Projet ANR: MOSAICO
Multi-layer Orchestration for Secured and low lAtency applICatiOns
Projet No.: ANR-19-CE25-0012

Compte-Rendu Meeting #9
02-03/12/2021
Montimage - Paris

Participants:

- Orange : Bertrand Mathieu, Joël Ky, Stéphane Tuffin, Olivier Dugeon
- ICD-UTT : Guillaume Doyen, Marius Letourneau, Hichem Magnouche
- Montimage : Edgardo Montes De Oca, Huu Nghia Nguyen, Manh-Dung Nguyen
- CNRS-Loria : Thibault Cholez, Philippe Graff, Xavier Marchal

L'agenda proposé est validé.

Administratif/Dissémination :

- Site Web

L'UTT a mis à jour le site Web.

- Réunions

La prochaine réunion plénière en physique aura lieu à IMT Rennes (10-11/03/2022), sous organisation UTT. Guillaume doit vérifier la disponibilité d'une salle à ces dates.
PA Guillaume : Réserver la salle et préparer l'organisation de cette réunion

La réunion plénière suivante aura lieu les 23-24 juin à Orange à Lannion
PA Bertrand : Organiser cette réunion

La prochaine après-midi de réunion technique prévue le jeudi 09 décembre est annulée. La réunion de suivi de projet du 16 décembre est maintenue.

- Dissémination

Le papier sur les micro-services GPU a été soumis par le Loria à NOMS le 1er novembre. La notification est prévue pour le 17 décembre.

L'UTT va soumettre à NetSoft un papier sur le modèle d'orchestration de micro-services, deadline le 15 décembre.
PA UTT : Finaliser le papier

Orange va reprendre le papier sur le positionnement du projet pour soumission à IEEE Comm Mag, avec objectif d'avoir une version complète pour fin décembre.
PA Stéphane : Faire 1er draft complet du papier
PA Tous : Relire, compléter, adapter, etc.

Le Loria et Orange rédigent un papier à soumettre à Elsevier Computer Communications sur la détection du trafic cloud gaming. Thibault prend la main pour la finalisation et on reboucle ensuite ensemble. La deadline est le 20 décembre.
PA Loria : Finir rédaction du papier
PA Loria/Orange : Finaliser

Les papiers HiPNet sélectionnés pour soumettre une version étendue au journal Springer JNSM vont être notifiés. Les 2 papiers MOSAICO (papier L4S sécurité et analyse cloud gaming) ont été retenus. La deadline est fin février.
PA Auteurs : rédiger un nouveau papier étendu.

Montimage envisage de soumettre un papier 5GReplay à Esorics, si possible, sinon à une conférence/journal plus tard. Stéphane voit avec l'équipe sécurité à Orange Innovation si intérêt et voit si possible d'offrir un accès à Montimage à certaines PF d'Orange Open Source comme Magma pour pouvoir faire des tests.
PA Stéphane : Contacter les équipes Orange

Orange propose de rédiger un article pour publication sur le site Web Hello Future, un site public, à forte audience, présentant les innovations d'Orange.
PA Bertrand/Stéphane : Rédiger cet article

Concernant la normalisation, des échanges ont eu lieu sur une possible présentation au 3GPP ou à l'IRTF NMRG pour présenter les aspects

monitoring low-latency. Mais c'est encore tôt, car il faudra y aller avec des résultats concrets.

- Rapport intermédiaire T0+18

Il faut commencer à rédiger ce rapport.
PA Bertrand : Rédiger une 1ère version
PA Guillaume/Thibault/Edgardo : Vérifier la 1ère version et éventuellement compléter

- Plan Gestion de Données

Bertrand a mis à jour le document et envoyé à Guillaume, Thibault et Edgardo. Edgardo a fait une passe pour Montimage. Thibault est en train de regarder pour mettre le dataset Cloud Gaming en Open Data sur un serveur Loria. Suite à cela, il mettra à jour le PGD. Guillaume doit aussi vérifier et éventuellement compléter le PGD.
PA Guillaume/Thibault : Relire et éventuellement compléter avant fin le 15/12

- Divers / relations ANR

Edgardo demande si on peut commencer à facturer et envoyer la facture à l'ANR.
PA Bertrand : Se renseigner et informer Edgardo.

Livrable D234

Le livrable D234 est prévu à T0+24 (initialement T0+18, plus le retard de 6 mois du à la Covid-19). Ce livrable est donc attendu pour décembre.
Nous avons bien avancé dans sa rédaction. Il manque les derniers inputs, notamment sur la section 5 (model for orchestrating micro services (UTT), Orchestration of micro-services running on GPU (Loria), Interaction between orchestrators and detectors for a close-loop solution (Montimage, Orange). Ces sections seront remplies dans le mois. Ensuite, il faudra finaliser le document (cohérence, liaisons, intro/conclusion, relecture, etc.).
PA Thibault/Hichem/Huu-Nghia/Olivier : Remplir leur partie
PA Guillaume : Faire les intros, conclusions, liaisons
PA Tous : Relire

Présentation du modèle d'orchestration de micro-services : Hichem

Hichem présente son modèle, qui est dynamique, pour restructurer les SFC, suivant la possibilité de parallélisme, mutualisation, suppression des redondances, etc.
Il intègre la notion de bifurcation interne nœud ou externe (avec mémoire partagée si interne et copie de paquet si externe) et choix P4 ou VM. Le modèle

mathématique prend en compte la latence des arcs, la latence de traitement de chaque micro-service et la latence de bifurcation/fusion. Ceci a amené à discussion et nous avons conclu qu'il faudrait évaluer de manière concrète le temps de traitement des micro-services pour juger de l'intérêt de paralléliser les traitements avec copie/fusion des paquets. L'UTT doit regarder. Guillaume propose de faire, à une prochaine réunion visio, une présentation d'Hichem sur les papiers ayant inspirant ses travaux pour évaluer l'intérêt réel de la parallelisation de paquets.
PA Hichem : Identifier les temps de traitements des micro-services vs copie/fusion.
PA Guillaume/Hichem : Proposer une date pour une présentation d'Hichem

Il y a aussi eu des échanges sur le micro-service de type « copie », faisable en P4 avec insertion d'un identifiant de paquet avec une solution comme INT par ex. Par contre la « fusion » doit être faite en micro-service VM. En effet, ce n'est pas faisable en P4, car il n'y a pas de gestion de buffer de paquets et c'est problématique si les paquets n'arrivent pas en même temps. Il y a donc un cout à considérer pour le micro-service fusion.
Dans une première réflexion, on pense plutôt à avoir les micro-services colocalisés dans le même datacenter (pour remplacer une machine physique existante).

Dans le modèle, il faut aussi prendre en compte le nombre de flux parallèles que peut traiter le serveur (qui est une contrainte matérielle, par ex mémoire pour la gestion des tables de flux). Ceci peut avoir un impact sur le modèle d'optimisation.
PA Hichem : Voir comment intégrer cela à son modèle.

Nous avons discuté du nombre potentiel de SFC. Si dans le plan de contrôle, on peut imaginer de nombreux services, on peut se poser la question pour le plan de données. Quels micro-services on peut avoir dans le dataplane ? Pour l'instant, c'est assez limité en nombre.
Orange propose de réfléchir à des micro-services 5G qui seraient possibles pour réaliser des fonctions des stations de base, notamment quand elles sont distribuées selon l'architecture CU/DU/RU.
PA Joël/Stéphane/Bertrand : Réfléchir à micro-services 5G

Enfin, il faut concevoir l'algorithme d'optimisation selon 2 approches : 1) on essaie d'optimiser tous les services, dans le cas où les services sont déjà déployés et opérationnels, mais on pense qu'on gagnerait à optimiser. Même si le modèle est long à converger, ce n'est pas gênant puisque les services sont fonctionnels. Après résolution du modèle, on peut concevoir la reprogrammation des services. Dans ce cas, le modèle doit prendre en compte

l'optimisation de tous les services ; 2) On essaie d'optimiser le déploiement d'un nouveau service, en se basant sur l'état existant du modèle (qui a déjà déployé les services précédents). Dans ce cas, le modèle ne fait pas la résolution pour tous les services, mais uniquement pour le nouveau. Il doit être plus rapide pour permettre le déploiement rapide de ce service.
PA Hichem : Voir si sa solution de résolution du modèle actuelle permet cela ou s'il est nécessaire de l'adapter.

Présentation Monitoring & Closed-loop & Architecture: Hichem et Olivier

Huu-Nghia rappelle ses premières analyses sur le High-precision Monitoring notamment le in band telemetry et son application à notre cas.
Il y a eu de longs échanges sur le positionnement de ce monitoring, les problèmes de latence apparaissant surtout dans les réseaux d'accès (5G par ex) mais pas dans le cœur ou collecte. Guillaume fait remarquer que c'est différent de ce que le monde académique peut avoir en tête et que ce serait peut-être utile de faire un Position Paper pour bien expliquer que le problème de latence a lieu sur les réseaux d'accès et pas ailleurs, en expliquant et justifiant.
Peut-être qu'avec la 5G/6G et l'augmentation du débit, on pourrait éventuellement voir de la congestion ailleurs, par ex en préagrégation, mais cela demande à être confirmé.

Finalement, nous pouvons imaginer faire du high precision monitoring de bout-en-bout, avec une partie cœur/collecte en IP et une partie radio RU/CU/DU avec les mêmes concepts mais appliqués à l'environnement radio. Stéphane précise qu'Orange a fait un outil Open Source, LatSeq, qui est un framework de monitoring Haute précision dans la station de base, dont le projet pourrait s'inspirer. Orange (Joël) pourrait être intéressé pour étudier cette problématique. On pourrait utiliser les infos mesurées par LatSeq, à adapter par rapport à l'éclatement de la station de base et insérer du INT pour cela.
PA Joël/Stéphane/Bertrand : Etudier architecture 5G radio et réfléchir à adaptation possible de L4S

Nous pourrions faire des tests de bout-en-bout, avec par ex une partie « radio » OAI à Orange Lannion, une partie Cœur 5G à MMT et le serveur Cloud Gaming à Loria. Une autre option serait d'avoir plutôt OAI/Cœur au même endroit (MMT et Orange) pour efficacité. Il faut y réfléchir pour la suite des testbeds.
PA Orange/Montimage : Réfléchir à testbeds

De même, il faut réfléchir aux scénarios de tests. Stéphane propose d'en fournir un premier jeu, en utilisation du cloud gaming et l'UTT viendra ensuite ajouter du trafic perturbateur.

Olivier présente des résultats, retour d'expérience de tests réalisés par Orange sur les mécanismes closed-loop.
Il y a plusieurs options pour faire du closed-loop: telemetry, IGP, etc.
3 cas d'usage ont été étudiés :
1) 1er use-case : contrôle de la latence avec L3-VPN. Suite aux tests, pour être plus efficace et rapide dans la réaction, la recommandation est de faire un closed-loop au niveau contrôleur SDN pour les problèmes réseaux (50-100 ms) et au niveau orchestrateur pour les problèmes service (3-10 s).
2) 2eme use-case : contrôle de débit : Cela fonctionne bien pour optimiser la bande passante globale mais le temps de réaction est trop long pour faire le reroutage des chemins SR-TE et cela génère des pertes de paquets.
3) 3ème use-case : utiliser une métrique étendue IGP. Cela permet de réagir automatiquement. En cas de problème, la solution peut « flooder », annoncer tout de suite aux voisins que le délai dépasse le seuil (et peut-être aussi directement au contrôleur via PCEP) et donc on peut imaginer avoir une réaction rapide. La topologie réseau calculée par IGP pourrait être envoyée comme input réseau au modèle d'orchestration d'Hichem. Par contre, la télémétrie génère beaucoup de données donc il y a un compromis à faire.

Il y a eu une discussion sur le niveau de réaction, par rapport à la boucle MAPE (Monitor Analyze Plan and Execute). Pour optimiser la latence, c'est mieux si la réaction est prise au niveau du nœud (si le nœud a les infos nécessaires), sinon, il faut remonter au contrôleur pour les problèmes de réseaux si une décision globale est à faire, et enfin remonter à l'orchestrateur s'il y a des problèmes impactant l'organisation des micro-services.

Présentation des mécanismes de détection de flux malveillants : Marius

Marius rappelle le papier HiPNet sur les tests avec 3 types de flux indésirables. Mais ces comportements de flux (par ex micro-bursts) peuvent être légitimes. Dans ce cas, comment (et doit-on) détecter ces flux ? Quels critères ? Comment ?
Une première proposition présentée par Marius est de se placer sur le routeur et de gérer un vecteur de mesures. Il faudrait pourvoir constituer une base de flux malveillants pour pouvoir comparer le vecteur mesuré du flux avec celui de la base de flux malveillants.

Dans la solution courante, Marius fait la détection avec des stats (pas de l'IA, car trop complexe et problème de dataset complet), avec analyse fréquentielle (transformée de Fourier), et étude de la co-déviation des vecteurs.
Cela pourrait fonctionner dans les cas idéaux avec des mesures unitaires, mais quid en conditions avec plusieurs flux concurrents ? Avec conditions réseaux différentes ? Faudrait-il plusieurs vecteurs pour représenter un type de flux malveillant selon différentes configurations ?
Chaque flux a ses propres caractéristiques d'impact soit sur latence, soit sur le débit, soit sur l'IAT, etc… Faut-il caractériser les flux malveillants ou plutôt détecter les flux malveillants en fonction de leur impact sur les flux LL ? Mais attention aux flux générés artificiellement ? Et attention de faire les mesures au bon moment et pouvoir mesurer les autre flux ?
Ce que Marius propose actuellement permet de voir l'impact d'un flux malveillant sur 1 flux légitime, mais quid de la détection si 1 flux malveillant sur X (X) flux malveillants. Et quid de la possibilité de cumuler les flux malveillants pour faire une attaque efficace ?

Marius se pose actuellement aussi la question de l'évaluation du détecteur. Il faudra y revenir plus tard.

Dans l'immédiat, pour le papier JNSM, il est conclu de focaliser sur l'analyse et de faire plus de tests (attaque corrélée par ex, et cela reprend aussi les remarques de reviewers) et de mentionner les corrélations. L'objectif est d'enrichir le papier avec des tests. Nous décidons de ne pas encore mentionner les travaux sur le détecteur dans ce papier, car c'est trop tôt et il faut encore y réfléchir.
PA Marius : Faire de nouveaux tests, enrichir les résultats/analyses avec différentes configurations (plusieurs flux malveillants, corrélation attaques, etc.).

Présentation de la détection du trafic Cloud Gaming: Philippe

Philippe présente les derniers résultats de son étude sur l'analyse du comportement réseau du trafic Cloud Gaming versus d'autres types de trafic (vidéo streaming, live video et visio). Cela inclut notamment la détection en ligne en temps réel, réalisé à la volée vs depuis un fichier de captures pcap.
Philippe a eu des résultats significativement différents au début, car il utilisait 2 librairies différentes (scapy et pyshark) qui semblent donner des résultats différents. Il faut vérifier si ces outils offrent réellement des valeurs différentes ou s'ils calculent sur des informations différentes (par ex longueur du paquet UDP vs longueur du paquet Ethernet, etc.).

Si ensuite, les valeurs sont toujours différentes, il faut vérifier quel outil est correct. Pour cela, refaire des tests simples, avec peu de paquets et vérifier à la main la véracité des 2 outils. Ceci est important pour choisir le bon et être certain que les valeurs mesurées sont correctes.
PA Philippe : Identifier les raisons des différentes valeurs scapy et pyshark. Identifier lequel fournit les bonnes valeurs et baser son implémentation sur cette solution.

Philippe a réalisé une 1ère version du détecteur en python, mis dans un container pour pouvoir être déployée dans le niveau VM de MOSAICO.
On pourrait avoir des problèmes de performances et de traitement en temps réel si tous les flux sont remontés à la VM, mais normalement, seuls les nouveaux flux seront remontés pour être identifiés et ensuite le contrôleur enverra la configuration au nœud pour lui indiquer de rediriger les paquets de cette session dans la bonne file d'attente L4S, sans avoir à remonter les paquets de ces sessions au détecteur. Il n'y aura donc que peu de nouvelles sessions en parallèle à gérer.

Nous avons eu une discussion sur l'architecture MOSAICO pour savoir où on devrait mettre le détecteur CG (par ex, vers l'UPF). Mais il faut faire attention au tunnel et au NAT d'adresses IP sur l'UPF. Et comme le tunnel est sécurisé, on ne peut pas déchiffrer partout. Orange va proposer l'architecture à retenir.
PA Orange : Proposer architecture réseau MOSAICO avec emplacement des différents micro-services discutés.

Pour le trafic cloud gaming, L4S ne s'applique pas tel quel, car les PFs de CG ne font pas actuellement d'ECN (et délivré en WebRTC/UDP et pas TCP). Il nous faut donc réfléchir à une solution alternative, par ex avoir 2 files strictes et faire un couplage spécifique, « à la » L4S. Cela reviendrait à « réinventer » L4S pour les services ne supportant pas accurate ECN. Une 2ème étape pourrait être d'adapter les services pour qu'ils puissent être compatibles avec accurate ECN.
Le problème du cas d'usage de CG est qu'avec WebRTC (et retour RTCP toutes les 300ms), la réaction serait plus lente que ce que nous pourrions espérer avec Accurate ECN (et d'autant plus vrai avec la lente adaptation des codeurs vidéos). Une option pourrait être d'avoir cette solution pour une solution un peu plus lente, et d'utiliser Packet Wash pour une réaction rapide, en attendant l'autre adaptation.

Pour le papier JNSM, il est décidé d'étendre le papier HiPNet en faisant des tests pour les réseaux à capacités variables avec Mahimahi. Pour cela, Orange va refaire des captures réseaux pour avoir un dataset récent. Ces jeux seront

utilisés par le Loria pour faire les tests en situation d'environnement mobile. Il est prévu de faire une réunion dédiée Orange/Loria sur cet aspect.
PA Stéphane/Bertrand/Joël : Faire les captures réseaux
PA Loria : Faire les tests cloud gaming avec Mahimahi et fichiers fournis par Orange

Photos souvenirs :