30 minutes durant la journée. Nous avons détaillé sur chaque créneau de 30 minutes durant la journée, car les gains globaux à l'échelle de la journée pour le QoS est assez faible. Ces résultats sont présentés dans le tableau 5.3. Le scénario de triage permet au mieux un gain de 3% pour le T-1/3 et au pire un gain de 2,7% pour le T-2/3 sur le QoS_{60} . Pour le QoS_{20} les scénarios de triage permettent des gains respectivement de 3,7% et 3,6% pour les scénarios T-1/3 et T-2/3. Le scénario de prépriorisation permet un gain de 0,3% et de 0,5% respectivement pour QoS_{60} et pour QoS_{20} . La combinaison des scénarios de triage et de prépriorisation permettent respectivement des gains sur le QoS_{60} de 3,0% et de 2,7% pour le PP_T-1/3 et le PP_T-2/3. Pour le QoS_{20} c'est là où on note la meilleure amélioration avec des gains de 3,7% et 3,6% pour le PP_T-1/3 et le PP_T-2/3.

Bien que raisonnables, ces modifications présentent déjà des améliorations par rapport au scénario initial. Néanmoins les résultats présentés sont pour tous les appels confondus. Il est intéressant de regarder les gains obtenus par type d'appels. Ces résultats sont présentés dans le tableau 5.4. On constate que les gains obtenus sur chaque type d'appels peuvent varier entre -1,1% jusqu'à 7,1% en fonction du type d'appel et du scénario.

Or, si on calcule le QoS_{20} et le QoS_{60} sur chaque créneau de 30 minutes, on fait apparaître des gains nettement plus importants sur certains créneaux, notamment

Tableau 5.3 – Valeur du QoS au global sur la journée dans la première configuration.

	Moy. QoS_{60}	δ AS-IS	Moy. QoS_{20}	δ AS-IS
AS-IS	95,4%		86,8%	
T-1/3	98,5%	3,0%	90,2%	3,4%
T-2/3	98,2%	2,7%	90,4%	3,2%
PP	95,7%	0,3%	87,3%	0,5%
PP_T-1/3	98,4%	3,0%	90,2%	3,7%
PP_T-2/3	98,2%	2,7%	90,0%	3,6%

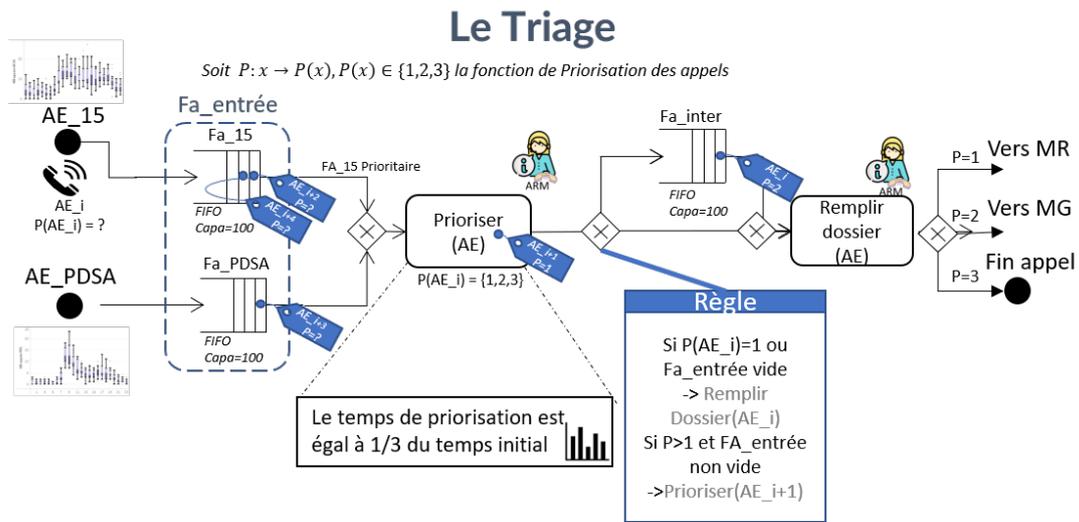


Figure 5.16 – Scénario de triage des appels

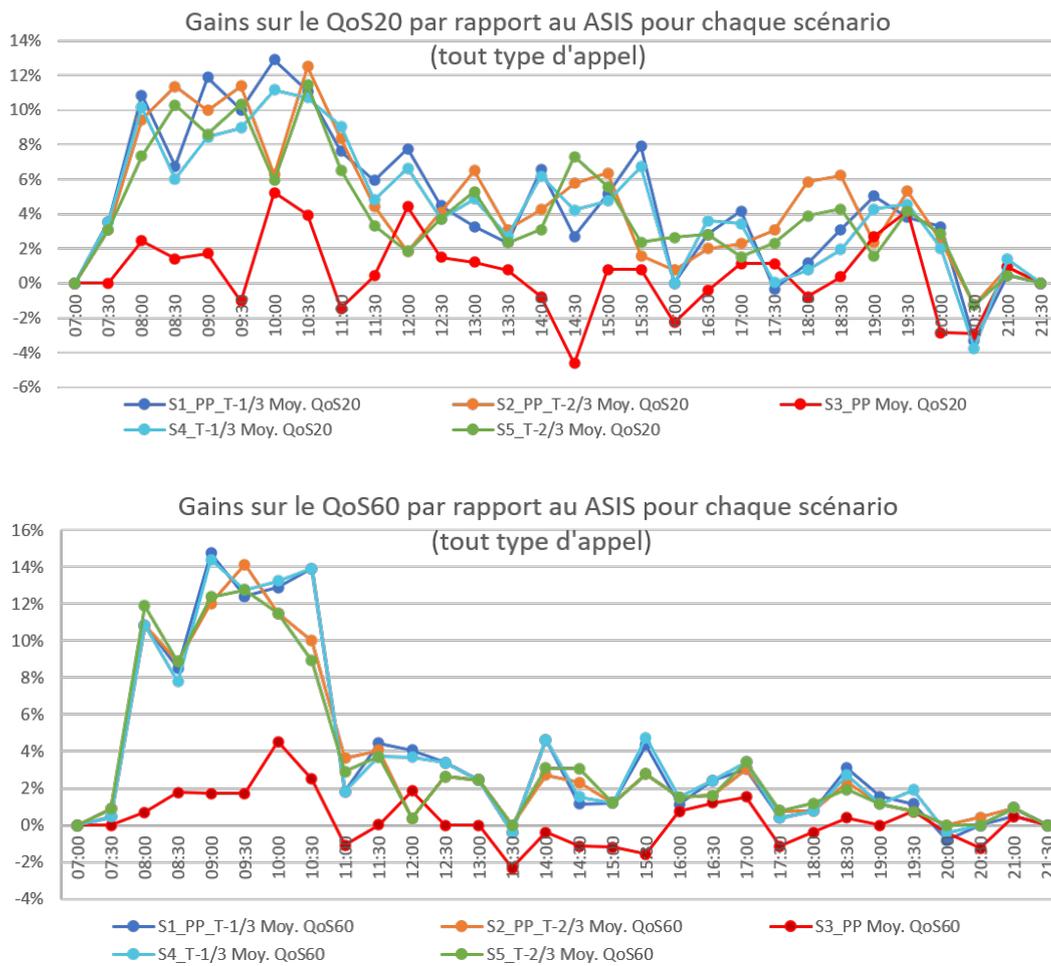


Figure 5.17 – Gains sur le QoS20 et QoS60 au global pour tous type d'appels, dans la configuration C1

Tableau 5.4 – Amélioration du QoS au global par priorité d'appels

		Moy. QoS20	δ AS-IS	Moy. QoS60	δ AS-IS
S1_PP_T-1/3	P1	93,6%	5,2%	99,3%	3,0%
	P2	94,5%	3,3%	99,3%	2,0%
	P3	76,1%	5,7%	95,6%	6,4%
S2_PP_T-2/3	P1	93,1%	4,8%	99,1%	2,7%
	P2	94,2%	3,0%	99,2%	1,9%
	P3	76,8%	6,3%	95,0%	5,8%
S3_PP	P1	90,3%	2,0%	97,6%	1,3%
	P2	92,5%	1,3%	98,3%	1,0%
	P3	69,4%	-1,1%	88,0%	-1,1%
S4_T-1/3	P1	92,8%	4,5%	99,0%	2,6%
	P2	94,3%	3,1%	99,1%	1,8%
	P3	76,1%	5,6%	96,2%	7,1%
S5_T-2/3	P1	92,2%	3,9%	98,6%	2,2%
	P2	93,5%	2,3%	99,0%	1,7%
	P3	76,7%	6,2%	95,5%	6,3%
S6_AS-IS	P1	88,4%	-	96,6%	-
	P2	91,2%	-	97,3%	-
	P3	70,5%	-	89,2%	-

Tableau 5.5 – Gains par rapport au AS-IS pour les appels P1 sur chaque créneau

Appels P1	S1_PP_T-1/3		S2_PP_T-2/3		S3_PP		S4_T-1/3		S5_T-2/3	
	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60
07 :30	3,0%	0,0%	3,0%	0,0%	0,0%	0,0%	3,0%	0,0%	3,0%	0,0%
08 :00	9,6%	4,3%	10,6%	5,3%	4,3%	2,1%	7,4%	4,3%	5,3%	3,2%
08 :30	8,9%	6,7%	14,4%	7,8%	4,4%	3,3%	6,7%	4,4%	13,3%	7,8%
09 :00	17,9%	10,5%	11,6%	8,4%	3,2%	4,2%	14,7%	7,4%	10,5%	6,3%
09 :30	9,3%	11,3%	12,4%	12,4%	2,1%	4,1%	7,2%	11,3%	8,2%	10,3%
10 :00	16,8%	11,6%	5,3%	10,5%	8,4%	9,5%	13,7%	10,5%	7,4%	10,5%
10 :30	10,5%	10,5%	12,6%	6,3%	6,3%	4,2%	12,6%	10,5%	10,5%	4,2%
11 :00	10,8%	0,0%	9,7%	1,1%	1,1%	-1,1%	10,8%	0,0%	6,5%	-1,1%
11 :30	3,4%	1,1%	2,2%	2,2%	0,1%	1,1%	-3,4%	1,1%	3,4%	1,1%
12 :00	4,4%	2,2%	2,2%	1,1%	4,4%	1,1%	3,3%	2,2%	1,1%	1,1%
12 :30	0,0%	3,4%	0,0%	2,3%	0,0%	3,4%	0,0%	2,3%	0,0%	2,3%
13 :00	7,7%	2,2%	6,6%	2,2%	2,2%	0,0%	7,7%	2,2%	6,6%	2,2%
13 :30	2,2%	-1,1%	5,6%	0,0%	5,6%	-1,1%	3,4%	-1,1%	5,6%	0,0%
14 :00	14,4%	4,4%	8,9%	2,2%	2,2%	1,1%	12,2%	4,4%	10,0%	2,2%
14 :30	8,4%	2,1%	6,3%	4,2%	-6,3%	-2,1%	7,4%	1,1%	8,4%	4,2%
15 :00	1,1%	1,1%	2,3%	1,1%	2,3%	0,0%	3,4%	1,1%	1,1%	1,1%
15 :30	6,3%	4,2%	0,0%	4,2%	2,1%	2,1%	7,4%	4,2%	-1,1%	3,2%
16 :00	0,0%	2,2%	2,2%	2,2%	-2,2%	1,1%	-1,1%	2,2%	2,2%	2,2%
16 :30	3,9%	2,6%	3,9%	2,6%	0,0%	1,3%	3,9%	2,6%	3,9%	2,6%
17 :00	4,5%	3,4%	2,2%	3,4%	2,2%	2,2%	3,4%	3,4%	0,0%	3,4%
17 :30	0,0%	0,0%	4,5%	-1,1%	5,6%	-1,1%	1,1%	-1,1%	3,4%	-1,1%
18 :00	1,1%	1,1%	1,1%	1,1%	2,2%	1,1%	2,2%	1,1%	1,1%	1,1%
18 :30	3,4%	1,1%	2,2%	0,0%	0,0%	-1,1%	0,0%	0,0%	0,0%	-1,1%
19 :00	3,3%	0,0%	2,2%	0,0%	3,3%	0,0%	1,1%	0,0%	1,1%	0,0%
19 :30	4,4%	1,1%	5,6%	0,0%	5,6%	1,1%	4,4%	1,1%	3,3%	0,0%
20 :00	-1,1%	-1,1%	0,0%	-2,2%	1,1%	0,0%	-1,1%	-1,1%	-2,2%	-2,2%
20 :30	-3,3%	1,1%	1,1%	1,1%	-3,3%	0,0%	-2,2%	1,1%	0,0%	0,0%
21 :00	1,3%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
21 :30	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Tableau 5.6 – Gains par rapport au AS-IS pour les appels P2 sur chaque créneau

S1_PP_T-1/3		S2_PP_T-2/3		S3_PP		S4_T-1/3		S5_T-2/3	
Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60
0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,02	0,00
0,11	0,05	0,09	0,05	0,03	0,01	0,10	0,02	0,05	0,07
0,04	0,03	0,03	0,03	0,03	0,02	0,03	0,01	0,02	0,02
0,09	0,08	0,08	0,06	0,02	0,07	0,04	0,07	0,04	0,06
0,06	0,09	0,12	0,11	0,01	0,04	0,06	0,08	0,10	0,10
0,10	0,04	0,00	0,04	0,04	0,03	0,08	0,05	0,00	0,03
0,09	0,08	0,11	0,07	0,07	0,04	0,09	0,06	0,04	0,05
0,05	-0,01	0,04	0,01	0,02	0,01	0,05	0,00	0,01	0,01
0,05	0,02	0,04	0,02	0,01	0,01	0,07	0,02	0,04	0,02
0,01	0,02	-0,02	0,01	0,01	0,02	0,01	0,02	0,00	0,00
0,08	0,03	0,02	0,01	0,03	0,00	0,06	0,03	0,02	0,01
-0,02	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00
0,03	0,00	0,04	0,00	0,03	-0,01	0,03	0,00	0,04	0,00
0,00	0,01	0,00	0,01	0,00	0,00	-0,01	0,01	0,00	0,01
0,03	0,00	0,05	0,00	0,00	-0,02	0,05	0,00	0,05	0,00
0,06	0,02	0,05	0,02	0,00	0,02	0,06	0,02	0,05	0,02
0,10	0,03	0,09	0,02	0,05	0,02	0,09	0,03	0,09	0,03
0,01	0,01	0,00	0,01	0,01	0,01	0,03	0,01	0,01	0,01
0,00	0,00	-0,03	0,00	-0,03	-0,01	0,01	0,00	-0,01	0,00
0,03	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00
-0,03	0,01	0,01	0,00	-0,03	0,01	-0,01	0,01	0,00	0,00
0,02	0,00	0,03	0,00	-0,02	0,00	0,02	0,00	0,00	0,00
0,00	0,01	0,03	0,01	0,00	0,01	-0,01	0,01	0,04	0,01
0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,01	0,00	0,01
0,05	0,03	0,07	0,01	0,05	0,01	0,05	0,03	0,05	0,01
-0,04	0,00	0,00	0,00	-0,04	-0,02	-0,04	0,00	0,00	0,00
0,02	0,00	0,00	0,00	0,02	0,00	0,02	0,00	0,00	0,00
-0,03	-0,03	0,00	0,00	-0,03	0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tableau 5.7 – Gains par rapport au AS-IS pour les appels P3 sur chaque créneau

S1_PP_T-1/3		S2_PP_T-2/3		S3_PP		S4_T-1/3		S5_T-2/3	
Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60
0,06	0,01	0,04	0,02	0,00	0,00	0,06	0,01	0,04	0,02
0,12	0,22	0,09	0,21	0,00	-0,01	0,13	0,25	0,11	0,25
0,07	0,15	0,16	0,15	-0,03	0,00	0,08	0,17	0,15	0,16
0,09	0,25	0,10	0,21	0,00	-0,06	0,07	0,28	0,11	0,24
0,14	0,17	0,10	0,19	-0,06	-0,03	0,13	0,19	0,13	0,18
0,12	0,22	0,13	0,19	0,03	0,01	0,12	0,23	0,10	0,20
0,13	0,22	0,14	0,16	-0,01	0,00	0,10	0,24	0,19	0,17
0,07	0,06	0,11	0,08	-0,07	-0,03	0,11	0,05	0,11	0,08
0,09	0,09	0,07	0,07	0,00	-0,02	0,10	0,07	0,03	0,07
0,16	0,07	0,05	-0,01	0,07	0,02	0,14	0,06	0,04	0,00
0,06	0,04	0,09	0,04	0,02	-0,03	0,05	0,05	0,08	0,04
0,02	0,04	0,10	0,04	0,00	0,00	0,05	0,04	0,07	0,04
0,02	0,00	0,00	0,00	-0,05	-0,04	0,02	0,00	-0,02	0,00
0,04	0,07	0,03	0,04	-0,04	-0,02	0,06	0,07	-0,01	0,05
-0,03	0,01	0,06	0,02	-0,06	0,00	0,01	0,03	0,08	0,04
0,08	0,01	0,11	0,01	0,00	-0,04	0,05	0,01	0,10	0,01
0,08	0,05	-0,01	0,02	-0,03	-0,07	0,05	0,06	0,02	0,02
-0,01	0,00	0,00	0,01	-0,05	0,00	-0,01	0,01	0,04	0,01
0,04	0,04	0,04	0,02	0,01	0,03	0,05	0,04	0,05	0,02
0,05	0,05	0,03	0,05	0,00	0,02	0,05	0,06	0,03	0,06
0,01	0,00	0,03	0,03	0,00	-0,03	0,00	0,01	0,03	0,03
0,01	0,01	0,12	0,01	-0,03	-0,02	-0,01	0,01	0,09	0,02
0,05	0,06	0,12	0,05	0,01	0,01	0,06	0,06	0,08	0,05
0,09	0,03	0,03	0,02	0,03	0,00	0,09	0,02	0,03	0,02
0,02	0,00	0,04	0,01	0,02	0,00	0,04	0,02	0,04	0,01
0,11	-0,01	0,06	0,02	-0,06	0,00	0,08	0,00	0,09	0,02
-0,06	-0,01	-0,04	0,00	-0,05	-0,03	-0,08	-0,01	-0,03	0,00
0,01	0,02	0,02	0,02	0,03	0,01	0,03	0,02	0,01	0,02
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tableau 5.8 – Qualité de service et gains par rapport à l’AS-IS pour chaque scénario

	AS-IS		Triage				Prériorisation				Triage + Prériorisation			
	QoS60	QoS20	QoS 20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS IS	QoS60	δ AS IS
07 :00	1,00	0,99	1,00	0,5%	1,00	0,0%	0,99	0,0%	1,00	0,0%	1,00	0,5%	1,00	0,0%
07 :30	0,99	0,95	0,99	3,6%	1,00	0,4%	0,95	0,0%	0,99	0,0%	0,99	3,6%	1,00	0,4%
08 :00	0,84	0,63	0,73	9,8%	0,95	10,8%	0,66	3,1%	0,86	1,7%	0,74	10,5%	0,95	10,8%
08 :30	0,85	0,64	0,72	7,8%	0,95	9,6%	0,67	2,5%	0,88	2,5%	0,73	8,5%	0,95	10,3%
09 :00	0,76	0,50	0,57	7,4%	0,91	15,1%	0,51	1,0%	0,79	3,1%	0,61	10,9%	0,91	15,4%
09 :30	0,82	0,59	0,70	10,7%	0,94	12,4%	0,61	2,4%	0,83	1,0%	0,71	11,7%	0,94	12,1%
10 :00	0,82	0,56	0,68	11,8%	0,93	11,5%	0,58	2,1%	0,85	3,8%	0,69	13,6%	0,93	11,1%
10 :30	0,81	0,58	0,65	6,4%	0,95	13,2%	0,60	1,8%	0,85	3,2%	0,65	6,8%	0,95	13,2%
11 :00	0,91	0,67	0,79	11,6%	0,96	4,7%	0,68	0,4%	0,89	-1,4%	0,78	10,1%	0,96	4,7%
11 :30	0,94	0,83	0,91	8,5%	0,99	4,8%	0,86	2,9%	0,95	1,1%	0,92	9,6%	1,00	5,5%
12 :00	0,96	0,82	0,88	6,3%	0,99	3,3%	0,85	3,0%	0,97	1,8%	0,89	7,4%	0,99	3,7%
12 :30	0,97	0,89	0,92	2,3%	0,99	1,9%	0,89	0,0%	0,97	0,4%	0,92	3,0%	0,99	1,9%
13 :00	0,99	0,93	0,96	3,3%	1,00	1,2%	0,93	0,0%	0,98	-0,4%	0,94	1,6%	1,00	1,2%
13 :30	0,99	0,90	0,95	4,7%	0,99	0,4%	0,91	0,8%	0,97	-1,9%	0,94	4,3%	0,99	0,4%
14 :00	0,93	0,81	0,89	8,1%	1,00	6,6%	0,81	-0,4%	0,94	0,8%	0,90	8,5%	1,00	6,6%
14 :30	0,96	0,82	0,85	3,1%	0,97	0,8%	0,80	-2,3%	0,95	-1,2%	0,84	1,5%	0,97	0,4%
15 :00	0,98	0,89	0,93	4,0%	1,00	1,2%	0,90	0,8%	0,97	-1,6%	0,94	4,4%	1,00	1,2%
15 :30	0,96	0,79	0,89	9,5%	1,00	3,6%	0,82	3,2%	0,95	-1,6%	0,90	10,7%	1,00	3,2%
16 :00	0,97	0,90	0,91	1,1%	0,99	1,9%	0,89	-1,5%	0,97	-0,4%	0,91	1,1%	0,99	1,5%
16 :30	0,98	0,89	0,94	5,2%	1,00	2,0%	0,89	0,0%	0,98	-0,4%	0,93	4,4%	1,00	2,0%
17 :00	0,98	0,89	0,95	6,1%	1,00	2,3%	0,92	2,7%	0,98	0,0%	0,96	6,8%	1,00	1,9%
17 :30	0,98	0,89	0,88	-1,1%	0,98	0,4%	0,89	0,8%	0,98	0,0%	0,87	-1,5%	0,98	0,4%
18 :00	0,99	0,91	0,93	2,3%	1,00	0,8%	0,91	0,8%	0,97	-1,6%	0,93	2,7%	1,00	0,8%
18 :30	0,97	0,87	0,91	4,3%	0,99	1,9%	0,90	3,5%	0,98	0,4%	0,92	5,4%	1,00	2,3%
19 :00	0,99	0,95	0,96	0,4%	0,99	0,4%	0,95	-0,8%	0,98	-0,4%	0,97	1,2%	1,00	0,8%
19 :30	0,97	0,86	0,94	7,2%	0,99	2,7%	0,87	0,8%	0,97	0,4%	0,93	6,4%	0,98	1,9%
20 :00	0,97	0,92	0,92	0,4%	0,98	0,4%	0,93	1,6%	0,98	1,2%	0,93	1,6%	0,97	0,0%
20 :30	0,99	0,96	0,94	-2,1%	1,00	0,4%	0,96	0,0%	0,99	0,0%	0,94	-1,7%	1,00	0,4%
21 :00	0,99	0,98	0,99	0,9%	1,00	1,4%	0,97	-0,5%	1,00	0,9%	0,98	0,0%	1,00	0,9%
21 :30	1,00	1,00	1,00	0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%

ceux qui présentent un nombre d’appels élevés. Un premier constat est qu’entre 00h00 et 07h00 ainsi qu’entre 21h30 et 23h30, le QoS_{20} est de 100%, tous les appels sont décrochés en moins de 20 secondes. C’est pourquoi dans le tableau 5.8 nous ne présentons que les créneaux entre 07h00 et 21h30. La mise en place du triage permet d’obtenir un gain pouvant aller jusqu’à 15% par rapport au scénario AS-IS. La figure 5.17 présente les gains en pourcentage par rapport au scénario de base pour chacun des scénarios. Par exemple, sur le créneau de 09h00, le QoS_{60} passe de 0,76 à 0,91. Ce sont sur les créneaux où il y a le plus d’appels, soit entre 7h30 et 11h00 où on observe les améliorations les plus sensibles, entre 8% et 15% pour les scénarios de triage et les scénarios mixtes. Ce qu’il est intéressant de noter c’est que le scénario de triage permet de passer au-dessus des 90% d’appels décrochés en moins de 60 secondes entre 7h30 et 11h00.

Une autre chose que l’on peut voir est que sur certains créneaux, la prériorisation ou le triage a un impact légèrement négatif (au plus -5%) sur les performances. De plus, on constate que le scénario PP est le moins performant de tous. Cette perte de performance se doit d’être expliquée. Pour ce faire, il est nécessaire d’avoir plus de précisions. Si on détaille les gains par niveau de gravité des appels (voir figures 5.25, 5.27 et 5.20) on constate que les gains sont bien meilleurs pour les appels de priorité P1 et que la perte de performance a lieu sur les appels de priorité P3.

Pour les appels P1, le scénario de triage seul améliore jusque 18% du QoS_{60} sur certains créneaux, comme celui de 9h00, entre un temps de triage à 1/3 ou 2/3 on note une différence de 3% entre les deux scénarios. Donc, même avec un temps de triage très important on note des gains non négligeables en matière de décroché

des appels. Le scénario de PP permet des gains pouvant aller jusque 8% (comme à 10h00). Néanmoins on note une baisse en moyenne pour les appels P1 notamment sur le créneau de 14h30. Cette diminution de production est due au fait que lorsqu'un appel est décroché, les ARM ne vont pas interrompre l'appel en cours même s'il y a de nouveaux appels à décrocher. De ce fait, si les 4 ARM décrochent des appels qui sont des P3 et qu'un appel P1 arrive en file d'attente, ce dernier va attendre qu'un des ARM ait fini de traiter son appel courant. Ce cas de figure n'est pas possible dans les autres scénarios, c'est pourquoi les résultats sont globalement meilleurs.

Pour les appels P2, les gains vont jusqu'à 12% d'amélioration pour le QoS_{20} , ce sont encore une fois les scénarios proposant du triage seul ou mixé à la prépriorisation qui permettent d'obtenir les meilleurs résultats. Le QoS_{60} est amélioré de 8 à 10 % lors du premier pic d'appels au niveau des scénarios de triage et des scénarios mixtes, le scénario de prépriorisation ne permet une amélioration que de 4%.

On constate enfin une légère baisse de productivité sur les appels P3 : en mettant en place du triage ou de la prépriorisation, on accepte de dégrader les performances de décroché pour les appels de faible gravité afin d'augmenter la performance pour les appels de forte gravité. C'est pourquoi le scénario seul de prépriorisation a un effet globalement négatif sur les appels P3, car ils ne sont pas prioritaires dans ce cas de figure. On constate néanmoins l'intérêt des scénarios de triage et des scénarios mixtes, car ils permettent tout de même d'aller améliorer aussi le décrocher des appels P3. En effet, la particularité du triage est de mettre l'accent sur le décroché des appels pour qu'il soit décroché puis évalué le plus rapidement possible. On améliore ainsi les performances globales du système en termes de décroché des appels.

Effet sur le temps d'attente maximum

Pour chaque scénario, nous avons regardé le temps maximum d'attente sur l'intégralité des répliques afin de voir l'effet des modifications organisationnelles sur les points les plus extrêmes. La figure 5.21 présente les résultats obtenus pour tous les appels confondus. On a un maximum de temps d'attente qui est à 274,8 secondes sur le créneau de 9h00. C'est entre 07h00 et 21h30 où les temps d'attentes maximums sont élevés et globalement supérieurs à 1 minute. Au niveau des scénarios, on remarque que la prépriorisation a des performances globalement équivalentes au AS-IS. Les scénarios de triage permettent de réduire le temps d'attente maximal à 200 secondes. Les scénarios mixtes ont des résultats qui dépendent de la durée du triage, ils sont sensiblement meilleurs si ce dernier est à 1/3, car on observe une réduction du maximum à 200 secondes contre une amélioration bien moins nette et plus proche du cas initial pour le 2/3.

Lorsqu'on étudie les appels de gravité P1, on constate que les scénarios de triage permettent une amélioration globale des performances. On obtient une réduction du temps maximal d'attente pouvant aller jusqu'à 50% par rapport au scénario AS-IS. Sur le créneau de 09h00, on passe d'un maximum de 192 à 102 secondes d'attente au maximum, ce qui est très significatif pour un appel d'urgence. En revanche, plusieurs points ont un maximum qui est supérieur au scénario principal sur les créneaux de 11h00, 13h30, 14h30, 17h30, 18h30 et 20h30. Nous avons donc identifié les répliques

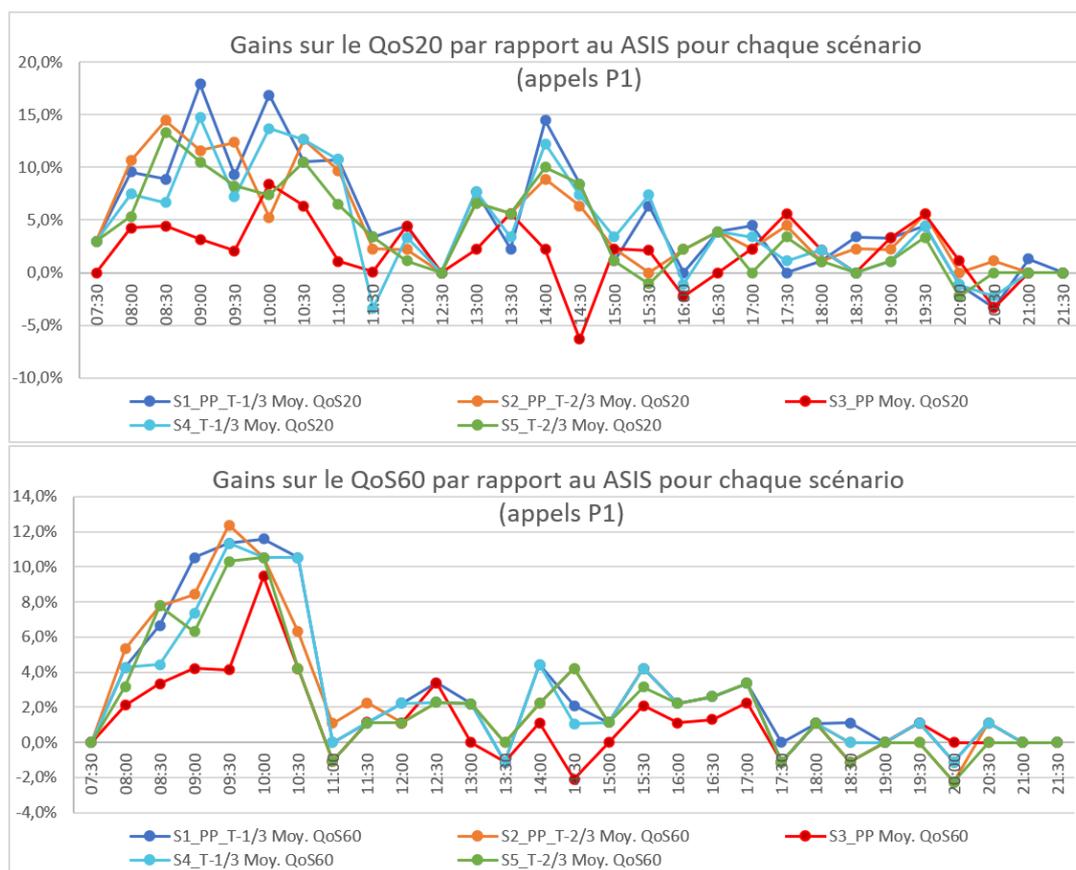


Figure 5.18 – Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C1

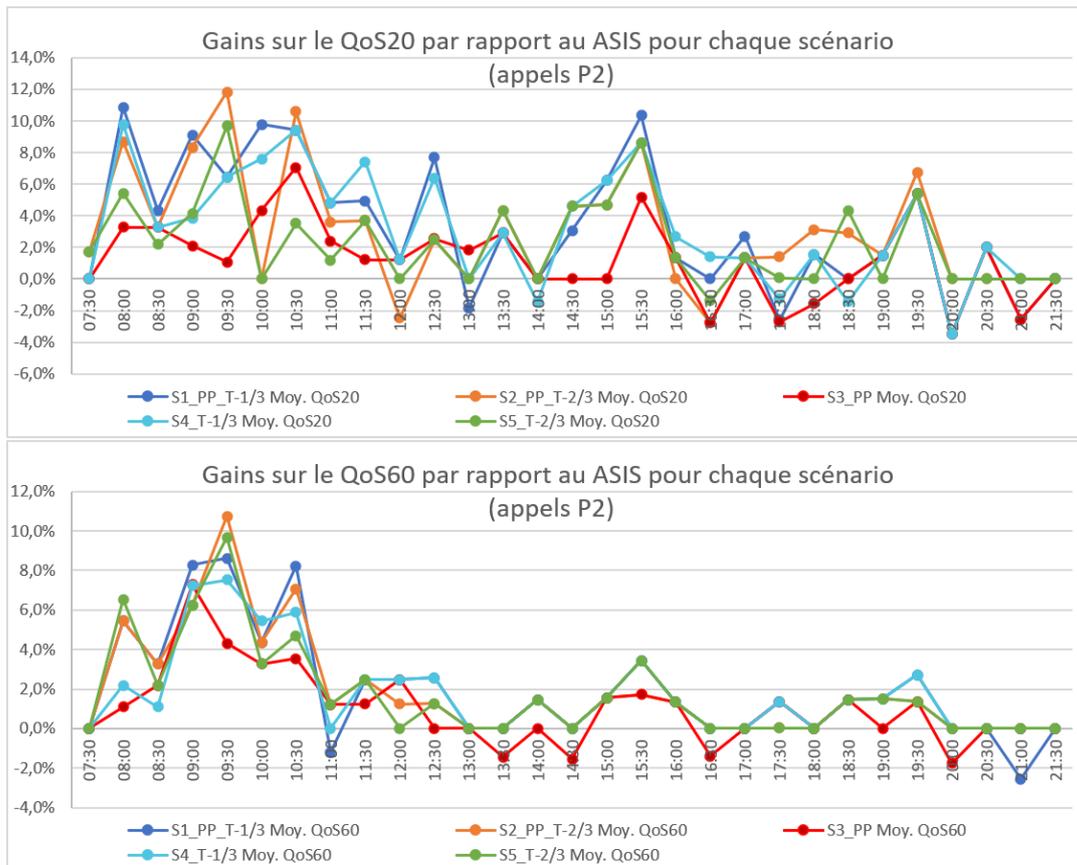


Figure 5.19 – Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C1

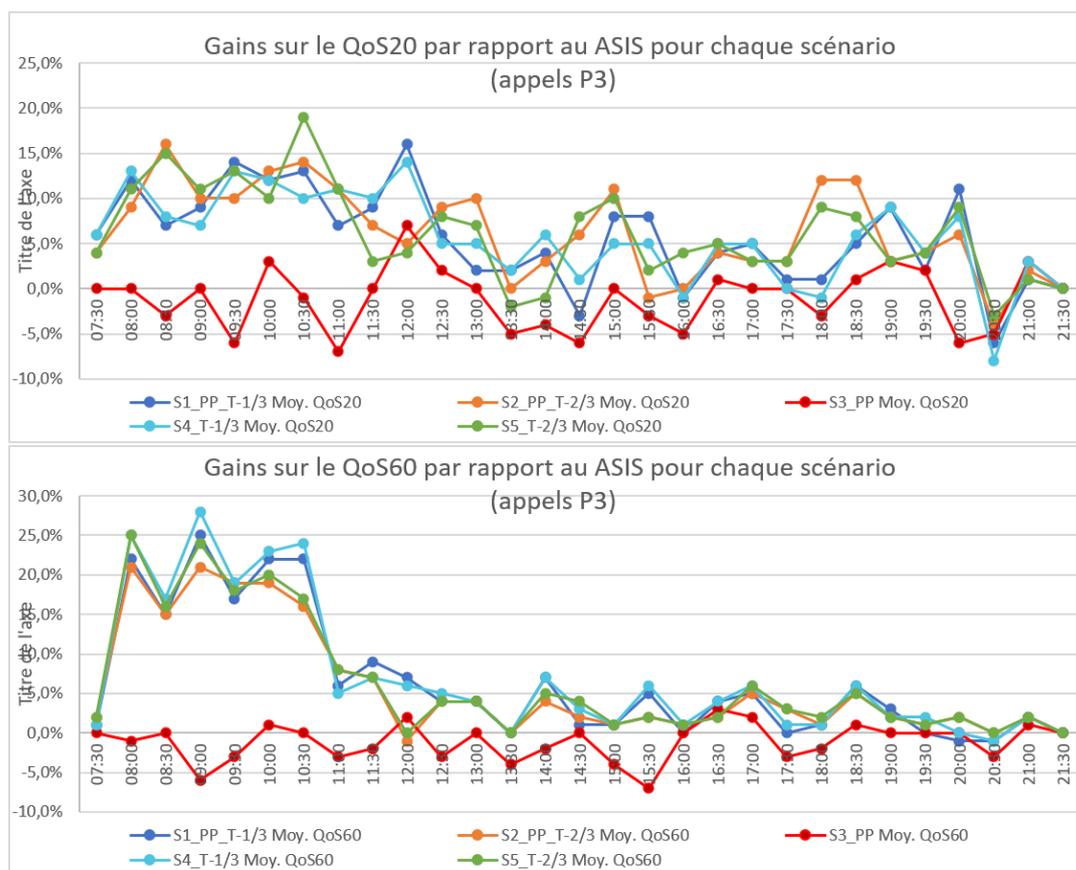


Figure 5.20 – Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C1

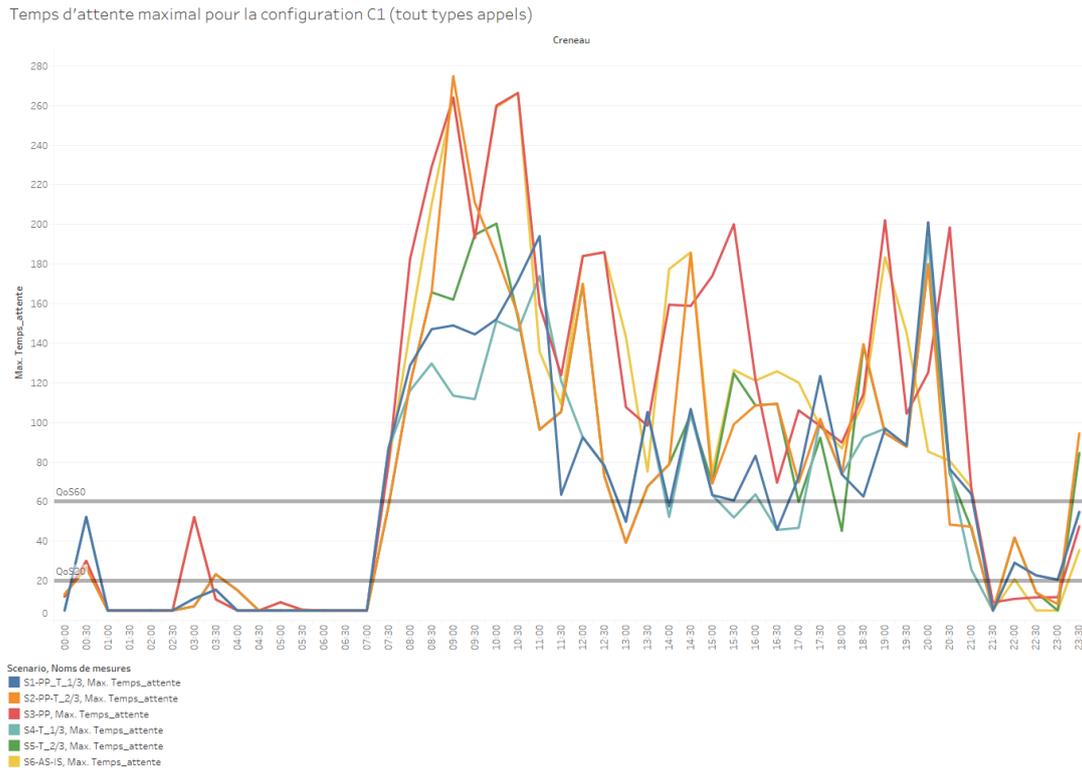


Figure 5.21 – Temps d’attente maximum pour chaque scénario de la configuration C1

et les appels avec un temps d’attente aussi long pour expliquer la cause de cette attente. Cette vérification directement dans le log de données de la simulation nous permet d’aller vérifier le comportement du modèle et de voir si cet écart au niveau des temps d’attente est une erreur inhérente au modèle ou si les temps d’attentes sont longs, car les agents étaient déjà en ligne. A titre d’exemple, nous allons détailler un cas ci-dessous :

Sur le créneau de 18h30 de la réplcation 12, on constate que l’appel 531 attend 92,4 secondes avant d’être décroché. Il arrive dans le centre d’appels à 18h42m48 et il n’est pas décroché avant 18h44m20. Pour comprendre pourquoi il n’est pas décroché, nous regardons ce que font les ARM durant cette période.

- L’ARM 1 est en train de remplir le dossier de l’appel 528 depuis 18h42m28 jusque 18h44m20 puis il priorise l’appel 532 qui est de priorité 1 aussi.
- L’ARM 2 est en train de remplir le dossier de l’appel 530 depuis 18h42m46 jusque 18h46m00.
- L’ARM3 passe un appel sortant externe de 18h40m35 jusque 18h44m18 puis il décroche l’appel 531 à 18h44m20.
- L’ARM 4 remplis le dossier de l’appel 527 de 18h42m38 jusque 18h44m05.

Donc il n’y a pas d’ARM disponible pour décrocher l’appel avant 18h44m20, car ils sont déjà tous soit en train de traiter un appel entrant, soit en train de passer un appel sortant. Pour tous les autres appels, on retrouve le même cas, les appels attendent, car les ressources sont déjà en ligne sur d’autres appels. On constate

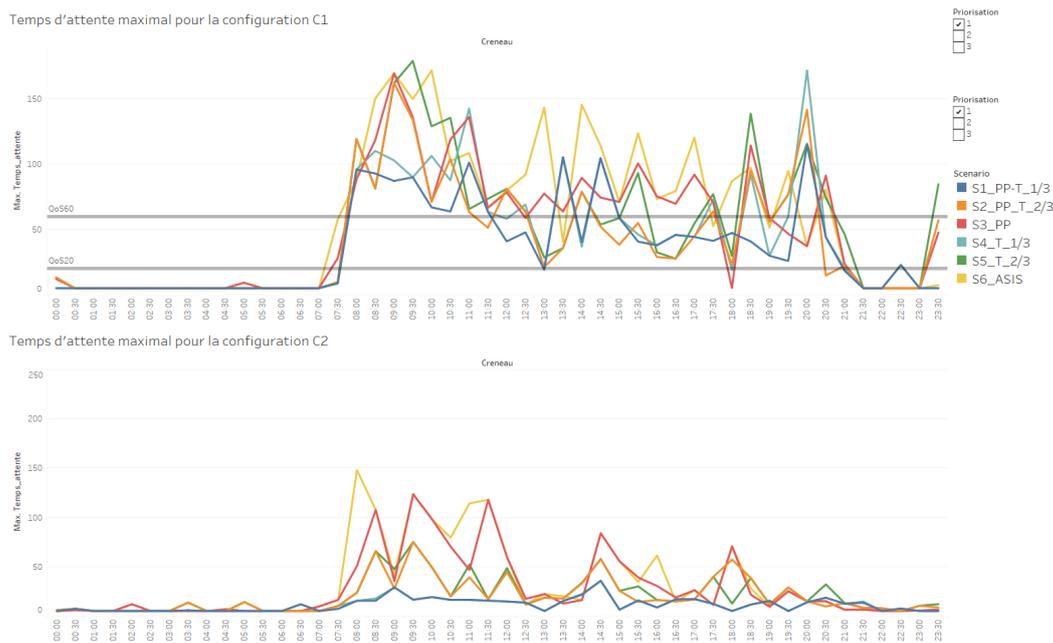


Figure 5.22 – Temps d’attente maximum pour chaque scénario de la configuration C1 : appels P1

que les ressources sont souvent occupées avec des appels sortants. Le traitement prioritaire des appels sortants semble donc avoir un impact fort sur la performance des centres d’appels. En effet, que ce soit pour les appels de 11h, 13h30, 14h30, ou 20h il y a trois ARM sur quatre qui sont en communication sur un appel sortant. Donc les temps d’attente longs ne sont pas dû à une erreur inhérente au modèle mais sont le résultat d’une réalité organisationnelle associée au scénario étudié.

Nous avons effectué le même travail sur le scénario de prépriorisation. Au global, la prépriorisation permet une amélioration sur le temps d’attente maximum de 11,8 secondes en moyenne avec un écart type de 35,6 secondes. Nous avons donc certains cas où les performances sont réduites avec l’utilisation de la prépriorisation. En effet, la pire performance est une augmentation du temps d’attente de 84,4 secondes par rapport au pire cas du scénario initial. Par contre, la meilleure amélioration est une diminution de 103,2 secondes par rapport au cas initial. Dans le tableau 5.9 nous détaillons les activités courantes des ARM dans les cas où le scénario de prépriorisation ne donne pas de meilleurs résultats que le scénario initial. On observe que les ARM sont plus de la moitié du temps en train de passer un appel sortant plutôt que de décrocher un appel entrant. Comme les appels sortants ne sont pas interruptibles même si des appels entrants sont en attentes d’être décrochés, cela engendre des temps d’attentes plus longs. Les temps d’attentes ne sont donc pas dus à la nature de la gestion des files d’attente, mais à la limite de la contrainte des activités non interruptibles.

De ce constat, nous avons envisagé un nouveau scénario d’amélioration dans lequel nous relaxons la priorité sur les appels sortants.

Tableau 5.9 – Activité courante des ARM plutôt que de décrocher l'appel entrant

	ARM 1	ARM 2	ARM 3	ARM 4
09h00	AS	AS	AS	AS
11h00	Décrocher 223	AS	Décrocher 221	AS
13h30	AS	AS	Décrocher 316	Décrocher 315
15h30	AS	AS	Décrocher 423	AS
19h00	AS	Décrocher	AS	AS
20h30	Décrocher 551	AS	AS	AS
22h30	Décrocher	Décrocher	Décrocher	AS

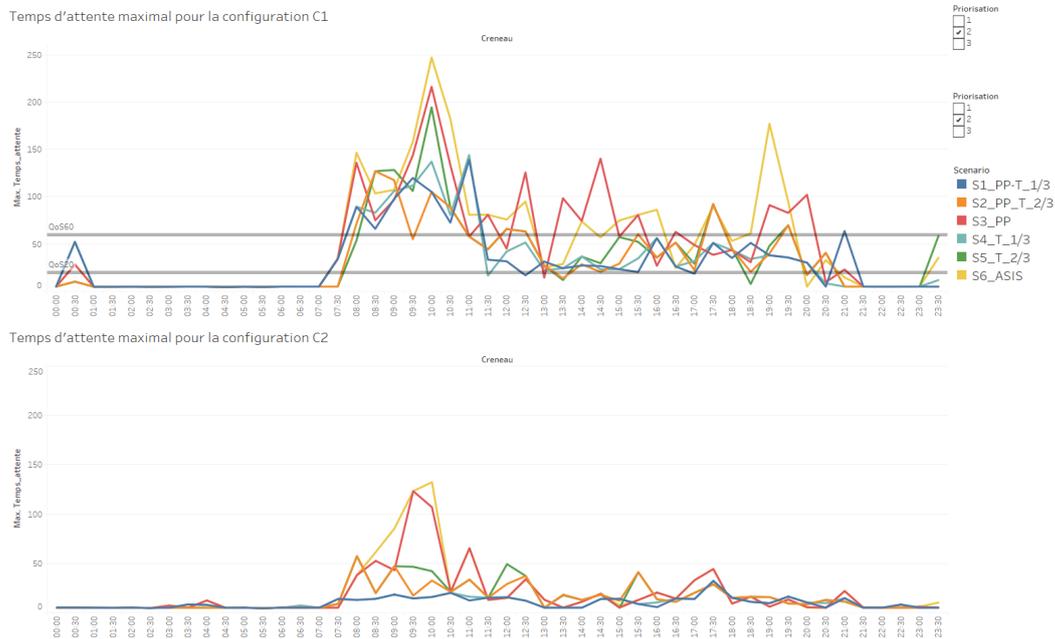


Figure 5.23 – Temps d'attente maximum pour chaque scénario de la configuration C1 : appels P2

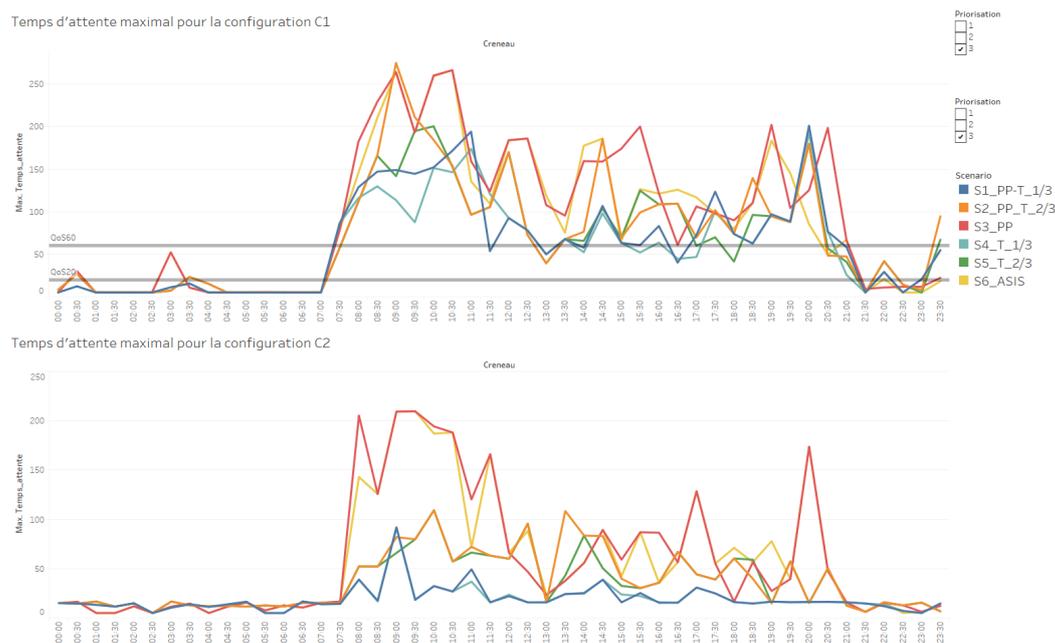


Figure 5.24 – Temps d’attente maximum pour chaque scénario de la configuration C1 : appels P3

La mise en place d’un système de triage ou de prépriorisation permet de gagner du temps de façon significative sur le décroché des appels. Concernant les appels les plus urgents de priorité P1 : Les scénarios mixtes de triage permettent d’améliorer le QoS_{20} de 3 à 5% en moyenne et de 12 à 18% au maximum. Le triage mixte améliore également le QoS_{60} de 3% en moyenne et de 12% au maximum. Les scénarios de prépriorisation seuls permettent une amélioration de 1 à 2% des QoS en moyenne et une amélioration au maximum de 8 à 9 %. Les scénarios de triage seuls permettent une amélioration de 4% en moyenne sur le QoS_{20} et de 13 à 15% au maximum et de 3% en moyenne et de 11% au maximum sur le QoS_{60} . Toutefois on observe des pertes de temps non négligeables dus à la non-disponibilité des ressources. Ainsi, nous souhaitons tester l’hypothèse suivante : *Autoriser l’interruption des appels sortants au profit du décroché des appels entrants peut-il avoir un impact positif sur la performance du décroché ?*

5.2.4 Une piste d’amélioration : modifier la priorisation des appels sortants

Dans cette configuration nous proposons d’autoriser la préemption des ressources pour le décroché des appels. Prenons un exemple pour détailler ce que cela signifie : jusqu’à présent, si un ARM commence un appel sortant et qu’un appel entrant arrive pendant ce temps, l’ARM termine son appel sortant puis va décrocher l’appel entrant. Dans ce nouveau scénario, si un appel entrant arrive alors qu’un ARM

est en train de passer un appel sortant, il peut mettre en pause son appel sortant pour décrocher l'appel entrant afin de s'en occuper. Dans ces conditions nous avons relancé une simulation des scénarios dans cette configuration que nous nommerons : Configuration C2. Les résultats de ces simulations sont comparés au scénario AS-IS de la configuration C1.

Effet sur la qualité de service

Dans la table 5.10 nous comparons les QoS_{20} et les QoS_{60} des différents scénarios à ceux du AS-IS modifié avec la règle des appels sortants pouvant être préemptés. Tous les types d'appels sont concernés et indifférenciés dans un premier temps. Plusieurs constats peuvent être faits :

- Bien qu'on observe toujours de légères variations négatives par rapport à l'AS-IS, celles-ci sont maintenant inférieures à 1% pour les appels de catégorie P3.
- Le scénario de triage améliore le QoS_{20} jusqu'à 21% du AS-IS sur certains créneaux. En comparaison, dans la première configuration le gain maximum avec le scénario de Triage pour le QoS_{20} était de 15%.
- Le scénario de prépriorisation a des gains moindres (au maximum 2,8% pour le QoS_{20} , et 2,1% pour le QoS_{60}).
- Le scénario de triage permet d'atteindre l'objectif de 99% des appels décrochés en moins de 60 secondes sur l'intégralité des créneaux. De plus, il permet d'avoir plus de 90% des appels décrochés en moins de 20 secondes sur l'ensemble des créneaux.
- Les organisations combinées du Triage et de la prépriorisation permettent une amélioration des QoS_{20} et QoS_{60} pouvant aller respectivement de 22,1% et 13,4%. Les gains sont les mêmes que pour le scénario de triage. La combinaison des deux est un peu meilleure sur certains créneaux comme à 12h00 où 98% des appels sont décrochés en moins de 20 secondes pour 97% avec seulement le Triage. Le triage seul semble suffisant pour améliorer les performances du centre d'appels.

Dans la table 5.11 nous comparons les qualités de service pour les différents scénarios avec la nouvelle configuration au scénario AS-IS de la configuration initiale. Cela permet de se rendre compte de la totalité des gains acquis. On observe des améliorations non négligeables. Dans le cas initial, on pouvait avoir jusqu'à 50% des appels qui n'étaient pas décrochés en moins de 20 secondes sur certains créneaux. La prépriorisation permet de passer à un minimum de 71% des appels décrochés en moins de 20 secondes. Le triage améliore encore ce résultat avec un minimum de 90% des appels décrochés en moins de 20 secondes. Cela correspond donc à une amélioration pouvant aller de 21 à 40% selon le scénario.

Effet sur le temps d'attente maximum

Cette nouvelle configuration permet aussi de réduire le temps maximal avant le décroché des appels. En effet, préempter les ressources sur les appels sortants au profit

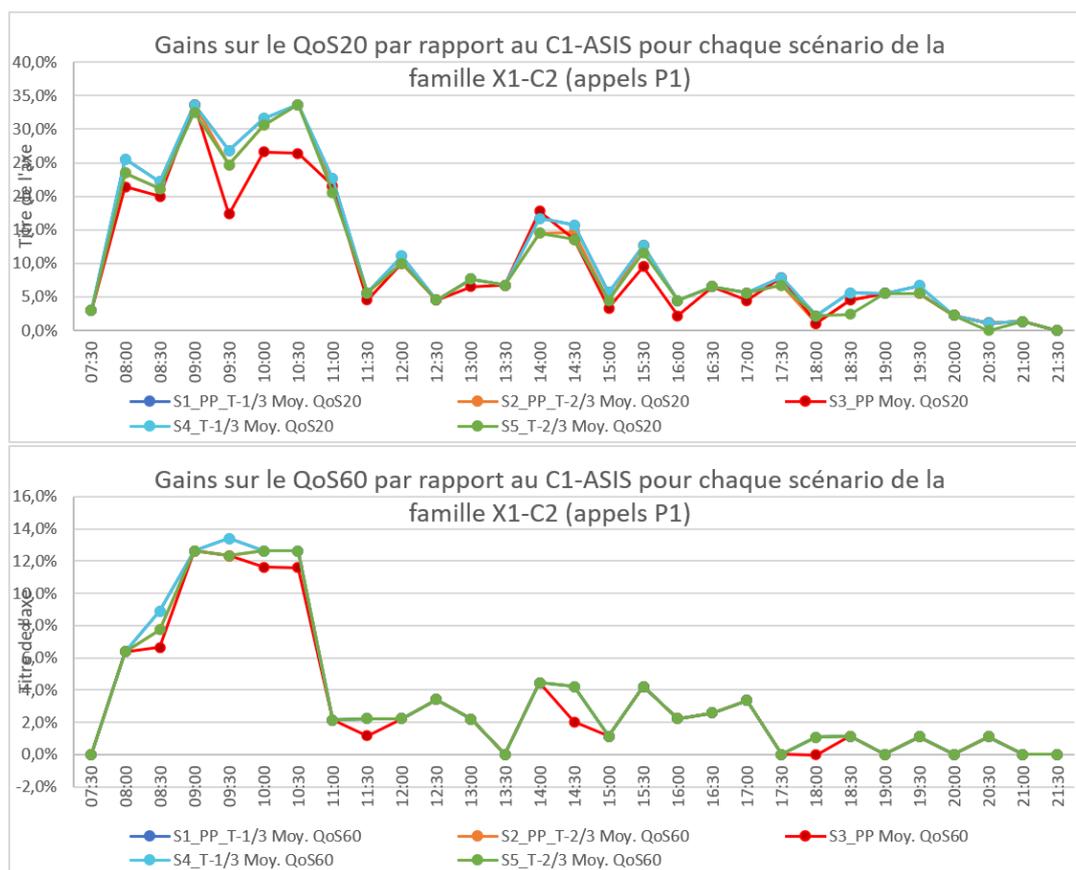


Figure 5.25 – Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C2

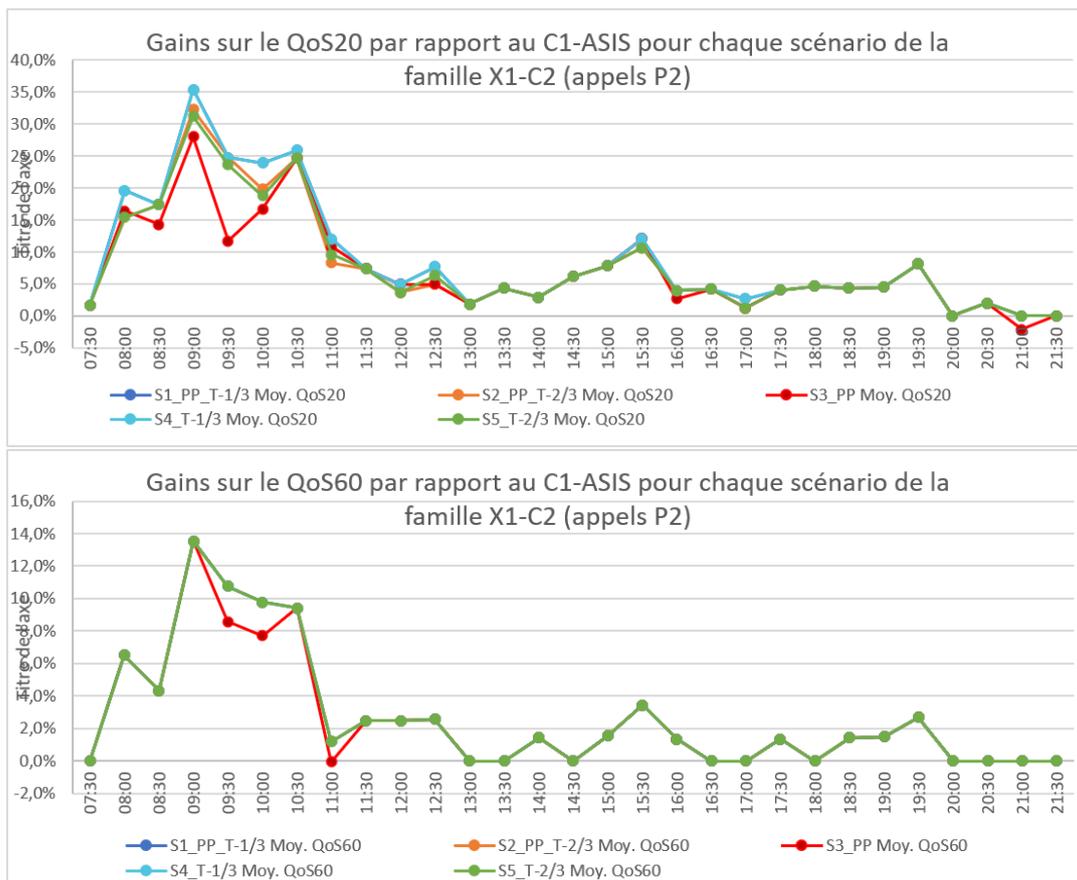


Figure 5.26 – Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C2

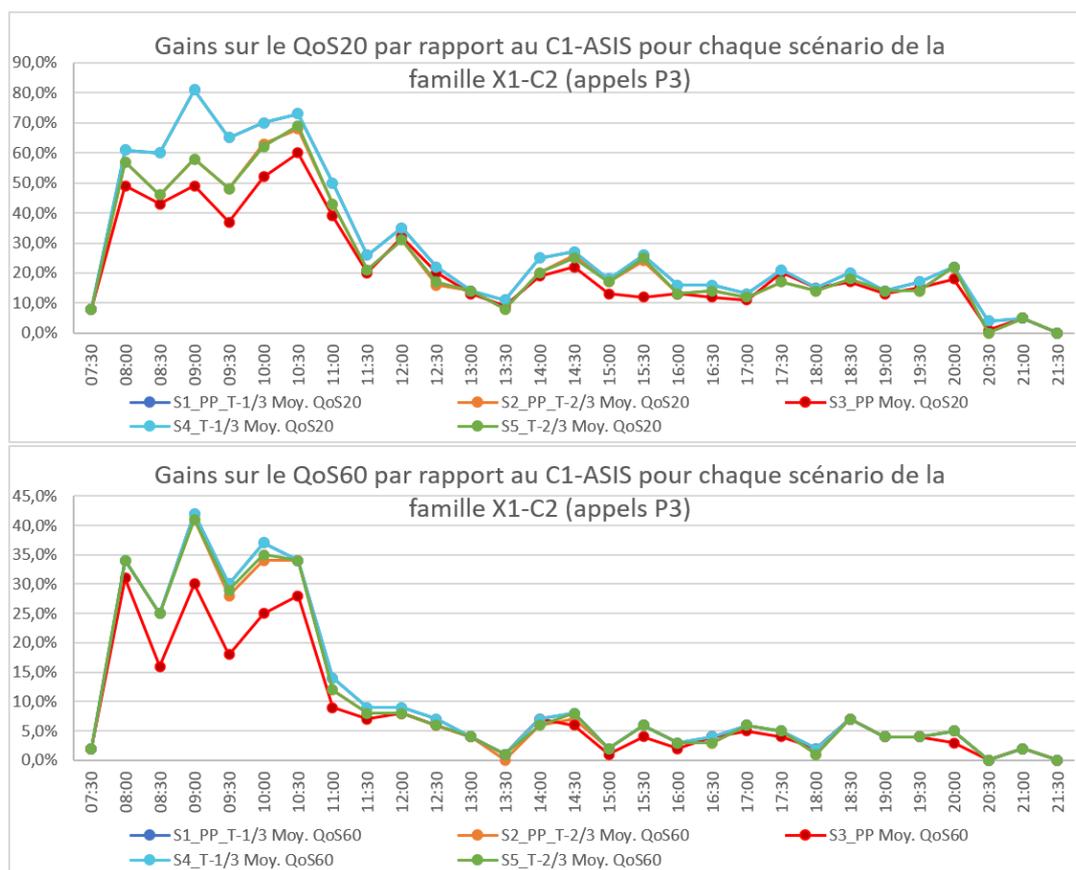


Figure 5.27 – Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C2

Tableau 5.10 – Configuration 2 : Comparaison du QoS entre les scénarios et l’AS-IS.

AS préemptés - Scénarios comparés entre eux														
Creneau	S4-AS-IS		S1 - PT&PP				S2-Pré-Priorisation				S3 - Triage			
	QoS20	QoS60	QoS20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS-IS	QoS60	δ AS-IS
07 :00	1,00	1,00	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%
07 :30	0,99	1,00	1,00	0,9%	1,00	0,0%	0,99	0,0%	1,00	0,0%	1,00	0,9%	1,00	0,0%
08 :00	0,82	0,95	0,98	15,7%	1,00	4,9%	0,82	0,0%	0,95	0,0%	0,98	15,7%	1,00	4,9%
08 :30	0,87	0,96	0,99	12,1%	1,00	3,9%	0,87	0,0%	0,96	0,0%	0,99	11,7%	1,00	3,9%
09 :00	0,70	0,86	0,90	20,3%	0,99	13,4%	0,71	1,4%	0,88	1,7%	0,90	20,3%	0,99	13,4%
09 :30	0,72	0,87	0,94	22,1%	1,00	12,4%	0,74	2,4%	0,89	2,1%	0,93	21,7%	1,00	12,4%
10 :00	0,77	0,87	0,92	15,7%	1,00	12,2%	0,79	2,8%	0,89	1,7%	0,92	15,7%	1,00	12,2%
10 :30	0,81	0,91	0,97	16,1%	1,00	8,9%	0,82	1,1%	0,93	1,4%	0,98	16,4%	1,00	8,9%
11 :00	0,86	0,95	0,98	11,9%	1,00	4,7%	0,88	2,2%	0,96	0,4%	0,98	11,9%	1,00	4,7%
11 :30	0,92	0,97	0,97	5,6%	1,00	2,6%	0,92	0,0%	0,97	-0,7%	0,97	5,6%	1,00	2,6%
12 :00	0,94	0,97	0,99	4,4%	1,00	3,3%	0,94	0,4%	0,97	0,7%	0,98	4,1%	1,00	3,3%
12 :30	0,96	0,98	0,99	3,4%	1,00	1,5%	0,96	0,4%	0,98	0,0%	0,99	3,0%	1,00	1,5%
13 :00	1,00	1,00	0,99	-0,4%	1,00	0,0%	0,99	-0,4%	1,00	0,0%	0,99	-0,4%	1,00	0,0%
13 :30	0,98	0,99	0,98	0,0%	1,00	0,8%	0,98	0,4%	1,00	0,4%	0,98	0,0%	1,00	0,8%
14 :00	0,95	1,00	0,99	4,2%	1,00	0,4%	0,94	-0,8%	0,99	-0,8%	0,99	4,2%	1,00	0,4%
14 :30	0,95	0,98	0,98	2,7%	1,00	1,5%	0,96	0,8%	0,98	0,0%	0,98	2,7%	1,00	1,5%
15 :00	0,96	0,99	0,98	2,4%	1,00	1,2%	0,97	0,8%	1,00	0,8%	0,98	2,4%	1,00	1,2%
15 :30	0,94	0,98	0,96	2,4%	1,00	2,4%	0,93	-0,8%	0,98	0,0%	0,96	2,0%	1,00	2,4%
16 :00	0,97	0,99	1,00	2,6%	1,00	1,1%	0,96	-0,8%	0,99	0,0%	0,99	2,3%	1,00	1,1%
16 :30	0,98	1,00	0,99	0,4%	1,00	0,4%	0,98	0,0%	1,00	0,0%	0,99	0,4%	1,00	0,4%
17 :00	0,96	0,98	1,00	3,8%	1,00	1,9%	0,96	-0,4%	0,97	-0,8%	1,00	3,8%	1,00	1,9%
17 :30	0,97	0,99	0,98	1,9%	1,00	0,8%	0,96	-0,8%	0,99	-0,4%	0,98	1,9%	1,00	0,8%
18 :00	0,95	1,00	0,99	3,9%	1,00	0,4%	0,96	1,2%	0,99	-0,4%	0,99	3,9%	1,00	0,4%
18 :30	0,96	0,98	0,98	2,3%	1,00	1,6%	0,95	-0,8%	0,98	0,0%	0,98	1,9%	1,00	1,6%
19 :00	0,99	1,00	1,00	0,4%	1,00	0,0%	0,99	0,0%	1,00	0,0%	1,00	0,4%	1,00	0,0%
19 :30	0,96	0,99	0,98	1,9%	1,00	0,8%	0,96	0,0%	0,99	0,0%	0,98	1,9%	1,00	0,8%
20 :00	0,96	0,98	0,99	2,8%	1,00	2,0%	0,96	0,0%	0,98	0,0%	0,99	2,8%	1,00	2,0%
20 :30	0,99	1,00	1,00	1,3%	1,00	0,4%	0,99	0,0%	1,00	0,4%	1,00	1,3%	1,00	0,4%
21 :00	0,99	0,99	1,00	0,5%	1,00	0,9%	0,99	0,0%	0,99	0,0%	1,00	0,5%	1,00	0,9%
21 :30	1,00	1,00	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%

Tableau 5.11 – Comparaison des QoS entre les scénarios de la nouvelle configuration et l’AS-IS de la configuration originale.

AS préempté - Scénario comparé à l’AS-IS sans préemption														
	S4-AS-IS_1		S1 - PT&PP				S2-Pré-Priorisation				S3 - Triage			
	QoS20	QoS60	QoS20	δ AS-IS-1	QoS60	δ AS-IS-1	QoS20	δ AS-IS	QoS60	δ AS-IS-1	QoS20	δ AS-IS-1	QoS60	δ AS-IS-1
07 :00	0,99	1,00	1,00	0,5%	1,00	0,0%	1,00	0,5%	1,00	0,0%	1,00	0,5%	1,00	0,0%
07 :30	0,95	0,99	1,00	4,9%	1,00	0,9%	0,99	4,0%	1,00	0,9%	1,00	4,9%	1,00	0,9%
08 :00	0,63	0,84	0,98	34,3%	1,00	15,7%	0,82	18,5%	0,95	10,8%	0,98	34,3%	1,00	15,7%
08 :30	0,64	0,85	0,99	34,8%	1,00	14,9%	0,87	22,7%	0,96	11,0%	0,99	34,4%	1,00	14,9%
09 :00	0,50	0,76	0,90	40,2%	0,99	23,4%	0,71	21,3%	0,88	11,7%	0,90	40,2%	0,99	23,4%
09 :30	0,59	0,82	0,94	34,8%	1,00	17,6%	0,74	15,2%	0,89	7,2%	0,93	34,5%	1,00	17,6%
10 :00	0,56	0,82	0,92	36,6%	1,00	18,1%	0,79	23,7%	0,89	7,7%	0,92	36,6%	1,00	18,1%
10 :30	0,58	0,81	0,97	38,9%	1,00	18,6%	0,82	23,9%	0,93	11,1%	0,98	39,3%	1,00	18,6%
11 :00	0,67	0,91	0,98	30,4%	1,00	9,1%	0,88	20,7%	0,96	4,7%	0,98	30,4%	1,00	9,1%
11 :30	0,83	0,94	0,97	14,8%	1,00	5,9%	0,92	9,2%	0,97	2,6%	0,97	14,8%	1,00	5,9%
12 :00	0,82	0,96	0,99	17,0%	1,00	4,4%	0,94	12,9%	0,97	1,8%	0,98	16,6%	1,00	4,4%
12 :30	0,89	0,97	0,99	9,8%	1,00	3,0%	0,96	6,8%	0,98	1,5%	0,99	9,4%	1,00	3,0%
13 :00	0,93	0,99	0,99	6,5%	1,00	1,2%	0,99	6,5%	1,00	1,2%	0,99	6,5%	1,00	1,2%
13 :30	0,90	0,99	0,98	8,1%	1,00	1,2%	0,98	8,5%	1,00	0,8%	0,98	8,1%	1,00	1,2%
14 :00	0,81	0,93	0,99	18,1%	1,00	6,6%	0,94	13,1%	0,99	5,4%	0,99	18,1%	1,00	6,6%
14 :30	0,82	0,96	0,98	15,8%	1,00	3,5%	0,96	13,8%	0,98	1,9%	0,98	15,8%	1,00	3,5%
15 :00	0,89	0,98	0,98	9,1%	1,00	1,6%	0,97	7,5%	1,00	1,2%	0,98	9,1%	1,00	1,6%
15 :30	0,79	0,96	0,96	17,0%	1,00	3,6%	0,93	13,8%	0,98	1,2%	0,96	16,6%	1,00	3,6%
16 :00	0,90	0,97	1,00	9,4%	1,00	2,6%	0,96	6,0%	0,99	1,5%	0,99	9,1%	1,00	2,6%
16 :30	0,89	0,98	0,99	10,0%	1,00	2,0%	0,98	9,6%	1,00	1,6%	0,99	10,0%	1,00	2,0%
17 :00	0,89	0,98	1,00	10,6%	1,00	2,3%	0,96	6,5%	0,97	-0,4%	1,00	10,6%	1,00	2,3%
17 :30	0,89	0,98	0,98	9,8%	1,00	2,3%	0,96	7,2%	0,99	1,1%	0,98	9,8%	1,00	2,3%
18 :00	0,91	0,99	0,99	8,6%	1,00	1,2%	0,96	5,8%	0,99	0,4%	0,99	8,6%	1,00	1,2%
18 :30	0,87	0,97	0,98	11,2%	1,00	2,7%	0,95	8,1%	0,98	1,2%	0,98	10,9%	1,00	2,7%
19 :00	0,95	0,99	1,00	4,3%	1,00	1,2%	0,99	3,9%	1,00	1,2%	1,00	4,3%	1,00	1,2%
19 :30	0,86	0,97	0,98	11,7%	1,00	3,4%	0,96	9,8%	0,99	2,7%	0,98	11,7%	1,00	3,4%
20 :00	0,92	0,97	0,99	7,3%	1,00	2,8%	0,96	4,5%	0,98	0,8%	0,99	7,3%	1,00	2,8%
20 :30	0,96	0,99	1,00	4,2%	1,00	0,8%	0,99	2,9%	1,00	0,8%	1,00	4,2%	1,00	0,8%
21 :00	0,98	0,99	1,00	1,9%	1,00	1,4%	0,99	1,4%	0,99	0,5%	1,00	1,9%	1,00	1,4%
21 :30	1,00	1,00	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%

des appels entrants permet de réduire le temps d'attente maximal à 58,3 secondes avec le scénario de triage et à 116,6 secondes avec le scénario de prépriorisation. Cette amélioration engendre un gain de temps pouvant aller respectivement jusqu'à 109 et 152 secondes pour la prépriorisation et le Triage. C'est une amélioration de plus de deux minutes par rapport à l'AS-IS de la configuration initiale. Les scénarios de triage et de prépriorisation apportaient déjà une amélioration significative pouvant aller jusque 108 secondes. Avec la nouvelle règle de gestion des appels sortants il est possible de gagner 48 secondes supplémentaires. La figure 5.28 représente la comparaison du temps maximum d'attente avant décroché entre les deux configurations, pour chaque scénario, pour les appels de gravité supérieure. Encore une fois, nous ne présentons les résultats qu'à partir de 07h00 jusque 21h30, car les créneaux précédents ne sont pas significatifs. Sur cette figure nous avons fait apparaître deux lignes de références correspondant au QoS_{20} et au QoS_{60} . Si la courbe est au-dessus de cette ligne alors cela signifie qu'il y a des appels qui sont décrochés en plus de 20 ou 60 secondes. Plusieurs résultats ressortent de ce graphique :

- Dans la deuxième configuration, il y a deux scénarios qui permettent de faire en sorte que tous les appels soient décrochés en moins de 60 secondes sur chaque créneau de la journée. Ce sont les scénarios de triage et celui combinant triage et prépriorisation.
- Le scénario de prépriorisation seul donne des résultats assez peu satisfaisants, car il donne des résultats assez similaires à l'AS-IS de la deuxième configuration, sur certains créneaux on observe une diminution du temps maximum pouvant aller jusque 68% de celui du AS-IS. À 12h00 on passe d'un maximum à 121,6 secondes à 38,6 secondes. On note toutefois une dégradation du temps sur le créneau de 13h30 où le temps maximal du AS-IS est de 20,9 secondes, et avec le scénario de prépriorisation il passe à 86,7 secondes. Ce résultat étant surprenant, nous allons étudier le log de données plus en détail afin de comprendre ce qu'il s'est passé pour avoir une telle baisse de performance.
- Concernant le scénario de triage, il y a quelques créneaux où le temps d'attente est supérieur à celui du scénario AS-IS, lorsque ce temps d'attente était supérieur à 20 secondes nous avons aussi vérifié dans le log la raison de cette modification.

Prenons l'appel 304 de la réplique 89. Il arrive dans le centre d'appels à 13h03m15 et ne sera décroché que 40,7 secondes plus tard. La figure 5.29 détaille les activités par lesquelles passe chaque appel. Il est ainsi facile de retracer ce qu'il s'est passé et pourquoi l'appel 304 n'a pas pu être décroché plus tôt. On voit sur la figure que l'ARM 1 commence à remplir le dossier de l'appel 302 à 13h03m11 et il finit à 13h03m53. L'ARM 2 priorise l'appel 303 de 13h02m59 à 13h03m21, il termine son activité après le décroché du 304. L'ARM 3 est en train de remplir le dossier du 301 entre 13h02m48 et 13h04m04. L'ARM 4 remplit le dossier du 300 entre 13h02m28 et 13h03m26 appel de priorité 1 qui génère ensuite un appel interne que l'ARM 4 continue de traiter. C'est donc l'ARM 1 qui après avoir traité l'appel 302 va décrocher l'appel 304 pour le traiter. Ce temps d'attente long n'est pas dû à la nature du scénario, mais à la disponibilité des ressources. Nous sommes typiquement dans un cas où beaucoup d'appels arrivent en même temps. Pour palier ces attentes on pourrait envisager d'autoriser l'interruption de l'activité *Remplir le dossier* pour les appels de gravité 3.

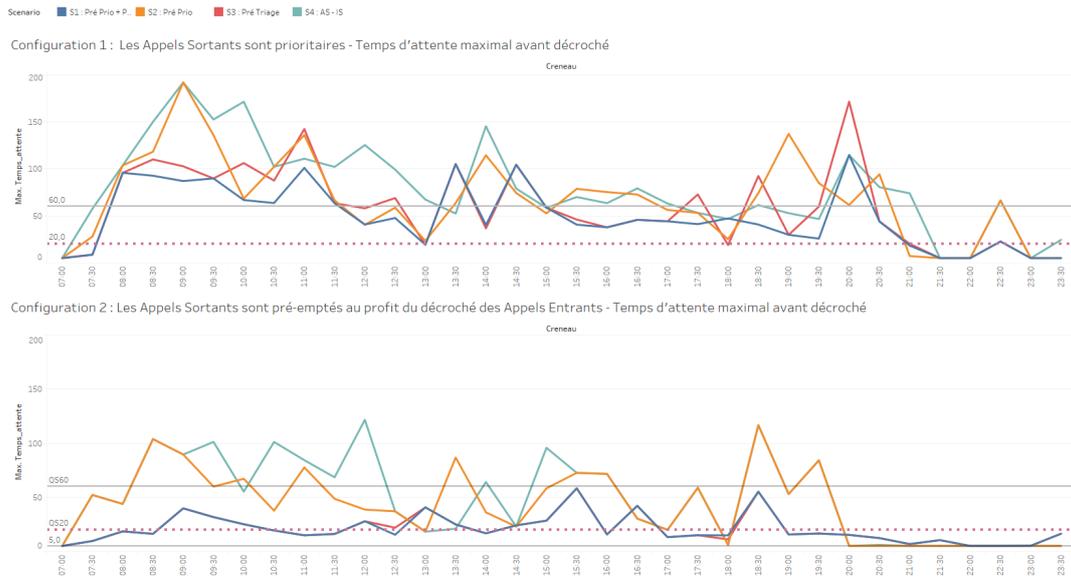


Figure 5.28 – Comparaison des temps d’attentes maxima pour les appels de gravité P1 entre les deux configurations

Conclusions sur les expérimentations

La modification de l’organisation du décroché de l’appel permet des améliorations sensibles sur la réduction du temps d’attente. Bien que ces gains doivent être mis en regard des hypothèses faites lors de la création du modèle, ils restent significatifs.

On identifie notamment une réduction du temps maximal d’attente pour les appels les plus urgents (P1) pouvant aller jusque 40% sur certains créneaux. Ce qui correspond à une diminution pouvant aller jusqu’à 1min30 sur certains appels. De plus les modifications permettent de respecter les critères nationaux de décrochés des appels et faire en sorte que tous les appels soient décrochés en moins de 60 secondes. Ainsi une combinaison idéale serait de mettre en place un triage des appels combiné à la prépriorisation de ceux-ci et à la possibilité d’interrompre les appels sortants au profit des appels entrant. Cette combinaison permet des gains significatifs notamment pour les appels de la priorité P1 qui nous le rappelons sont les plus critiques

5.2.5 Gérer une crise : comment le centre d’appels réagit-il confronté à une augmentation soudaine et significative des appels ?

Dans cette section, nous souhaitons évaluer la réaction du centre d’appels face à une augmentation significative et brutale des appels tout au long de la journée. Nous proposons de doubler le profil d’arrivée des appels ce qui pourrait correspondre à

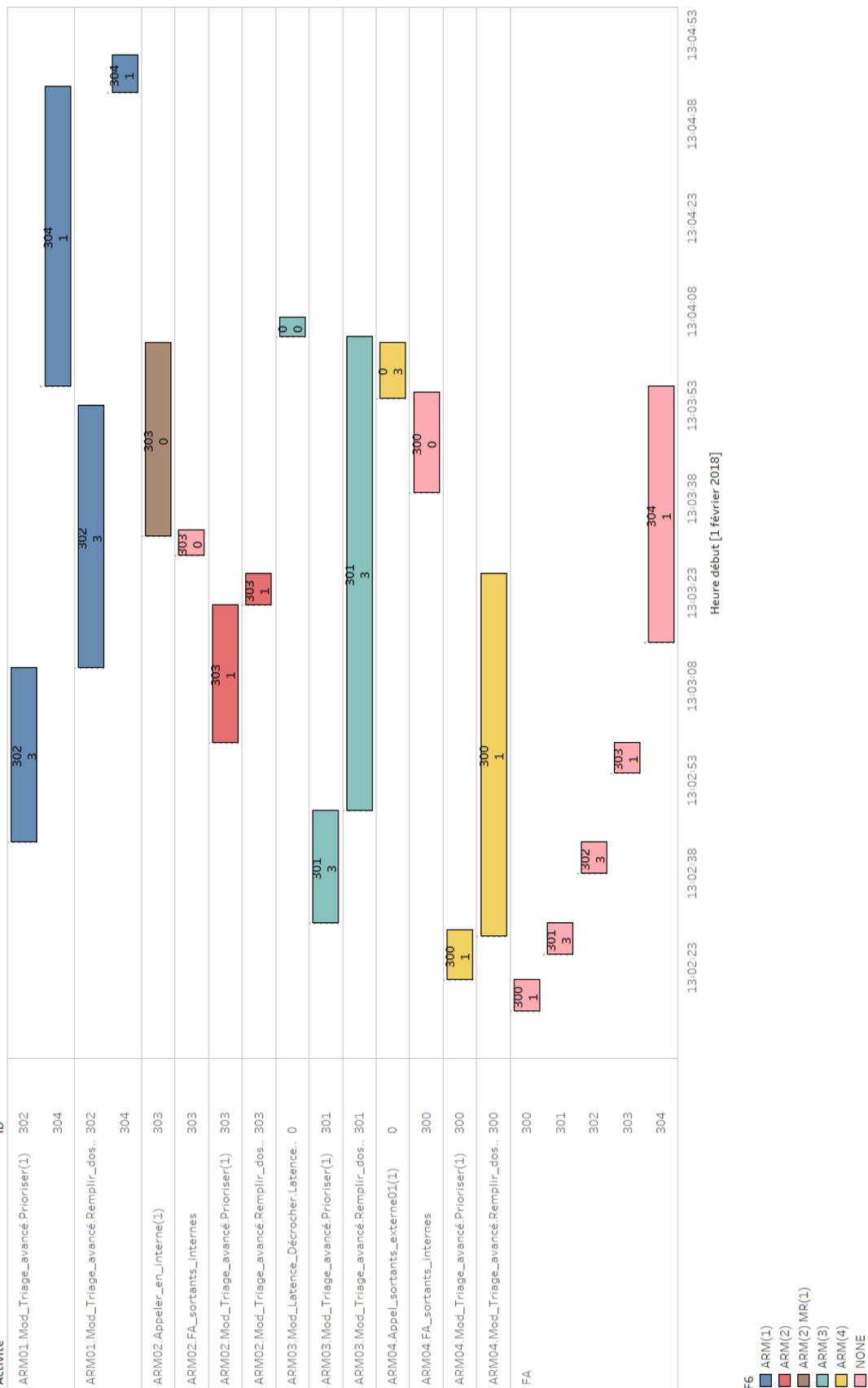


Figure 5.29 – Détail de la réplication 89 pour le scénario de Triage dans la configuration 2

un scénario d'épidémie par exemple de grippe saisonnière. Nous regardons l'effet de cette augmentation sur chacun des créneaux horaires en termes de temps d'attente maximum, de QoS_{20} et de QoS_{60} . Nous étudions dans un premier temps l'effet sur la configuration C1, puis nous appliquons la configuration C2. Les données sont analysées pour chacun des scénarios de prépriorisation, de triage ainsi que pour les scénarios mixtes. Si on regarde la courbe d'arrivée du nombre d'appels par créneau (voir figure 5.30) on constate comme précédemment qu'on a plusieurs périodes. De 00h00 à 7h30, durant la nuit profonde il y a assez peu d'appels et la qualité de service est plutôt bonne. Puis, il y a un premier pic d'appels entre 8h00 et 10h00 et un second pic d'appels entre 14h et 16h. Suite à cela, la situation se stabilise vers 21h30 avec une diminution du nombre d'appels. Pour la suite, et comme précédemment nous réduisons l'horizon d'étude aux créneaux entre 7h30 et 21h30 qui sont donc les plus intéressants, car c'est là où le centre d'appels montre des difficultés de décroché des appels.

Doubler la demande sur la configuration C1

Effet sur la qualité de service

Le fait de doubler le nombre d'appels reçus a un effet très négatif sur la qualité de service. La figure 5.30 présente les valeurs de QoS pour chaque scénario ainsi que le nombre moyen d'appels sur chaque créneau. On constate que la qualité de service diminue lorsque le nombre d'appels augmente. Que ce soit le QoS_{20} ou le QoS_{60} , dès 8h00 ils valent respectivement 5,7% et 22,8% (voir figure 5.30). Cela implique que 81,2% des appels attendent plus de 60 secondes dès 8h00 si le nombre d'appels double. Le centre d'appels est donc rapidement saturé si on reste avec un nombre de ressources fixes. Le pire créneau est celui de 10h00 où le QoS_{20} et le QoS_{60} valent respectivement 1,6% et 8,7%. On constate que tous les scénarios testés engendrent une amélioration globale de la qualité de service et permettent de réduire l'effet de l'augmentation de la durée d'attente des appels.

La figure 5.31 présente les gains par rapport au scénario AS-IS pour les appels de priorité 1. On constate que les scénarios mixtes qui combinent prépriorisation et triage permettent d'aller chercher des gains supérieurs à 50% du scénario initial pour le QoS_{60} , cela implique qu'après l'application de ces scénarios, dans le pire des cas 40% des appels sont décrochés en moins de 60 secondes. Même si ce n'est pas suffisant vis-à-vis des objectifs de performance en régime nominal, cela permet d'atténuer significativement l'effet néfaste de l'augmentation du nombre d'appels sur les performances du centre d'appels en situation de crise.

Ensuite, c'est le scénario de prépriorisation qui permet les meilleurs gains sur le QoS_{60} avec jusqu'à 50% d'amélioration sur les créneaux les plus chargés. Il est intéressant de noter que ce n'est pas le cas pour le QoS_{20} où la prépriorisation s'avère moins performante avec des gains moindres et une amélioration au maximum de 15%. Ce qui le rend moins performant que les scénarios mixtes qui vont chercher un gain pouvant aller jusque 25%.

Les scénarios de triage seuls permettent d'améliorer le QoS_{60} de 30%. On note que lorsqu'on augmente le nombre d'appels, le fait d'avoir un triage plus long (ce qui est



Figure 5.30 – Qualité de Service pour chacun des scénarios. Demande double, Configuration C1

le cas pour le scénario $S5_T - 2/3$) donne des écarts de performance plus importants que ce que l'on avait pu constater avec un nombre d'appels normal. Concernant le QoS_{20} , les scénarios de triage sont meilleurs que celui de prépriorisation avec des gains supérieurs à 10% après 12h30.

Concernant les appels de priorité P2, les gains sur la qualité de service sont représentés dans la figure 5.32. On constate que c'est encore une fois les scénarios mixtes qui permettent d'obtenir les meilleurs résultats et qui réduisent le QoS_{60} de 50% au maximum lors du premier pic d'appels. et lors du deuxième pic d'appels, on gagne 20% par rapport au scénario AS-IS. Le QoS_{20} est amélioré de 30% sur le premier pic d'appels ainsi que sur le second pic d'appels.

Le scénario de prépriorisation permet une amélioration au global du QoS_{60} avec des gains qui vont jusque 35% sur le premier pic d'appels et des gains de 10% sur le deuxième pic d'appels. Concernant le QoS_{20} on observe des gains de 15% lors du premier pic d'appels et des gains de 10% lors du deuxième pic d'appels. Le scénario de prépriorisation est plus efficace que le scénario T-2/3 sur le premier pic d'appels, sur le deuxième pic les performances sont équivalentes.

Les scénarios de triage T-1/3 et T-2/3 améliorent le QoS_{60} entre 20% et 30% durant le premier pic, on note une différence de performance de 20% entre T-1/3 et T-2-3, le fait d'augmenter le temps de triage va effectivement avoir un effet plutôt important sur la qualité de service lorsque le centre d'appels est saturé. Lors du deuxième pic d'appels T-1/3 améliore QoS_{60} de 20% et T-2/3 l'améliore de 10%. Concernant

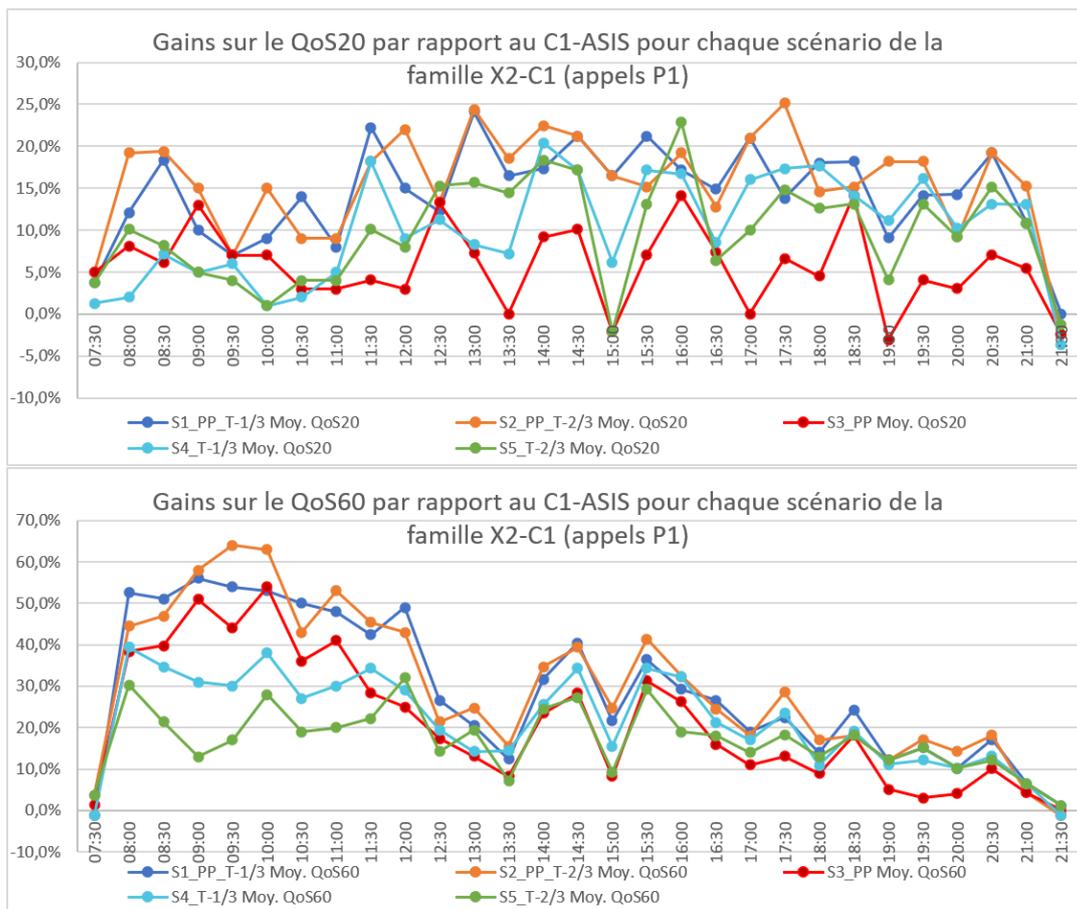


Figure 5.31 – Demande doublée : Gains sur les QoS_{20} et QoS_{60} pour les appels P1 dans la configuration C1

QoS_{20} les scénarios de triage permettent d'obtenir 5% d'amélioration lors du premier pic et au maximum 15% lors du deuxième pic. Le triage seul reste moins efficace que les scénarios mixtes.

La figure 5.20 montre les gains pour les appels de priorité 3. Pour le QoS_{60} , le premier constat est que le scénario de prépriorisation a un effet négatif sur le décroché des appels P3 (au bénéfice des appels de priorité supérieur). Le constat est similaire pour le QoS_{20} : il y a quelques gains qui allant jusque 5%, mais il y a des effets négatifs pouvant aller jusque -5%.

Les scénarios de triage et les scénarios mixtes permettent une amélioration du QoS_{60} surtout après le premier pic d'appels et durant le deuxième. Durant le premier pic d'appels il n'y a pas vraiment de scénario qui permettent d'améliorer le décrocher des appels P3, car la priorité est mise sur les appels P1 et P2. Toutefois on note quand même des améliorations pouvant aller jusque 20% pour le QoS_{20} avec le scénario T-1/3, et jusque 30% sur le QoS_{60} avec les deux scénarios de triage.

Effet sur le temps d'attente maximum

Le temps d'attente maximum si on double les appels monte jusque 800 secondes lors du premier pic d'appels, puis 400 secondes lors du second pic d'appels. Le temps d'attente maximum est globalement le même que ce soit pour les appels P1, P2 ou P3 concernant le AS-IS. Ce qui est logique puisqu'il n'y a pas de priorisation entre les types d'appels.

Pour les appels P1 et P2, ce sont les scénarios mixtes et le scénario de prépriorisation qui permettent d'obtenir la meilleure réduction du temps d'attente maximal dans la première configuration. Ils permettent notamment de maintenir l'attente en dessous de 200 secondes. Les scénarios de triage seuls sont peu performants et on note des temps d'attente pouvant aller jusque 600 secondes lors du premier pic d'appels.

Pour les appels P3, on constate que le scénario de prépriorisation engendre un temps d'attente plus important que les autres scénarios. Lors du pic d'appels, ce sont les scénarios PP-T-1/3 et T1/3 qui permettent la meilleure réduction du temps d'appels.

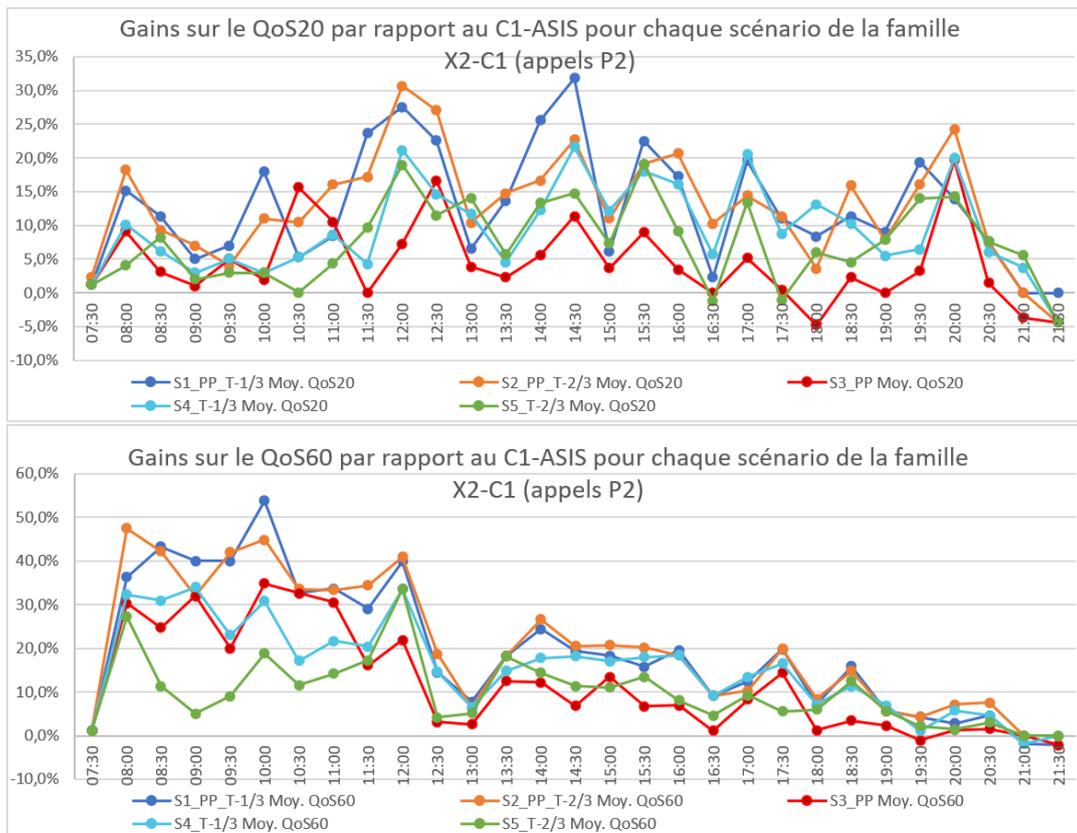


Figure 5.32 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C1

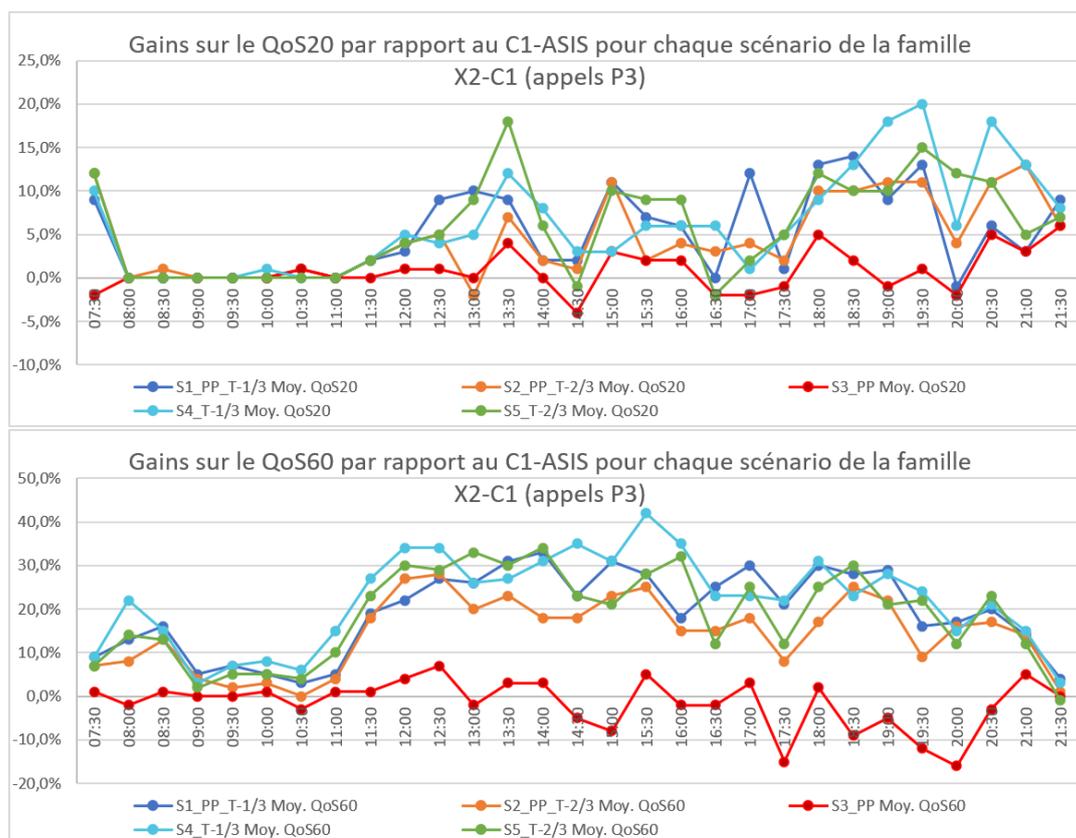


Figure 5.33 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C1

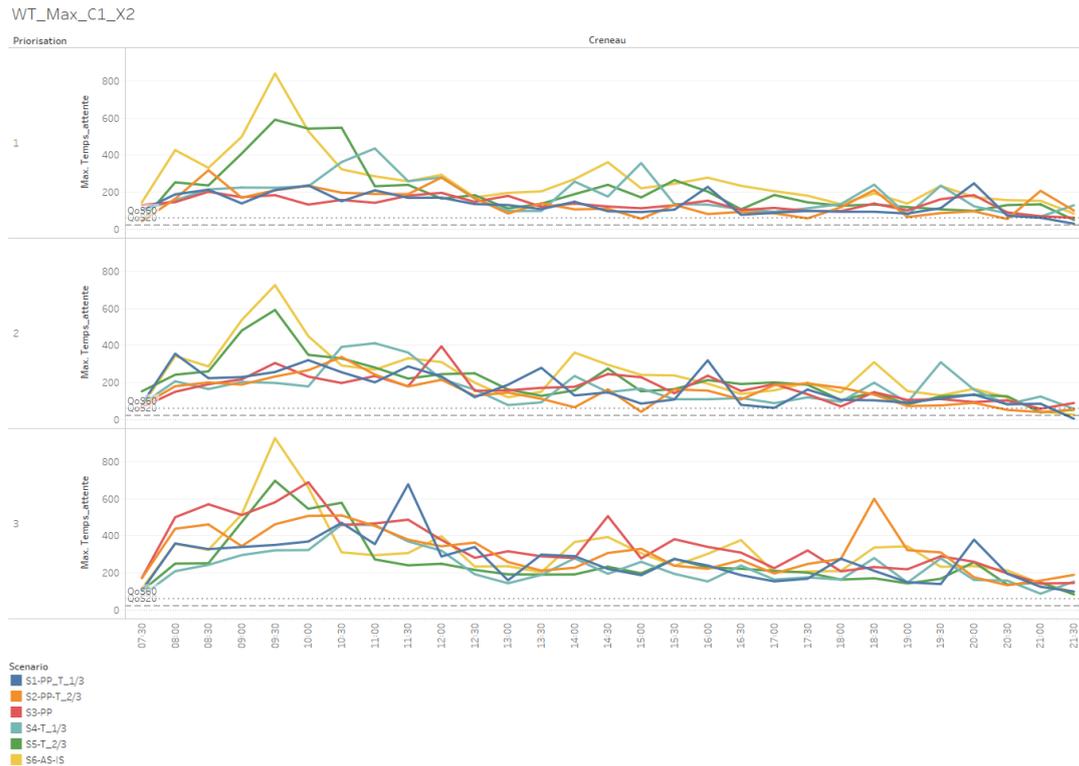


Figure 5.34 – Effet de l’augmentation du nombre d’appels sur le temps d’attente maximum en configuration C1

Doubler la demande sur la configuration C2

Nous avons déjà pu constater que les scénarios proposés permettent d’améliorer la prise en charge des appels lorsque le nombre d’appels double. Même si cela n’est pas suffisant pour faire en sorte que le centre d’appels ne soit pas saturé. Nous souhaitons maintenant tester l’effet des scénarios dans la configuration C2 (cf. section 5.2.4) et voir la réaction du système à l’augmentation du nombre d’appels. La figure 5.35 montre la valeur du QoS_{20} et du QoS_{60} pour chaque scénario. On observe qu’au global deux types de scénarios se démarquent : le scénario mixte PP_T-1/3 et le scénario de triage T-1/3 et les autres scénarios. Au niveau du QoS_{20} les deux premiers scénarios permettent de le maintenir au-dessus de 60% là où l’AS-IS descend jusqu’à 20% au plus haut du premier pic d’appels. Les autres scénarios de PP, et de T-2/3 améliorent la situation, mais uniquement d’une dizaine de %. Les différences entre les scénarios sont plus nettes au niveau du QoS_{60} . En effet, au niveau du premier pic d’appels, on constate que le scénario de prépriorisation améliore le QoS_{60} de 16% , le scénario de T-2/3 l’améliore de 32%, le scénario mixte PP-T-2/3 l’améliore de 37 % et les deux scénarios PP-T1/3 et T1/3 permettent de conserver un QoS_{60} supérieur à 90%.

Effet sur la qualité de service

Les figures 5.36, 5.37 et 5.38 présentent les gains sur chaque type d’appels comparé au scénario AS-IS initial, soit celui de la première configuration.

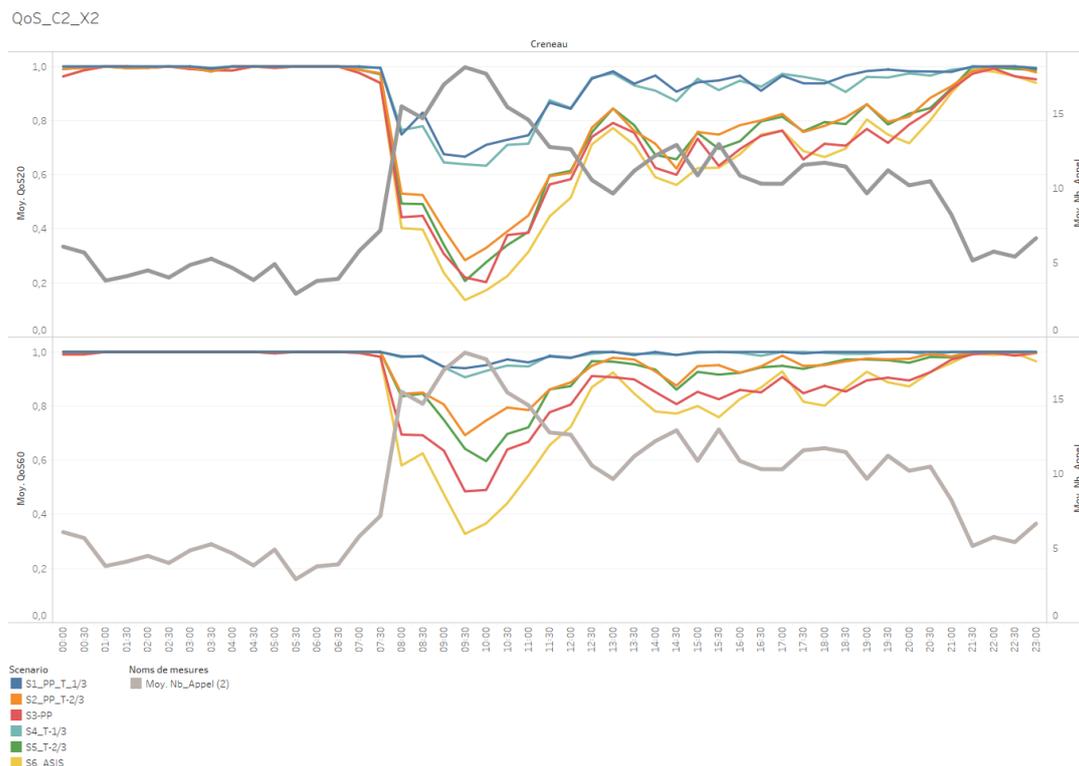


Figure 5.35 – Qualité de Service pour chacun des scénarios. Demande double, Configuration C2

Pour les appels de type P1 et P2, on retrouve les deux scénarios T-1/3 et PP-T-1/3 qui permettent une amélioration de 90% par rapport au scénario initial. Ces scénarios permettent de faire en sorte que même si on double le nombre d'appels, 95% des appels P1 sont décrochés en moins de 20 secondes. Les scénarios de PP-T-2/3, T-2/3 et PP permettent aussi une amélioration par rapport au AS-IS qui permet de gagner 50% d'appels décrochés en moins de 20 secondes lors du pic d'appels. Au niveau du QoS_{60} ce sont les scénarios mixtes et le scénario T-1/3 qui sont les plus efficaces avec une amélioration très sensible de 80 à 90 % lorsque le pic d'appels est le plus important.

Pour les appels de type P3, on améliore un peu moins le QoS_{20} car la priorité est une fois encore mise sur les appels P1 et P2. Toutefois ce sont quand même des gains de près de 50% qui peuvent être observés durant le premier pic d'appels pour les scénarios PP-T1/3 et T-1/3. Lors du deuxième pic d'appels, on observe des gains de 80% qui permettent de faire en sorte que 95% des appels soient décrochés en moins de 20 secondes. Pour le QoS_{60} on note la faiblesse du scénario de prépriorisation qui ne permet pas d'améliorer autant les performances des appels P3 par rapport aux autres scénarios.

Effet sur le temps d'attente maximum

Avec la mise en place de la configuration C2, il est intéressant de noter que le scénario PP-T-1/3 permet de réduire le temps d'attente maximal à 162,5 secondes. De plus, si ce n'est pour trois créneaux, on note que le temps d'attente maximal est inférieur

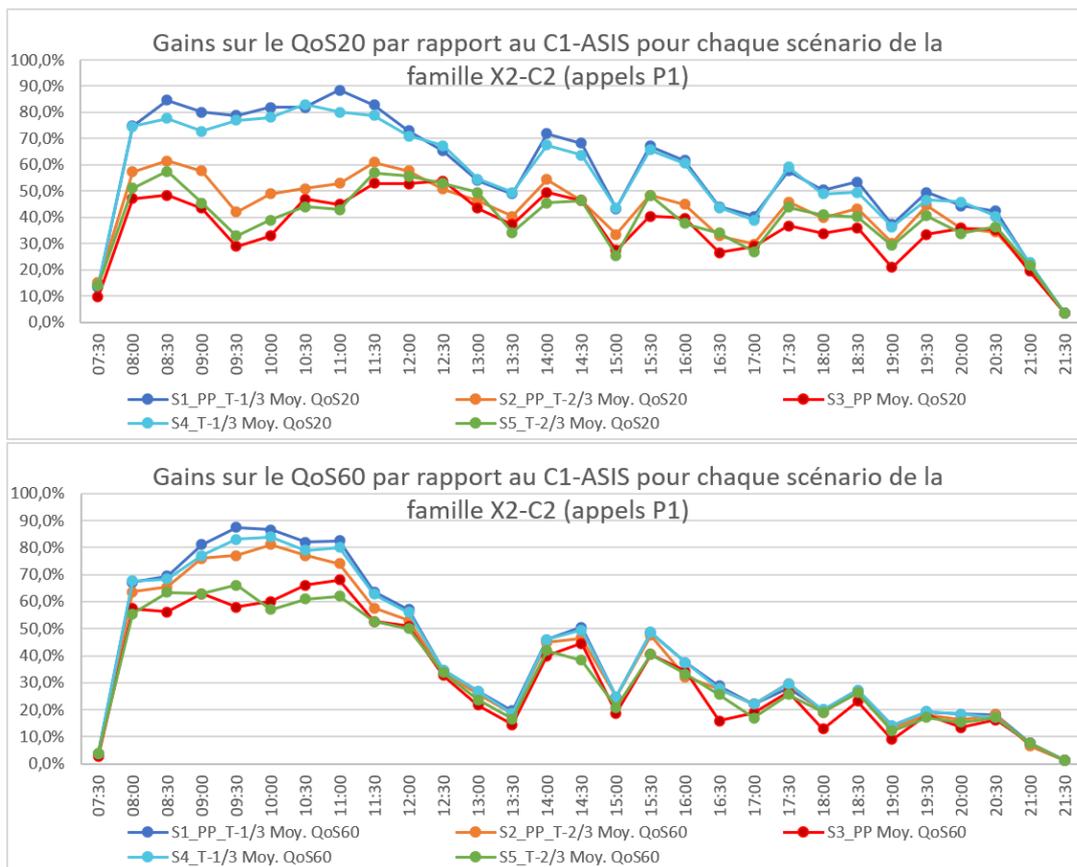


Figure 5.36 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C2

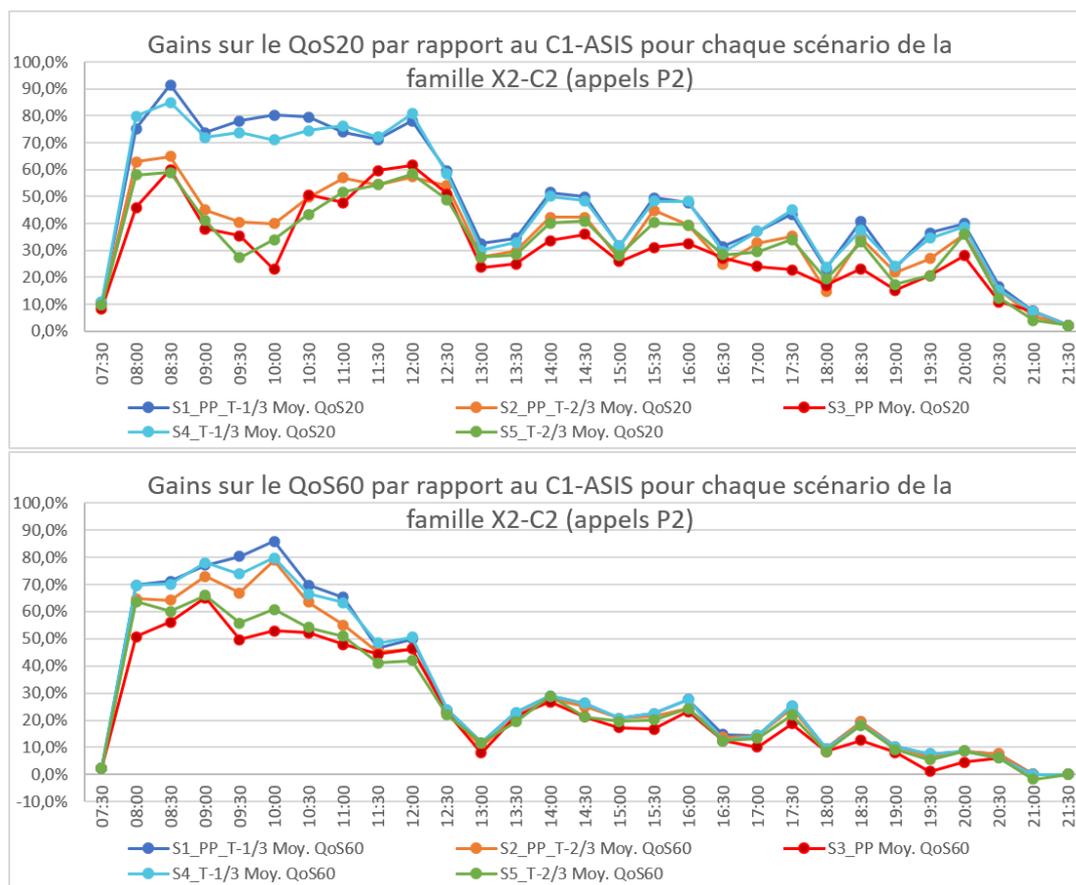


Figure 5.37 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C2

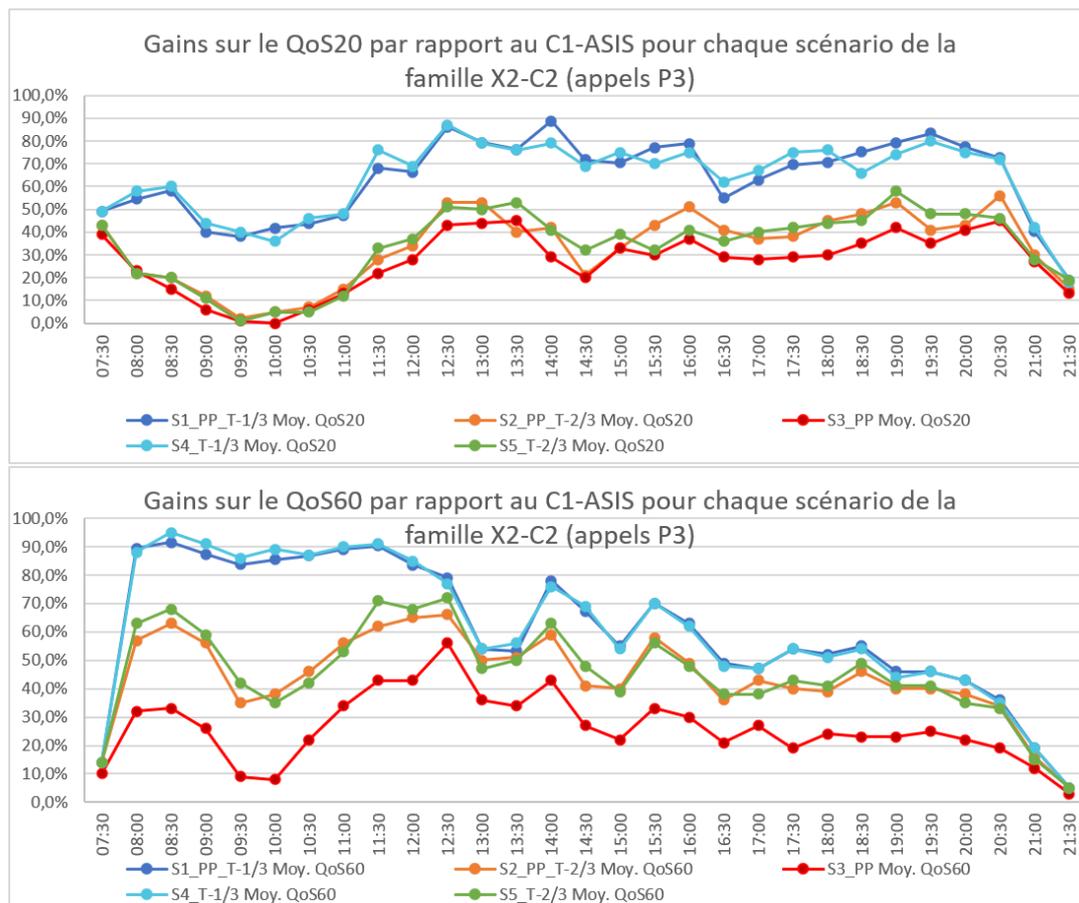


Figure 5.38 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C2

à 60 secondes, donc le QoS_{60} est de 100% sur ces créneaux pour les appels P1 et P2. Le scénario de T-1/3 maintient temps d'attente maximum à 244,6 secondes lors du premier pic d'appels puis dès 12h, le temps d'attente reste inférieur à 60 secondes pour les appels P1 et P2. Le scénario mixte PP-T-2/3 permet de réduire le temps d'attente maximum durant le premier pic à moins de 200 secondes. On note néanmoins un pic d'attente à 326 secondes à 11h30.

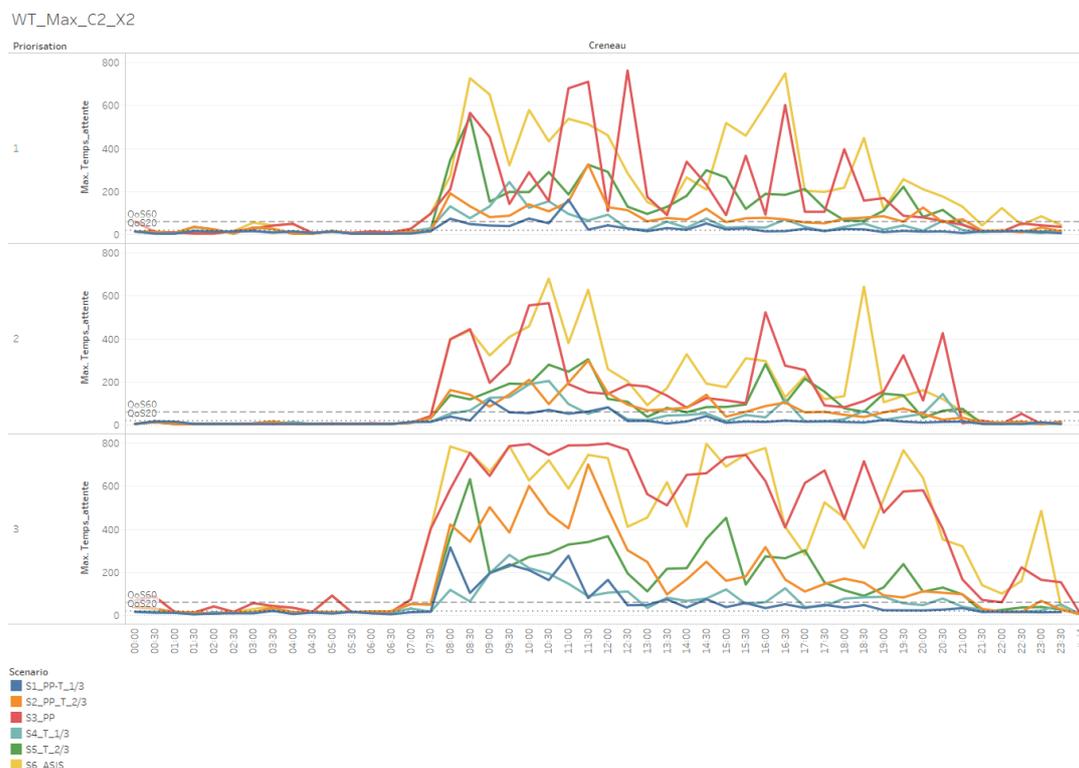


Figure 5.39 – Effet de l'augmentation du nombre d'appels sur le temps d'attente maximum en configuration C1

Dans le cas où le nombre d'appels double, on constate que la mise en place de scénarios de modification de l'organisation a un effet positif et tend à améliorer significativement les performances du système.

Le premier constat est qu'il semble évident que la priorité doit être mise sur le décroché des appels entrants, le fait de pouvoir interrompre un appel sortant afin de décrocher un appel entrant améliore le QoS_{20} moyen sur la journée de 51% à 71% et le QoS_{60} de 66% à 83%.

Le scénario de prépriorisation permet de réduire le temps d'attente des appels au global. Toutefois, tout comme pour l'AS-IS, il a tendance à légèrement augmenter le temps d'attente des appels P3. On note néanmoins une légère amélioration du QoS_{20} de 5% et du QoS_{60} de 7%. La faiblesse principale de ce scénario réside dans la difficulté de sa mise en oeuvre car il s'appuie sur des nouvelles technologies qui permettent d'obtenir de l'information supplémentaire

sur la gravité de l'appel en amont du décroché. Ce scénario serait donc un objectif à atteindre en complément du triage.

Le scénario de triage présente par contre un réel intérêt pour les centres d'appels d'urgence, il est plutôt facile à mettre en place, car il ne nécessite pas d'information supplémentaire en amont de l'arrivée de l'appel dans le centre. On peut soit envisager d'affecter un ARM spécifiquement à ce poste, soit demander aux ARM d'interrompre les appels courants pour décrocher les nouveaux, ce qui peut être humainement plus éprouvant. Si on arrive à garder le temps de triage en deçà d'un certain seuil (qui n'est pas identifié pour le moment) on peut prétendre à maintenir un QoS_{20} à 95% et un QoS_{60} supérieur à 95% même si on double le nombre d'appels. Il serait donc intéressant d'identifier le temps optimal de priorisation de l'appel pour lequel on observe une réduction de la qualité de service tout en laissant un temps suffisant à l'ARM pour identifier la gravité de l'appel.

Les scénarios mixtes de prépriorisation et de triage sont également intéressants et obtiennent en moyenne sur la journée les mêmes résultats que le triage seul. Dans la mesure où la prépriorisation n'est pas aisée à mettre en place, il serait plus intéressant pour les centres d'appels de mettre en place du triage qui en moyenne sur la journée permet d'obtenir un QoS_{60} de 100% ainsi qu'un QoS_{20} de 99%. Ainsi, les objectifs de performances fixés par la HAS seraient respectés.

5.3 Synthèse

Dans ce chapitre, nous avons proposé des scénarios de nouvelles organisations pour les SAMU. Ces scénarios ont été évalués au regard de plusieurs indicateurs de performance le QoS_{20} et le QoS_{60} ainsi que le temps d'attente maximal. Les contributions principales ainsi que les limites de ce chapitre sont présentées dans le tableau 5.12. La simulation de scénarios de triage, de prépriorisation et de scénarios mixtes en régime nominal et lors d'une augmentation soudaine du nombre d'appels constitue la contribution principale de ce chapitre. Plusieurs résultats intéressants doivent être soulignés :

- La mise en place d'un triage rapide des appels permet de respecter les objectifs gouvernementaux pour la QoS.
- La mise en place de prépriorisation améliore le décroché pour les appels de priorité P1 au détriment des appels de priorité P3.
- Le triage des appels permet de réduire le temps d'attente maximal de 152 secondes pour les appels P1 (soit 97% de gains), 230 secondes pour les P2 (soit 95% de gains) et 239 secondes pour les P3 (soit 92% de gains).
- La prépriorisation permet de réduire le temps d'attente maximal de respectivement 134, 171 et 156 secondes (soit, 89, 97, et 85% de gains) pour les appels P1, P2, P3.
- Même lorsque le nombre d'appels double, la mise en place du triage permet de maintenir un QoS_{20} supérieur à 70% et un QoS_{60} supérieur à 95%.

Ainsi, la mise en place du triage rapide des appels dans les SAMU semble une bonne piste pour améliorer les performances de décroché des centres d'appels d'urgence. Toutefois, ces conclusions sont toujours à nuancer vis-à-vis des hypothèses sous-jacentes au modèle. Notre modèle pourrait être ajusté, notamment en améliorant sa capacité à prendre en compte les appels abandonnés. De plus, nous n'avons pas différencié les appels de la population et les appels bilans des équipes d'effecteurs. Ces derniers ne suivent pas tout à fait le même processus et les temps de sessions seraient légèrement différents par rapport aux appels de la population. Qui plus est, notre modèle pourrait gagner en précision concernant la définition des temps de sessions qui sont considérés comme indépendants du profil d'arrivée. Or, la tendance montre que plus il y a d'appels dans la file d'attente plus les temps de sessions des appels ont tendance à diminuer. Enfin, il faudrait aussi étudier plus en profondeur le lien entre les appels entrants et l'émission d'appels sortants pour identifier s'il existe une relation entre ces deux types d'appels.

Tableau 5.12 – Contributions et Limitation : Chapitre 5

Chapitre 5	
Contributions	Évaluation de la mise en place d'un scénario de triage et/ou de prépriorisation
	Évaluation de ces scénarios lors d'une augmentation soudaine du nombre d'appels
	Préconisations organisationnelles pour les SAMU
Limitations	Les appels abandonnés ne sont pas considérés
	Les appels bilans et les appels des autres praticiens ne sont pas différenciés des appels entrants
	Les temps de sessions sont considérés comme indépendants du profil d'arrivée
	Les émissions des appels sortants sont considérées comme indépendants des appels entrants