le cnam Master Réseaux Informatiques d'Entreprise
RIE07 2012 – 2014

Grenoble INP FORMATION CONTINUE

Mémoire

présenté et soutenu le 4 juillet 2014 par
Cynthia Lopes do Sacramento

SDN et le passage au Cloud Computing

Un regard sur les aspects réseaux et les apports de SDN dans les data centres

---

Jury :	Romain Kobylanski
	François Miller
	Véronique Panne
Tuteur académique :	Jean-Luc Parouty
Tuteur :	Claude Casery
Entreprise :	Bull

Remerciements

Le présent document peut être considéré comme le résumé de mon cycle d'études « Master Réseaux Informatique d'Entreprise ».

Je tiens à remercier tous ceux qui, à l'Institut National Polytechnique de Grenoble et à l'entreprise Bull, m'ont accompagnée dans ce parcours.

À tous les intervenants pédagogiques de la formation RIE, j'adresse mes remerciements pour les connaissances transmises, spécialement à Messieurs Kobylanski, Miller et Parouty ainsi qu'à Madame Panne, pour leurs conseils et leur accompagnement.

J'exprime toute ma gratitude à mes collègues de Bull pour leur accueil et leur sens de la convivialité pendant ces deux années d'alternance. Toute ma reconnaissance va spécialement aux membres de l'équipe OSI pour m'avoir intégrée : tout d'abord à Claude Casery, mon tuteur, pour son rôle clé dans mon développement professionnel, mais aussi à Julien et à Laure pour leur aide technique et leur disponibilité.

Merci également aux membres de l'équipe OMNIS, de m'avoir précédemment accueillie lors de mon stage de licence, et d'être devenus rapidement mes grands amis de France. Leur façon d'être, leurs encouragements (particulièrement de la part de Martine) et leur drôlerie m'ont beaucoup inspirée et ont contribué à l'aboutissement de mon expérience à l'étranger.

Dans l'équipe Storeway, je suis particulièrement reconnaissante envers Emmanuel de m'avoir initiée au sujet SDN, thème choisi pour ce mémoire, ainsi qu'à Renata, ma compatriote, pour sa compréhension et son aide pour déterminer l'équivalence d'un vocabulaire particulier.

J'exprime ma gratitude aussi à tous ceux qui ont relu mon document et m'ont permis de l'améliorer, et particulièrement à Christine pour ses explications, corrections et sa patience.

Enfin, je remercie ma famille et mes amis qui m'ont soutenue même de loin, ainsi que tous les nouveaux amis rencontrés localement pour lesquels j'ai développé une forte affection depuis mon arrivée. Tous ont contribué d'une manière ou d'une autre à la réussite de mon parcours.

Table des matières

Table des figures

Introduction

Les centres de traitement de données évoluent aujourd'hui à un rythme intense pour accompagner l'explosion constatée dans l'utilisation (en volume et en diversité) de données. L'accélération de l'innovation dans l'informatique impose une rénovation constante des infrastructures des entreprises. La virtualisation a permis aux centres de données d'améliorer la productivité de leurs serveurs, mais pour arriver à l'agilité souhaitée, les data centres doivent faire évoluer leurs réseaux. Cette étude analyse les applications **Software-Defined Networking : Réseau Informatique Défini par Logiciel (SDN)** pour distinguer quels sont les apports dans le contexte actuel et futur des data centres et habiliter le passage au **cloud computing**.

La plupart des infrastructures de **Technologie de l'Information (TI)** n'ont pas été construites pour supporter la croissance explosive des données constatée aujourd'hui et n'ont pas la capacité de traitement de l'information demandée. Plusieurs centres de traitement de données sont devenus hautement distribués et relativement fragmentés par rapport aux besoins des différents profils des clients. Ils se trouvent donc limités dans leur capacité à évoluer rapidement et à supporter l'intégration des nouveaux types de technologies ou à se mettre à l'échelle des besoins de ses utilisateurs.

Lorsqu'ils sont équipés d'infrastructures performantes, partagées et dynamiques ainsi que des outils nécessaires pour adapter les ressources à la demande, les **Système d'Information (SI)** peuvent répondre efficacement aux besoins métiers. Ainsi, les structures pourraient se focaliser sur l'innovation et l'adaptation des ressources selon les priorités stratégiques de leurs métiers. Cela faciliterait la prise de décisions, qui pourrait être concentrée sur l'information en temps-réel. [1]

Alors que le coût du réseau dans un data centre est estimé à 15% du total, sans être un des plus élevés, il est largement établi qu'il représente un élément clé pour la réduction des coûts et l'augmentation du retour sur investissement. Les coûts d'investissement dans les serveurs ont été évalués à 45% des coûts des data centres. Malheureusement la charge

utile des serveurs est remarquablement basse, arrivant à seulement 10% d'utilisation dans certains exemples. [2]

La technique de la virtualisation a permis le partage des processus entre machines, mais des contraintes réseau continuent à limiter l'agilité dans les data centres. L'agilité est définie par la capacité de placer tout service n'importe où dans le data centre, tout en assurant la sécurité, la performance et l'isolation entre tous les services. Les designs des réseaux conventionnels dans un data centre empêchent cette agilité ; par nature ils fragmentent à la fois les réseaux et la capacité des serveurs, limitant et réduisant la croissance dynamique des pools de serveur et de traitement de l'information. [3]

L'agilité est donc un élément clé ; certaines entreprises s'évertuent à déployer de nouvelles applications ou faire évoluer les existantes au rythme de la croissance de leur business. Selon le sondage mené par AlgoSec avec 240 professionnels de l'informatique, 25% des organisations participantes doivent attendre plus de 11 semaines pour qu'une nouvelle application soit mise en ligne (et dans 14%, ce temps dépasse 5 mois). Les résultats révèlent également que 59% des entreprises ont besoin de plus de huit heures pour réaliser un changement de connectivité dans une application. [4]

Cependant, lors du passage au Cloud, les entreprises réalisent que la virtualisation des serveurs est considérablement limitée par les designs Ethernet classiques et les contrôles de sécurité réseau traditionnels. Avec l'augmentation de la virtualisation au sein des data centres, quatre tâches majeures deviennent critiques :

— Prévention de la congestion du trafic ;
— Réduction de la complexité mise en place des politiques réseau et maintien du niveau de service ;
— Élimination des points aveugles qui conduisent à des pannes ;
— Scellement des failles de sécurité pour protéger les données. [5]

Cette étude a pour but de démontrer comment SDN peut être appliqué aux réseaux pour permettre aux data centres de passer au Cloud Computing, et dépasser les limites du réseau actuel. Dans le premier chapitre, le contexte des data centres sera défini. Ensuite, les problématiques dans l'aspect réseau seront exposées. Enfin, le dernier chapitre présentera SDN et démontrera ses apports dans ce cadre.

Chapitre 1

Les évolutions stratégiques des Data Centres

Ce chapitre a pour but de définir un data centre et d'en analyser les problématiques, enjeux et solutions possibles, en vue de comprendre sa situation actuelle ainsi que ses limites par rapport aux nouveaux besoins et défis business. Les éléments les plus importants de la conception et de l'architecture du data centre seront présentés ainsi que les difficultés qui ont amené à faire évoluer le mode de livraison de services informatiques vers une approche Cloud Computing.

1.1 Data centres et leurs objectifs

Un data centre (aussi nommé « ferme de serveurs ») est un répertoire centralisé pour le stockage, le management et la distribution de données et d'informations. Typiquement, un data centre est une installation physique utilisée pour loger des systèmes informatiques et ses composants associés, tels que systèmes de télécommunication et stockage.

Les data centres traditionnels hébergent historiquement de nombreuses applications relativement peu sollicitées, chacune s'exécutant dans une infrastructure matérielle dédiée qui est isolée et protégée des autres systèmes dans la même installation physique. Ces data centres accueillent du matériel et du logiciel pour multiples unités organisationnelles ou même diverses entreprises. Différents systèmes informatiques au sein d'un tel data centre ont souvent très peu d'éléments en commun en termes de matériel, logiciel ou infrastructure de maintenance, et en général ne communiquent pas entre eux.

Les tendances de l'informatique vers une approche "serveur" et l'explosion en popularité des services sur internet ont changé ce scénario. Des infrastructures data centre entières ont été dédiées à un seul acteur pour faire fonctionner ses services offerts. Dans ce cadre, un data centre appartient à une seule organisation et utilise des matériels et plateformes logicielles relativement homogènes qui partagent une couche commune de systèmes de management. Ces data centres dédiés exécutent un nombre réduit d'applications (ou services internet) beaucoup plus sollicitées ; l'infrastructure commune de management permet alors une meilleure flexibilité de déploiement.

Ces infrastructures sont montées pour absorber la croissance des applications déployées et la haute disponibilité exigée pour ces services, visant en général 99,99% (une heure au maximum de temps d'arrêt par an). Il est difficile d'atteindre un fonctionnement libre des failles dans un regroupement de systèmes matériel et cela devient encore plus complexe avec le grand nombre de serveurs impliqués.

Les infrastructures de ces data centres doivent être dimensionnées précisément en fonction de la charge des applications supportées. Par conséquent, de nouvelles approches ont été proposées pour la construction et l'opération de ces systèmes qui doivent être conçus pour tolérer un nombre important de failles avec très peu ou aucun impact sur la performance et disponibilité des services offerts. [6] [7]

1.2 Organisation d'un data centre et difficultés

Un data centre est en général organisé en rangées de racks où chaque rack contient des dispositifs modulaires tels que serveurs, switches, baies de stockages ou équipements spécialisés. Les composants essentiels de l'infrastructure, branchés aux racks des data centres d'entreprises, sont classés en tant que compute, stockage et réseau et représentent la base sur laquelle les applications business sont construites.

Un châssis se présente avec ses propres ventilateurs, source d'alimentation, fond de panier et module de gestion. Pour réduire l'espace occupé, des serveurs peuvent être compartimentés dans un châssis qui est glissé dans le rack. Un châssis fournit des slots où il est possible d'insérer des élément actifs modulaires (aussi connus en tant que « blades »). Par exemple, un seul châssis peut contenir 16 serveurs 1 U ; étant donné que les racks supportent 6 châssis, ils ont donc une capacité théorique de 96 éléments modulaires.

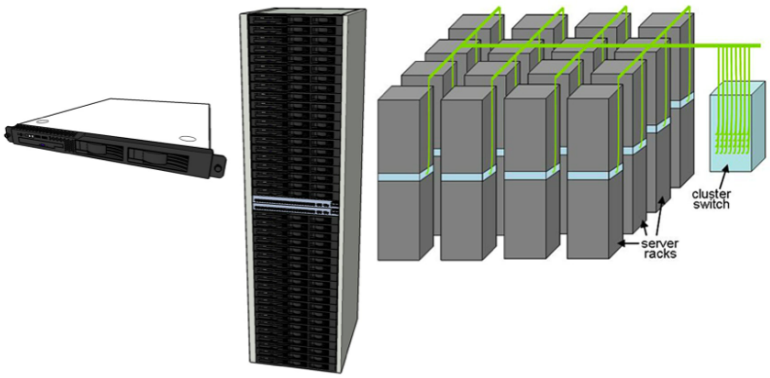

FIGURE 1.1 – Organisation de racks. [7]

La figure 1.1 montre l'organisation des racks dans un data centre. Un serveur occupant 1 U du rack est visualisé à gauche. Au centre on peut voir un rack et à droite un cluster de racks avec un swtich/routeur de niveau cluster. En général un ensemble de serveurs 1U est monté dans un rack et inter-connecté à un commutateur Ethernet local. Ces switches au niveau des racks, qui peuvent utiliser des liens de 1 à 10 Gbps, ont un nombre de connexions montantes vers un ou plus switches de niveau cluster (data centre).

Le stockage dans les data centres peut être proposé de diverses manières. Souvent le stockage de haute performance est logé dans des baies de stockage qui permettent un accès distant transparent aux données, indépendamment du nombre et des types des dispositifs de stockage physiques installés. Le stockage peut également être fourni dans une « brique de stockage » plus petite, localisée dans le rack ou slot de châssis ainsi que directement intégrée aux serveurs. Dans tous les cas, un accès réseau efficace au stockage est crucial.

Le problème le plus important dans cette structure est l'éventuelle insuffisance de bande passante. En général, les connexions montantes sont conçues pour supporter un certain taux de demandes excédentaires puisque la fourniture d'une bande passante entière n'est pas toujours possible. Par exemple, pour 20 serveurs à 1Gbps, chacun doit partager un lien Ethernet montant unique de 10Gps à un taux de demande excédentaire de 2. Cette situation peut être problématique si la charge réseau non locale au rack monte considérablement. Comme le stockage est traditionnellement fourni dans une baie séparée, tout le trafic de stockage traverse le lien montant dans le réseau stockage. Par exemple, l'archivage d'un gros volume peut consommer une importante bande passante. À mesure que les data centres augmentent en taille, une architecture réseau plus extensible devient essentielle.

La consommation d'énergie et les coûts liés sont devenus également une des préoccupations de la conception des data centres, car ils représentent une part importante de la totalité des dépenses. Actuellement les CPUs ne sont plus le seul élément cible d'amélioration de l'efficacité énergétique, car ils ne dominent plus la majorité de la consommation. Des problématiques de refroidissement et consommation d'énergie sont des facteurs de plus en plus critiques dans la conception de data centres.[7] [8]

1.3 Virtualisation et partage de ressources

Le besoin d'augmenter l'efficacité dans l'utilisation des ressources a conduit à une conception d'infrastructures avec partage et virtualisation des composants. La virtualisation fait référence à l'abstraction des ressources logiques de leurs couches physiques pour améliorer l'agilité, la flexibilité et la réduction des coûts et ainsi privilégier les activités du business. La virtualisation permet de consolider un ensemble de composants d'infrastructures sous-utilisés en un nombre de dispositifs plus petits et mieux utilisés, contribuant à l'économie des coûts et à l'efficacité de l'utilisation des composants.

La virtualisation de serveurs est une méthode pour abstraire le système d'exploitation de la plateforme matérielle. Cela permet aux multiples systèmes d'exploitation ou multiples instances du même système d'exploitation de coexister dans un ou plusieurs clusters de processeurs. L'image 1.2 illustre le partage de ressources par l'intermédiaire de la virtualisation.

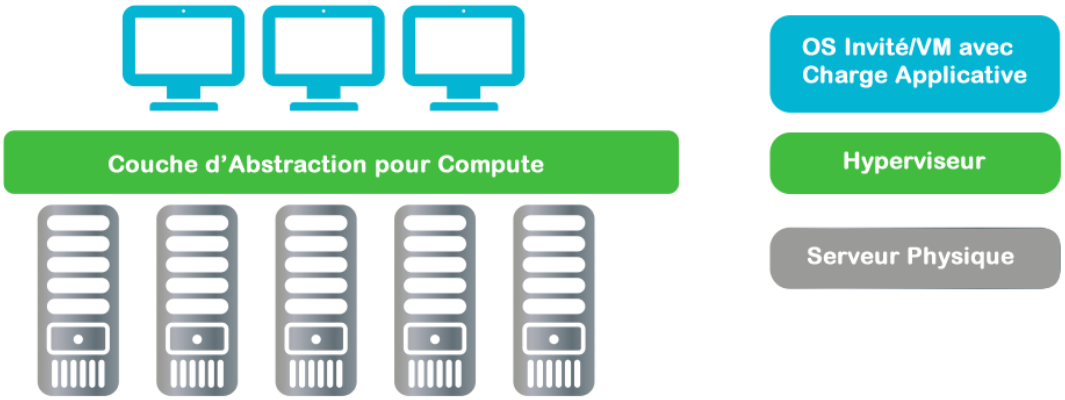

FIGURE 1.2 – Modèle d'infrastructure à ressources partagées. [9]

Un hyperviseur est inséré entre le système d'exploitation et le matériel pour réaliser la séparation entre le logique et le physique. Les instances de systèmes d'exploitation lancées sont appelées invités, ou systèmes d'exploitation invités. Selon la technologie utilisée, l'hyperviseur peut fournir l'émulation matérielle aux systèmes invités et gère l'allocation de ressources matérielles. Les principaux hyperviseurs disponibles sur le marché aujourd'hui sont : VMware ESXi, KVM basé sur Linux et supporté par Red Hat, Citrix XEN et Microsoft Hyper-V.

Ce modèle apporte des avantages pour l'efficacité dans l'utilisation de ressources avec des charges applicatives idéales. Cependant, quand une application commence à consommer plus de ressources que l'estimé, il peut arriver des scénarios où les systèmes d'exploitation invités n'ont pas assez de ressources, impactant ainsi la qualité du service dispensé.

Cette approche a apporté une maîtrise globale de gestion, supervision et outillage. Elle a aussi mis en évidence que le composant « **compute** » de l'infrastructure améliore nettement l'utilisation et l'automatisation d'allocation des ressources serveurs. Cette amélioration a été possible grâce à la programmation du contrôle de ressources fournies aux instances invitées. Toutefois, le développement de nouvelles solutions pour gérer la charge dynamique de certaines applications faisait toujours défaut. [9] [10]

1.4 Le besoin d'un modèle plus dynamique

Traditionnellement, les data centres d'entreprises sont conçus pour perdurer et atteindre les objectives déterminés. Cela veut dire que les éléments sous-jacents sont dimensionnés et construits pour supporter le pic de charge prévu en termes de performance, disponibilité et sécurité. Quand la croissance volumétrique escomptée ne correspond pas à la réalité, cette méthode de dimensionnement peut conduire à une situation de sous-dimensionnement ou sur-dimensionnement. Ce qui produit un effet négatif pour les investissements et les efforts de réduction de coûts.

En général, pour atteindre une meilleure disponibilité, les infrastructures sont amenées à une sous-utilisation des ressources. Comme la charge des applications varient continuellement dans les applications sur internet, il reste deux choix : soit sous-dimensionner la provision des ressources et perdre des clients ou alors sur-dimensionner et gaspiller les ressources.

Dans tous les cas, un plan détaillé de capacité est fait pour spécifier une série d'investissements importants en matériel et logiciel, dont la charge maximale est déterminée. L'image suivante illustre cette planification et les situations de problèmes de dimensionnement.

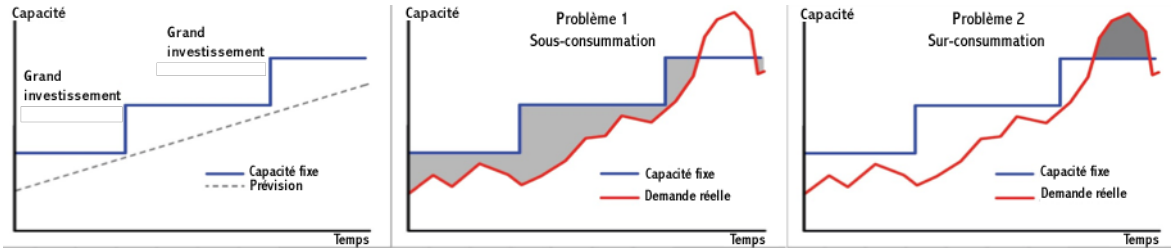

FIGURE 1.3 – Capacité fixe de ressources vs charge prévisionnelle. [11]

Face à cette problématique, un nouveau mode de livraison de service a été proposé pour aborder les défis du traitement des demandes pour la variation dynamique des charges applicatives. Avec la nouvelle tendance du Cloud Computing et l'**Infrastructure as a Service, Infrastructure en tant que service (IaaS)**, la conception de **clusters** hautement disponibles et leurs solutions extensibles peuvent être architecturées avec des requis non-fonctionnels comme base.

Avec sa nature adaptable, le modèle de livraison cloud permet aux ressources d'être étendues et réduites dynamiquement en fonction de la consommation. Une couche logicielle d'abstraction, implémentée par les hyperviseurs, virtualise le traitement des ressources physiques, permettant ainsi au processeur, à la mémoire et aux disques durs de s'adapter aux variations des demandes. Le réseau physique est également partagé mais n'a pas d'implémentation permettant l'adaptabilité de ces ressources. [9] [11]

1.5 Cloud Computing

Le Cloud Computing peut être défini comme un nouveau modèle de consommation et livraison de ressources et services de **TI**, et a pour principales caractéristiques :

— Libre service à la demande ;
— Mobilité d'accès au service réseau ;
— Location indépendante de services dans un pool de ressources commun ;

— Sécurité des ressources, administration automatisée et réglementation de l'utilisation de données ;
— Extensibilité et approvisionnement rapides ;
— Paiement à la consommation.

Les avancées importantes dans les domaines de virtualisation, réseau, approvisionnement et architectures multi-tenantes ont permis de faire évoluer radicalement les infrastructures de data centres. Le plus grand impact du Cloud Computing vient de l'instauration de nouveaux modèles de consommation et de livraison de services qui encouragent l'innovation.

L'évolution des data centres a permis de satisfaire une plus grande variété de besoins dans le monde du travail, ce qui implique la prise en compte de plusieurs facteurs lors de la conception de l'architecture selon les différents objectifs de chaque entreprise. Le Cloud Computing est donc le produit d'un nouveau paradigme pour les architectures des data centres.

Le Cloud Computing livre dynamiquement des services sur des réseaux à partir d'un ensemble abstrait de ressources. Ces ressources se retrouvent quelque part dans le « nuage » [1] et sont disponibles immédiatement à la demande. Les types de ressources ainsi que leur localisation sont transparents aux utilisateurs finaux. Ces utilisateurs ont pour principale préoccupation que leurs applications, données et contenus soient sécurisés et disponibles, avec un certain niveau de qualité de service.

Du point de vue de l'infrastructure, le Cloud Computing sollicite fortement les ressources mutualisées dans une grande variété de technologies (compute, stockage, réseau) pour leur allocation dynamique. Tout ceci dans un environnement automatisé, orchestré et logiquement diversifié, en conciliant une variété d'applications. L'orchestration permet de mutualiser les ressources au travers de multiples data centres pour une réponse dynamique aux besoins clients.

La virtualisation de serveurs a représenté un premier et important pas pour la viabilité de l'approche Cloud Computing. Toutefois, les deux autres éléments de base de l'infrastructure des data centres doivent accompagner ces changements pour permettre un accès complet aux services offerts par le Cloud. Les hyperviseurs ont permis la séparation des systèmes logiques des serveurs physiques dans le cas de la virtualisation de serveurs ; cette abstraction doit être également appliquée aux matériels réseaux et

1. symbole qui fait allusion à la représentation d'internet dans les topologies réseau

de stockage. Cela permettra la définition d'un data centre entièrement piloté par un logiciel qui gère les ressources physiques, en les activant dynamiquement selon la charge applicative.

L'image suivante illustre une vue conceptuelle d'un data centre basé sur ces trois éléments avec des couches d'abstraction permettant de sécuriser, orchestrer et livrer ses ressources au consommateur.

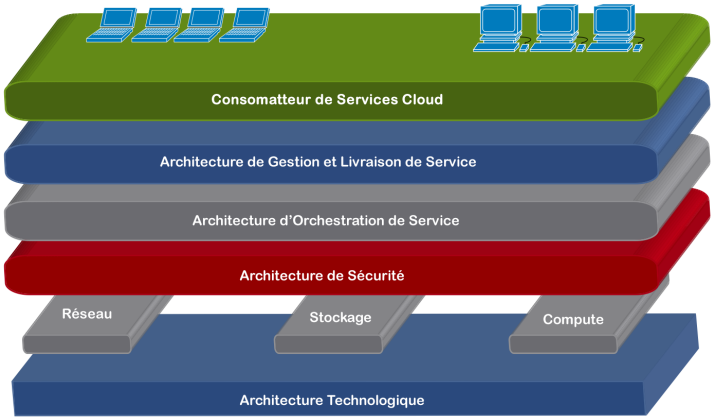

FIGURE 1.4 – Vue conceptuelle d'un data centre. [12]

L'abstraction de ces trois composants matériels est essentielle pour permettre la distribution de services Cloud au sein des data centres. L'adoption de la virtualisation des serveurs a déjà atteint son public ; en 2009 un sondage avait révélé que 77% des entreprises consultées déployaient au moins un système virtualisé dans leur data centre [13]. On observe que beaucoup de travaux se déroulent actuellement en recherche et développement pour acquérir un niveau équivalent de maturité pour les dispositifs réseau et stockage.

L'abstraction du stockage signifie la capacité à mutualiser les dispositifs physiques de stockage pour pouvoir les utiliser en tant que volumes de stockage logiques. Cela caractérise la virtualisation du stockage ou le **Software Defined-Storage, Stockage Défini par Logiciel (SDS)**. Pour l'aspect stockage, il est reconnu que des solutions **SDS** se trouvent disponibles sur le marché, telles que EMC ViPR, HP StoreVirtual, IBM SmartCloud Virtual Storage Center, Red Hat Storage Server entre autres.

Pour les réseaux, il se développe également une technologie fournissant une couche d'abstraction pour divers dispositifs réseau afin de permettre l'isolation logique et l'in-

dépendance du matériel. Il se trouve que **SDN** est une des approches proposées pour traiter la problématique de l'abstraction réseau et fait donc l'objet de cette étude. Le chapitre suivant démontrera en quoi les réseaux traditionnels ne sont pas adaptés aux exigences du Cloud Computing et analysera des exemples sur divers problèmes rencontrés. L'approche SDN et ses apports seront détaillés en dernière partie. [9] [12] [14] [15]

Chapitre 2

L'aspect bloquant du réseau

Dans ce chapitre, les principales problématiques data centre sous un aspect réseau seront présentées et analysées. Seront abordés quelques questionnements sur les réseaux classiques pour tenter de répondre aux nouveaux besoins data centres. Le chapitre présentera comment divers scénarios sont traités aujourd'hui et quelles en sont les limites.

2.1 Le rôle du réseau dans les projets de TI

Lors du développement de projets pour l'optimisation en **TI** telles que la consolidation de data centres et la virtualisation de serveurs, une attention spéciale doit être accordée au rôle critique des réseaux dans la planification, l'exécution et succès en général du projet. Il est souvent admis que des planifications supplémentaires concernant spécifiquement les réseaux auraient pu contribuer au succès de plusieurs projets.

Les principaux types de modifications dans ces projets incluent l'implémentation d'équipements réseaux supplémentaires pour augmenter ou améliorer la redondance, la capacité du flux, la sécurité et/ou la bande passante. Cependant, plusieurs pré-requis associés à ces changements ne sont pas, en général, identifiés au tout début des projets. Très souvent ils ne sont détectés qu'après les étapes initiales du projet, imposant un supplément de travail et l'ajout de coûts non anticipés.

Les aspects réseau d'un projet peuvent être difficiles à gérer, et des critiques sur le fonctionnement général sont fréquemment entendues. Pour des défis importants, il est nécessaire d'effectuer des analyses des causes précises et opportunes, de comprendre la

réactivité au niveau applicatif et de révéler les origines des problèmes de performance. Le simple achat d'équipement réseau ne permet pas nécessairement de répondre de manière satisfaisante et efficace aux besoins réels.

Pour prendre en charge la virtualisation complète d'une infrastructure de **TI** et continuer à optimiser le réseau, des décisions sur l'architecture doivent être prises dans le contexte de l'infrastructure existante, de la stratégie et des principaux objectifs du business. Sans le développement d'un plan réseau et la conception fonctionnelle associée, les transitions réseau peuvent être risquées et conduire à un contrôle réduit des services délivrés, des coûts potentiellement élevés, des résultats insuffisants et des problèmes inattendus en termes de performance ou disponibilité.

Traditionnellement un plan et une conception fonctionnelle solides suffisaient pour assurer le succès des projets avec un réseau réactif, optimisé, moins cher et répondant mieux aux engagements des services applicatifs. Alors que ce plan reste essentiel, il est difficile dans le contexte d'utilisation actuel de maîtriser complètement à l'avance la charge, dimension et tout autre pré-requis devant être assurés par les réseaux.

Face à ces difficultés et aux critiques reçues, il est facile de considérer les réseaux comme un élément bloquant pour le succès des projets. Même si la gestion et la planification peuvent être complexes, les réseaux représentent un moyen naturel pour gérer et renforcer les politiques liées aux risques, performance et coûts. Cependant, seul le réseau voit toutes les données, ressources connectées et interactions des utilisateurs à travers le Cloud. Le réseau est donc positionné de manière unique pour surveiller et mesurer l'usage et la performance des services distribués et de l'infrastructure. Les réseaux ont également un rôle central pour favoriser l'extensibilité et la disponibilité. Par exemple avec leur vue unique de bout-en-bout, les réseaux peuvent détecter la charge et la rediriger automatiquement selon les politiques de contrôle pré-définies.

Afin d'atteindre ce niveau de management et orchestration des ressources, on cherche à concevoir des architectures réseaux qui puissent s'étendre ou se rétracter ainsi que supporter des nouveaux services de façon dynamique et rapide en fonction des besoins immédiats. [16] [17] [18]

2.2 L'architecture réseau d'un data centre typique

La majorité des data centres d'aujourd'hui ont une structure réseau hiérarchique à trois niveaux : couche d'accès, agrégation/distribution et cœur (figure 2.1). La couche d'accès inter-connecte toutes les ressources partagées telles que serveurs, dispositifs de stockage et applications. Les réseaux d'agrégation (ou distribution) doivent fournir une haute bande passante pour la communication entre multiples réseaux d'accès. Le cœur du réseau est l'interface vers l'extérieur du réseau, qui peut être le lien avec des réseaux WAN, mobiles, VPNs ou autres types d'accès internet.

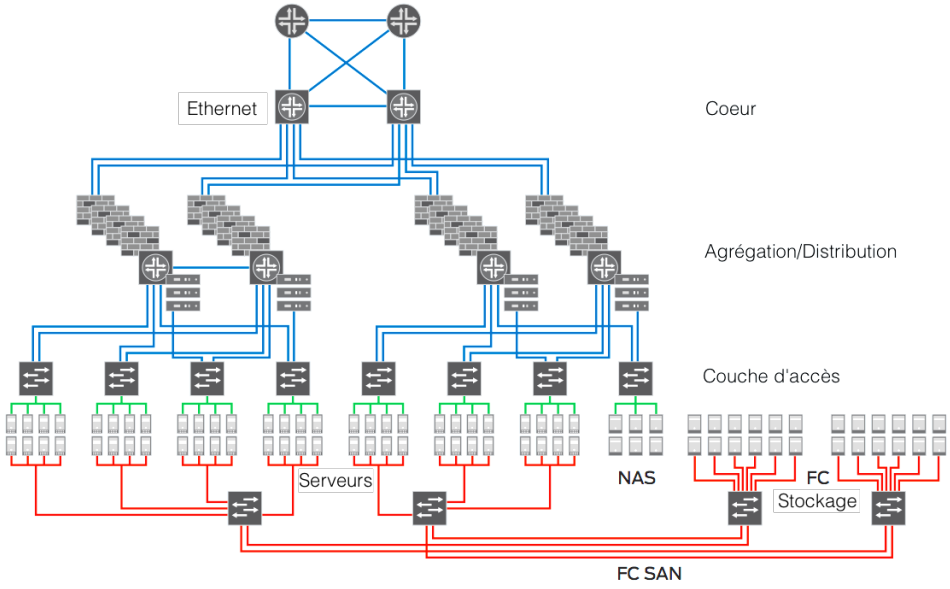

FIGURE 2.1 – Architecture réseau typique en trois tiers. [19]

Cette architecture a été conçue pour les applications client-serveurs dans des serveurs applicatifs dédiés. Dans cette conception, le flux du trafic se fait essentiellement à partir des serveurs en direction du cœur du réseau et ensuite en dehors vers internet ou autre (flux Nord/Sud).

Ce modèle d'architecture a été la principale référence de topologie réseau pour les data centres. Cependant, il devient complexe à le maintenir dans le contexte du Cloud Computing lorsque les applications actuelles commencent à être plus distribuées, avec plusieurs couches et orientées livraison de services. Ces changements dans les applications ont impacté les réseaux pour ce qui concerne le volume et les flux du trafic.

Le trafic réseau interne a augmenté avec la croissance du nombre d'applications et services déployés ; l'utilisation de systèmes de fichiers distribués avec des données stockées séparément a eu pour conséquence une charge réseau plus importante à travers le data centre. De ce fait, on constate une forte tendance de communications inter-serveurs ou inter-VMs, impliquant un flux Est/Oust plutôt que Nord/Sud. L'industrie estime même que 80% du trafic des applications Cloud Computing constitue des flux Est/Ouest.

Cette architecture révèle une multiples connexions des serveurs à divers types de réseaux tels que réseaux locaux, de stockage, communication entre processus. Cette disposition ajoute de la complexité et des coûts sous forme de câblage, nombre de ports, extensibilité, énergie, refroidissement etc.

Les data centres d'aujourd'hui exigent des réseaux agiles et flexibles pour pouvoir réagir rapidement aux changements et garantir une livraison efficace de services. Par exemple, des serveurs virtuels peuvent être déplacés de part et d'autre avec une simple commande en fonction de la demande. Par conséquent, les réseaux doivent s'adapter rapidement à ces changements pour éviter des perturbations du service. Cela suppose que les dispositifs apparaissent connectés au même réseau local, indépendamment de leur proximité physique. Il manque à cette architecture la flexibilité nécessaire pour effectuer ces types de changements, ce qui impacte défavorablement l'efficacité opérationnelle de tout le data centre.

Afin de mieux visualiser ces difficultés, quelques scénarios critiques pour les réseaux traditionnels seront étudiés en détail. Ces scénarios supposent des infrastructures basées sur cette architecture à trois niveaux (présentée précédemment) avec l'objectif de fonctionner en mode Cloud Computing. [19] [20] [21] [22]

2.3 La transformation des applications et exemples de scénarios critiques

Avec de multiples **Virtual Machine, Machine Virtuelles (VMs)** s'exécutant sur le même hôte, partageant une carte réseau unique au moyen d'un switch logiciel (ou virtuel), les applications ont donc moins de ressources réseau que lorsqu'un serveur physique leur est dédié. Cela peut conduire à des problèmes de performance réseau comme une bande passante réduite et un temps de latence augmenté et même à d'autres

difficultés comme la disponibilité de plages d'adresses IP ; les applications peuvent ne pas être en mesure de traiter ces questions dans le cas de ces restrictions.

Les plateformes typiques d'infrastructure virtuelle fournissent des logiciels pour faire migrer les instances de VMs actives d'un dispositif physique à l'autre ; VMware Distributed Resource Scheduler (DRS) avec VMotion et Proxmox VE Live Migration (**open source**) sont des exemples de ce type de solutions. Ces solutions ne connaissent pas l'état des applications ou du réseau. Par exemple, une VM peut être migrée au milieu d'une transaction bancaire sans prendre en compte l'état de persistance de l'opération, le nombre des connexions ou la charge réseau que la VM est en train de traiter. Cela peut causer des transactions non réussies, avec perte de données et problèmes plus graves pour les utilisateurs. Par ailleurs, ces technologies permettent de migrer une VM vers un serveur physique avec plus de cycles de CPU disponibles, mais elles n'ont pas d'informations sur la capacité réseau dans ce serveur.

Le déplacement dynamique de charges exige également que les VMs restent dans un VLAN commun, dans le même réseau au niveau 2. Pour pouvoir déplacer un réseau en dehors de son domaine niveau 2, il est nécessaire d'utiliser des procédures manuelles comme l'attribution d'adresses IP et mise à jour des entrées DNS pour les services déplacés. Pour maximiser cette flexibilité, des technologies émergent pour élargir le domaine des réseaux au niveau 2.

De nouveaux moyens qui utilisent l'encapsulation de tunnels tels que **Virtual eXtensible LAN (VXLAN)** et **Network Virtualization using Generic Routing Encapsulation (NVGRE)** étendent les réseaux couche 2 avec des réseaux couche 3. Même si cette capacité devient possible avec ces technologies, le trafic local aura toujours une meilleure performance et une latence inférieure s'ils restent dans un réseau niveau 2. [20] [23] [24]

2.3.1 Aspect multi-tenant

Cet exemple traite l'aspect multi-tenant des infrastructures Cloud. Pour simplifier, on considère une configuration avec deux réseaux : un pour le groupe d'ingénieurs et l'autre pour l'équipe des ventes. Pour réaliser cela avec les technologies réseaux traditionnelles, on utilise le concept de **Virtual Local Area Network, Virtual LAN (VLAN)**. Par exemple, le réseau d'ingénieurs peut être affecté au VLAN-1 et le réseau des ventes au

VLAN-2 comme illustré dans l'image 2.2.

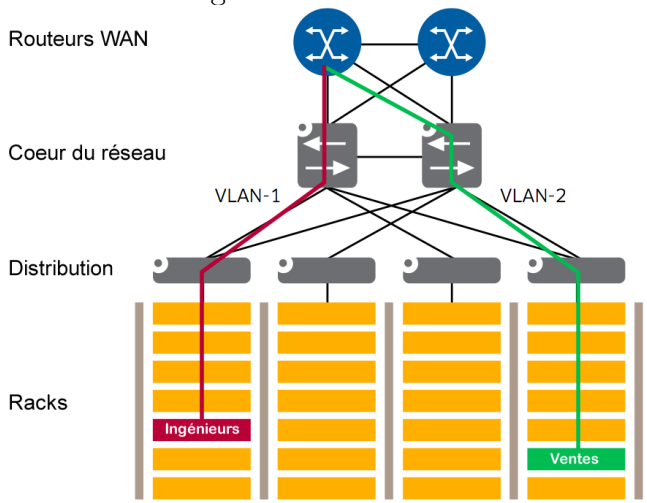

FIGURE 2.2 – Architecture réseau avec deux tenants. [25]

Cette attribution doit être effectuée par le système de gestion des switches virtuels dans l'hyperviseur. Ensuite, la configuration doit être réalisée dans tous les switches du réseau data centre et dans le switch physique auquel le routeur est connecté. Si le numéro identifiant du VLAN est déjà attribué, l'interconnexion n'est pas possible, cela veut dire qu'une administration continue des VLANs attribués doit être mise en place. Tout cela est fait de manière fastidieuse et manuelle par l'opérateur du réseau.

Chaque réseau est connecté à un routeur, qui doit être configuré pour prendre en compte ces VLANs. Le trafic doit toujours monter et descendre passant par ce routeur (réseau d'accès et de distribution), ce qui n'est pas très performant. La capacité du routeur peut limiter les communication inter-départements.

Pour fournir des adresses IP automatiquement aux applications clients, un serveur DHCP doit être attribué à chaque réseau. L'équipe d'opération réseau met en place un serveur DHCP pour chaque réseau en accord avec les configurations dans les routeurs.

Le routeur doit implémenter les règles de sécurité pour permettre le trafic d'applications business entre les deux départements et l'accès internet. Les responsables de la sécurité doivent appliquer les politiques définies dans les interfaces du routeur pour assurer que seuls les flux des trafics permis sont transférés.

Cet exemple illustre la complexité opérationnelle de la configuration d'un réseau pour supporter une application. L'intervention exige un haut niveau de manipulation ma-

nuelle et concerne différents éléments de l'architecture. La totalité de la procédure est complexe et susceptible de provoquer diverses erreurs. Tout changement associé, comme le déplacement d'un serveur, l'extension d'un des réseaux, la modification de la configuration des tenants, implique la répétition quasiment complète du procédé et de la validation.

Dans le cas où la situation décrite doit être réalisée simultanément pour multiples consommateurs, il y aura un grand délai d'implémentation. Dans un contexte cloud, le traitement de demandes réseau durant plusieurs jours ou semaines devient critique et inacceptable. [26] [25] [24]

2.3.2 Interconnexion WAN

Ce scénario illustre la connexion des réseaux tenants au WAN pour l'accès aux sites distants. Par exemple, on pourrait imaginer, dans le cas précédent, l'interconnexion des deux tenants proposés à un autre site de l'entreprise (image 2.3). Dans l'approche traditionnelle, un VLAN doit être créé entre le routeur du data centre et le routeur WAN. Il peut être identifié par exemple comme VLAN-3.

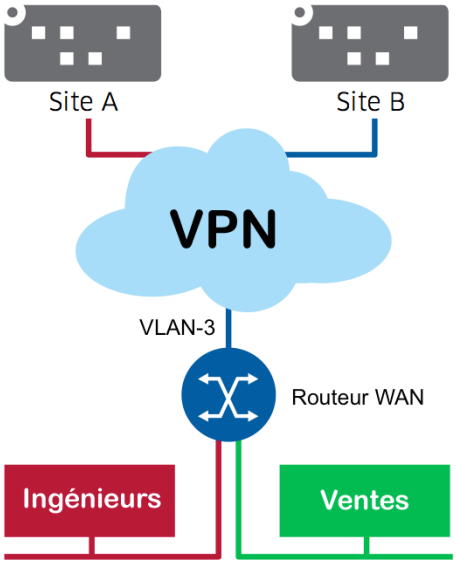

FIGURE 2.3 – Interconnexion des réseaux tenants sur deux sites. [27]

Un protocole de routage doit être défini pour fournir une résilience en cas de failles de communication. La configuration doit être appliquée dans les deux routeurs impliqués.

Pour offrir des mécanismes anti-failles aux niveaux 2 et 3 dans le WAN, plusieurs points VPNs doivent être déployés et configurés, ce qui ajoute de la complexité dans l'opération.

Traditionnellement, les réseaux du data centre et du WAN sont gérés par des responsables distincts. Cela implique la coordination entre deux équipes réseaux, en général via une structure formelle de projet et procédures manuelles. Par ailleurs, avec les demandes de divers tenants, ce modèle introduit des délais ainsi que des problèmes d'extensibilités supplémentaires lorsqu'ils requièrent des VLANs ou protocoles de routage dédiés.

Avec l'ajout de sites et tenants, une variété de pré-requis et configurations doivent être gérés et la complexité de déploiement de ce scénario devient critique. L'hétérogénéité des systèmes impliqués impose des défis difficiles à surmonter avec des délais de réalisation incompatibles. [27] [28]

2.3.3 Monitoring dans un data centre

Historiquement, les systèmes de monitoring se sont appuyés sur une variété de protocoles et une série fragmentée d'outils. Ces outils incluent la supervision de la performance du réseau, des applications et des analyses de sécurité, telles que **Intrusion Detection System, Système de Détection d'Intrusion (IDS)** et **Intrusion Prevention System, Système de Prévention d'Intrusion (IPS)**. En complément à l'inspection des paquets, ces outils utilisent plusieurs meta protocoles pour fournir des données sur les réseaux, tels que **SNMP**, NetFlow [29] et sFlow [30].

L'ajout de ces outils à l'opération du réseau est le premier pas vers une visibilité opérationnelle, permettant de détecter et solutionner les problèmes plus rapidement. La difficulté qui reste est d'alimenter ces outils avec la redirection du trafic vers leurs capteurs, pour qu'ils puissent superviser et gérer la quantité massive de données passantes d'une manière flexible et efficace.

La localisation des pannes est considérablement simplifiée si les données des paquets sont filtrées et rendues disponibles aux administrateurs du réseau. Précédemment, les flux de données circulaient principalement en direction Nord/Sud, le placement de la capture ou redirection du trafic était donc évident : dans les points d'entrée/sortie du réseau. Toutefois, avec les tendances actuelles de flux plutôt Est/Ouest, de mobilité des hôtes et de demandes plus importantes pour la performance, la supervision est devenue plus difficile et coûteuse dans les data centres modernes.

Les approches initiales pour la résolution de ces problèmes étaient l'agrégation du trafic drainé à des systèmes haute performance et envoyé aux outils de monitoring. Des dispositifs comme les **Network Packet Brokers (NPBs)** offrent des fonctionnalités additionnelles telles que manipulation ou enregistrement de la charge utile.

Le déploiement de ces outils pour chaque tranche réseau des tenants n'est pas efficace, particulièrement quand plusieurs instances d'outils sont requises pour le traitement d'un trafic excessivement important pour être supporté par une seule instance. En outre, l'ajout d'instances et outils exige des attentes répétitives pour la maintenance et implique diverses étapes pour le routage du trafic, connexion de systèmes de drainage et ensuite réversion à la configuration de routage initiale. Tout cela effectué manuellement par l'administrateur réseau. Une fois déployées, ces solutions fixent le trafic statiquement à un ensemble d'outils, ce qui peut conduire au cloisonnement des informations et à des difficultés pour les partager entre tenants. [31] [19]

2.4 Aspects de sécurité

Comme développé précédemment, on exige de nos jours que les applications délivrent immédiatement des informations et services spécifiques au contexte, à une latence réduite et à une haute performance. Parallèlement, le Cloud Computing et les applications orientées services introduisent des demandes plus spécifiques au niveau service. La mutualisation de ressources et les infrastructures multi-tenantes ont apporté de nouvelles préoccupations pour la sécurité qui n'existaient pas précédemment et qui ne sont donc pas traitées dans l'architecture traditionnelle. Cette section a pour but d'aborder quelques menaces et questions de sécurité introduites par la virtualisation et le Cloud Computing.

2.4.1 Points invisibles de la communication

Les applications de sécurité réseaux traditionnelles ne voient pas la communication entre VMs au sein du même hyperviseur, à moins que toutes leurs communications ne soient routées à l'extérieur de la machine hôte vers l'application de sécurité et re-routées à l'intérieur. Cette manière de traiter la problématique introduit un ralentissement considérable du réseau.

Dans le Cloud Computing, le moyen de traiter ce problème est d'intégrer des modules de sécurité dans chaque VM pour qu'elles puissent s'auto-protéger. Cette solution oblige la re-configuration de tous les systèmes en cas de mise à jour des politiques de sécurité. Il manque dans l'architecture actuelle un contrôleur central qui pourrait diffuser les politiques à tous les nœuds et VMs.

2.4.2 Attaques entre VMs et exposition de l'hyperviseur

Les serveurs virtualisés utilisent les mêmes systèmes d'exploitation et applications que les serveurs physiques. Les vulnérabilités retrouvées alors représentent des menaces pour les systèmes physiques et pour les environnements virtuels. De cette manière, quand un élément de l'environnement virtuel est compromis, l'ensemble du système est en position de risque dès lors qu'un système de sécurité intelligent et autonome n'est pas en place.

Dans ce scénario, un hacker peut attaquer un système invité, qui peut ensuite infecter d'autres VMs. Le risque augmente avec le nombre de VMs hébergées. Une protection capable de détecter des activités malveillantes au niveau des VMs doit être mise en place, indépendamment de la localisation des VMs dans l'environnement virtuel.

Un autre mode d'attaque concerne l'hyperviseur qui devient une cible d'attaque en raison de son rôle et ses responsabilités. Certaines attaques vont essayer de traverser l'espace d'isolation des VMs pour compromettre l'hyperviseur. Cela peut poser des problèmes plus graves lorsque des VMs qui contiennent des données sensibles finissent par résider dans le même hôte qu'une autre VM traitant des données moins critiques, entraînant différents niveaux de confiance dans un groupement de VMs. Il serait possible de séparer les VMs selon leurs niveaux de sécurité, mais ce principe contredit l'intention de la virtulisation de faire une meilleure utilisation des ressources.

Cette problématique implique un système de protection par VM, malgré la complexité déjà évoquée à synchroniser les politiques entre tous les systèmes. Cependant, sécuriser l'hyperviseur est également indispensable.

2.4.3 Lacunes sur les instantanées

La virtualisation et le Cloud Computing apportent des fonctionnalités telles que : approvisionnement, clonage, migration et désaffectation à la demande. Il en résulte que les VMs sont activées et inactivées à cycles rapides et il peut être difficile d'assurer une sécurité efficace de ces systèmes tout en les gardant à jour.

Après une période d'inactivité, l'écart des VMs par rapport à la base de sécurité peut être tellement important que la simple mise sous tension peut introduire des vulnérabilités considérables. Par ailleurs, même les VMs inactives peuvent être compromises. De cette manière, de nouvelles VMs peuvent être clonées à partir de templates avec une sécurité obsolète et être mises en ligne.

Quand ces VMs sont réactivées ou clonées, elles représentent une vulnérabilité instantanée à partir du moment où elles sont connectées au réseau. Une solution serait l'installation, pour chaque hôte, d'une VM dédiée pour la mise à jour des applications de sécurité des autres VMs qui sont lancées, ce qui permettrait de profiter des bénéfices de la virtualisation en toute sécurité.

2.4.4 Blocage de ressources

Quand des opérations gourmandes en termes de ressources sont affectées aux VMs, elle peuvent rapidement surcharger le système. Par exemple, quand les balayages ou mise-à-jour des antivirus sont réalisés simultanément sur plusieurs VMs dans un même serveur physique, le système peut ne pas supporter la consommation intensive des ressources (mémoire, CPU, réseau, stockage). Cela peut impacter la performance globale des applications en exécution.

Dans les systèmes physiques, des anti-virus installés dans chaque système d'exploitation consomment une mémoire additionnelle importante. L'application de cette architecture aux systèmes virtuels signifie une perte indésirable d'efficacité sur l'utilisation des ressources. Les produits qui ignorent la virtualisation proposent la randomisation (déclenchement à des moments aléatoires) ou le regroupement (regrouper les machines qui vont exécuter l'opération) pour éviter la charge intensive d'opérations simultanées.

Malheureusement, la randomisation ne permet pas d'éviter les longues périodes de haute consommation du système pour les cycles complets de balayage. Le groupement

ne contribue pas à la nature mobile de la virtualisation et exige une reconfiguration lors de la migration ou du clonage de VMs. [1] [32]

2.5 Complexité, Agilité et Sécurité : l'adaptation du réseau au Cloud

Les applications data centres ont évolué vers un modèle beaucoup plus dynamique pour accompagner et s'adapter aux demandes des consommateurs Cloud Computing. Ce nouveau mode de livraison et consommation exige une infrastructure beaucoup plus flexible et réactive pour fournir les ressources sollicitées.

En vue des scénarios étudiés, l'agilité est devenue cruciale dans ce contexte où les décisions se font de plus en plus en temps réel. Cet objectif ne peut pas être atteint sans la mise en place d'outils pour automatiser et orchestrer les éléments dans l'infrastructure. Dans l'aspect réseau, diverses opérations nécessaires sont toujours essentiellement manuelles et freinent le passage à ce modèle.

Le nombre de modules à gérer et leurs interfaces implique des contraintes pour le déploiement de nouvelles applications et services. La réponse aux divers cas d'utilisation Cloud Computing impose à la fois des opérations manuelles et une complexité de réalisation. Le développement d'outils pour automatiser ces scénarios introduit encore plus de complexité et ne s'adapte pas à tout data centre selon les modèles des équipements et protocoles utilisés.

En plus de ces exigences difficiles à traiter au moyen des réseaux traditionnels, l'environnement Cloud Computing exige une sécurité plus élevée du **SI**. Dans ce modèle, la performance, la haute disponibilité et la protection de données sont requises à un niveau plus important par rapport aux data centres traditionnels. L'architecture réseau actuelle n'est pas adaptée aux questions de sécurité ressortant de la virtualisation et du Cloud.

En résumé, il est difficile aujourd'hui d'atteindre le niveau d'agilité et sécurité exigé avec les technologies réseaux classiques sans introduire une complexité excessivement contraignante et donc bloquante pour l'implémentation du Cloud Computing. Dans le chapitre suivant, il sera analysé de nouvelles applications SDN qui proposent des solutions pour aborder ces problématiques.

Chapitre 3

Applications SDN et leurs apports aux data centres

Ce chapitre définira SDN et présentera les réponses aux problématiques réseau rencontrées en général dans les data centres et qui ont été débattues dans le chapitre précédent.

3.1 Définition de SDN

Les pré-requis réseaux des applications devraient être articulés dans un langage simple, utilisé pour demander les comportements réseau souhaités. La programmation des réseaux est aujourd'hui très limitée, introduisant un malheureux compromis : soit l'obligation de développer les applications par un paramétrage réseau excessivement détaillé, ou ignorer ces détails en traitant du réseau comme une boîte noire. Aucune de ces solutions n'est convenable ; la première implique des applications spécifiques à chaque type de réseau alors que dans la deuxième, le contrôle nécessaire n'est pas présent.

Un niveau approprié d'abstraction des réseaux est nécessaire pour les rendre programmables et améliorer l'utilisation des ressources. L'instanciation de services réseau doit pouvoir se faire instantanément, de manière alignée aux besoins des applications. Le réseau doit permettre l'établissement de connectivité interne et entre data centres, tout en gardant la cohérence avec les politiques (de sécurité, de disponibilité et de réglementation) définies par le prestataire cloud et ses tenants. Aujourd'hui, tout cela se fait lentement, manuellement avec un risque élevé d'erreurs, comptant sur les ordres

de travail pour établir une multitude de configurations selon le fournisseur de chaque équipement.

Les principes **SDN** proposent une couche d'abstraction pour permettre la programmation du réseau et autoriser le contrôle dynamique de services. SDN est un nouveau **paradigme** réseau défini comme une architecture qui a pour but de centraliser l'intelligence du réseau sur un contrôleur. Cette intelligence étant traditionnellement distribuée parmi plusieurs équipements réseaux réalisant une fonction spécifique dans l'infrastructure.

Ces dispositifs ont en général une fonctionnalité de commutation de paquets (**plan de données**) et une partie de traitement des données avec une logique spécifique selon les états et la configuration enregistrés (**plan de contrôle**). SDN propose de dissocier ces deux fonctions dans les dispositifs et d'agréger dans un contrôleur commun l'activité de traitement. L'image ci-dessous permet de visualiser les différences entre les architectures traditionnelle et SDN.

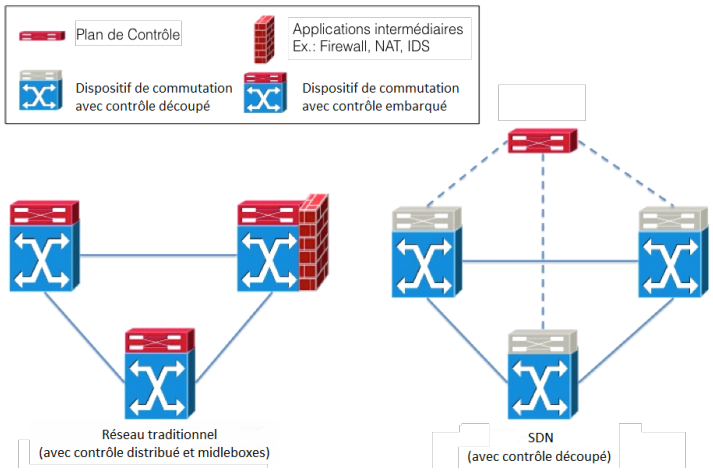

FIGURE 3.1 – Architectures : réseau traditionnel et SDN. [33]

L'expérience, lors de la construction des grands réseaux IP/MPLS, a montré que l'intelligence doit être poussée vers la périphérie pour autoriser la **scalabilité** des réseaux. Le principe doit être appliqué afin de promouvoir un cœur du réseau data centre simplifié et efficace. L'approche sépare les services réseaux de l'infrastructure physique permettant l'innovation parallèle dans les deux domaines.

Les leçons acquises lors de la conception des réseaux mobiles ont apporté la technique d'optimisation pour la mobilité et d'automatisation opérationnelle à large échelle. Cela

permet la création d'un modèle auto-instancié et dirigé selon les règles établies qui minimise les coûts d'implémentation et délais de livraison de service.

Cette solution apporte comme bénéfice l'introduction d'une interface commune et programmable de management, permettant la mise en place dynamique de services, indépendamment de la marque/modèle des dispositifs réseau. Avec SDN, l'administration réseau devient plus agile car un seul élément (le contrôleur) est à maîtriser au lieu d'avoir à configurer l'ensemble des équipements du système, ce qui accélère considérablement le temps de convergence du réseau pour l'accommodation de nouvelles applications déployées. [34] [35]

3.2 Virtualisation des fonctions réseau, NFV

Un concept qui est très souvent évoqué en parallèle à **SDN** est **Network Functions Virtualization, Virtualisation des fonctions réseau (NFV)**, une initiative fédérée à la base par les plus grands fournisseurs de services de télécommunication du monde avec l'**European Telecommunications Standards Institute (ETSI)** et qui a généré un grand intérêt chez les industriels. L'initiative a été créée pour aborder les principaux défis opérationnels et les coûts de gestion des applications réseau propriétaires et fermées qui sont actuellement déployées.

NFV pousse les technologies de virtualisation à consolider les applications réseau sur des serveurs, switches et baies de stockage physiques et standard de l'industrie, dans une couche logique qui met à disposition des ressources réseau. Une ressource réseau virtualisée représente une **Virtual Network Function, Fonction Réseau Virtuelle (VNF)**. Le principe des **VNFs** sert à augmenter la flexibilité à partager ressources et réduire les coûts de configuration et gestion. Un fournisseur de services peut mettre à disposition un ensemble d'applications et éléments d'infrastructure dans une plateforme qui fournit aux tenants les ressources dont ils ont besoin pour déployer leurs propres applications réseau adaptées à leurs objectifs métiers.

La virtualisation des fonctions réseau non seulement réduit les dépenses en équipements, mais apporte aussi d'autres bénéfices tels que l'extension agile d'applications, à une vitesse plus rapide, une plus haute disponibilité et une meilleure utilisation de ressources. Toutefois, pour bénéficier de ces apports, il est nécessaire que l'infrastructure réseau sous-jacente s'adapte rapidement et automatiquement. Par exemple, pour

migrer une fonction réseau vers un nouveau matériel, les politiques et les configurations associées à ce service doivent être provisionnées dans beaucoup d'autres équipements et fonctions. La complexité à configurer les réseaux dans un environnement si dynamique augmente énormément avec l'introduction de nouveaux éléments réseaux, c'est pourquoi les technologies qui supportent la programmation du réseau telles que SDN, pourront habiliter la virtualisation des fonctions réseau.

Bien que les développements de SDN et NFV puissent évoluer indépendamment, l'association des deux principes est d'un fort intérêt pour progresser dans les solutions cloud. SDN peut être employé en tant que technologie facilitatrice de la virtualisation des fonctions réseau, favorisant la consolidation des applications réseau dans des dispositifs industriels standard. [36] [37] [38]

3.3 SDN, NFV et le Cloud Computing

Les approches cloud permettent aux opérateurs réseau d'assurer une création et un déploiement de services plus rapides. Elles répondent également aux attentes croissantes sur la qualité et la performance des solutions, tout en traitant les charges trafics de plus en plus importantes.

Afin de supprimer les contraintes des réseaux dans les data centres, une plateforme innovante pour l'abstraction des fonctionnalités réseau et pour l'instanciation automatisée de services est proposée. Avec SDN, les fournisseurs de services cloud, opérateurs à l'échelle web et grandes entreprises technologiques peuvent construire une infrastructure réseau multi-tenante, robuste et extensible pour délivrer des espaces virtuels de compute, stockage et réseau prêts à l'usage pour des milliers de tenants et groupes d'utilisateurs.

Le **plan de contrôle** dans SDN fonctionne en lien avec le système de gestion cloud afin de configurer dynamiquement des éléments réseaux pour s'adapter aux décisions des systèmes d'orchestration pour le changement d'utilisations des ressources. SDN fournit l'infrastructure nécessaire pour réaliser cette capacité de NFV étendue.

La figure 3.2 affiche comment SDN et NFV s'intègrent dans une infrastructure Cloud pour atteindre les objectifs de virtualisation du réseau.

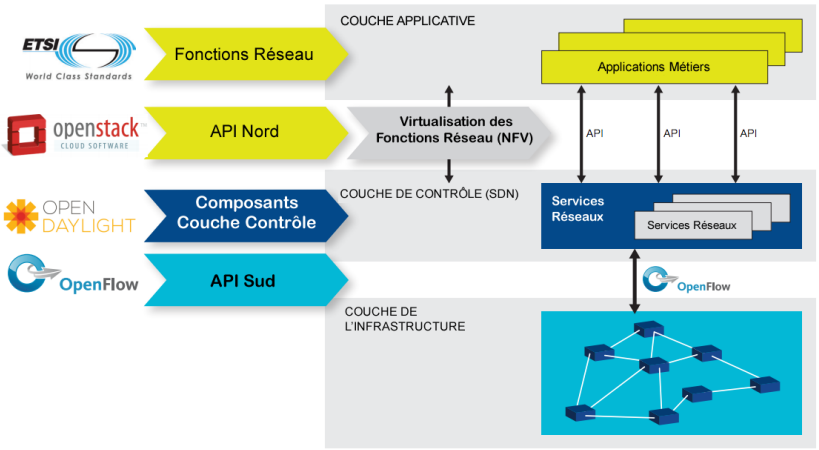

FIGURE 3.2 – SDN, NFV et Cloud Computing. [39]

L'image montre également les protocoles et standards ouverts qui émergent pour chaque technologie comme **OpenFlow** pour la communication entre les dispositifs réseau et le contrôleur SDN (**Open Daylight**) et **OpenStack** pour la gestion Cloud. Il est important de noter que de grands noms de l'industrie (comme Cisco, Microsoft, Google etc.) se sont engagés dans ces projets pour accélérer le développement de ces technologies.

Dans le Cloud, avec SDN et NFV travaillant ensemble, on peut dynamiquement étendre les fonctions réseau au sein des data centres. Par exemple, lorsque la charge réseau augmente, le contrôleur SDN peut demander au gestionnaire cloud d'instancier une nouvelle fonction réseau dans le Cloud pour redistribuer et répartir le trafic.

La couche de contrôle SDN apporte l'allocation de ressources extensible et en temps-réel pour les besoins des services réseaux. Cela permet que les services soient définis et provisionnés très rapidement via des portails en libre service. Le service est mis à disposition en quelques minutes, au lieu de jours ou même semaines comme il est traditionnellement nécessaire.

Cette plateforme, avec un contrôle intégré tout au long des domaines réseaux, expose des **Application Programming Interface, Interface de Programmations (APIs)** qui permettent aux applications d'instancier dynamiquement des ressources pour accomplir ces objectifs métiers. Un système de management de l'infrastructure fournit des capacités supplémentaires, comme une planification plus fiable, approvisionnement, activation, adaptation et le contrôle de nouvelles connexions de services.

SDN permet de coupler la gestion du Cloud au réseau programmable, achevant une intégration complète du réseau. Avec une orchestration commune des services, on obtient une réduction des coûts opérationnels pour l'approvisionnement et la supervision ainsi qu'une création flexible de services. Cela rend le réseau dynamique, adaptatif et agile, donc prêt pour le Cloud Computing. [39] [40]

3.3.1 Cas d'utilisation SDN-NFV : NFVIaaS

Pour illustrer un exemple d'utilisation de NFV favorisé par SDN, une étude de cas est proposée par l'**ETSI** : **NFVIaaS**, Infrastructure NFV en tant que service pour la virtualisation générale du réseau. L'exemple met en évidence les avantages du comportement hautement flexible et dynamique d'un réseau SDN.

Le cas d'utilisation NFVIaaS est présenté comme étant essentiel pour la livraison de services cloud. Dans ce scénario, le fournisseur de services peut les offrir en utilisant une infrastructure NFV (NFVI) d'un autre fournisseur de services. L'intérêt de cette capacité est qu'elle permet d'étendre la portée d'un fournisseur à des localisations où il ne maintient pas une présence physique.

La figure ci-dessous illustre le concept de NFVIaaS. Dans cet exemple, le fournisseur de service X offre un service de répartition de charges. Quelques clients de l'opérateur X ont besoin de ces services dans une localisation où l'entreprise ne possède pas d'infrastructure, mais le fournisseur Z, lui en possède.

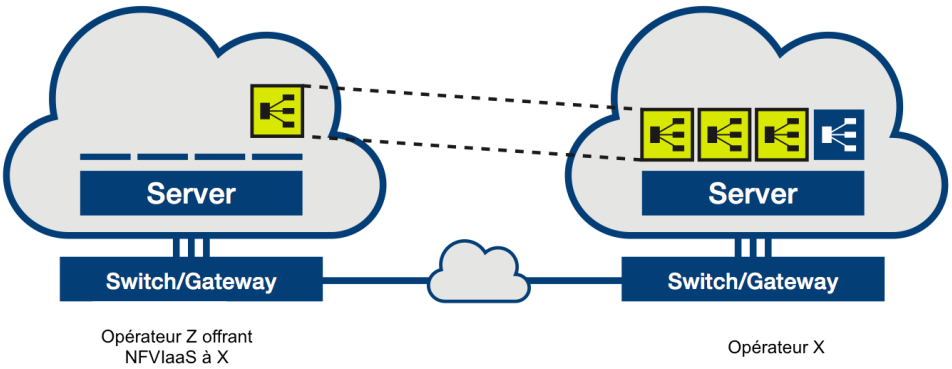

FIGURE 3.3 – Infrastructure NFV en tant que Service. [39]

NFVIaaS offre le moyen pour l'opérateur Z de louer son infrastructure NFV (compute,

stockage, réseau) à X, lui permettant ainsi l'accès à une infrastructure qui, ailleurs, serait d'un coût prohibitif à obtenir.

La valeur ajoutée à ce scénario est considérable. En éliminant le coût et la complexité à déployer du nouveau matériel pour les louer à coûts fixes, l'opérateur X peut déployer et étendre des services virtualisés rapidement. Z bénéficie de gains financiers sur une capacité excédante, associée à une plus grande efficacité de ses investissement en infrastructure NFV et SDN.

NFVIaaS représente également une simplification de déploiement par l'abstraction des différences parmi divers opérateurs, autrement gérées par de l'encapsulation de tunnels, de la **Quality of Service, Qualité de service (QoS)**, par un renforcement de politiques de sécurité et des procédures opérationnelles. Dans ce cas, d'autres requis importants sont à prendre en compte comme l'aspect multi-tenant pour l'isolation suffisante du trafic, le renforcement de politiques spécifiques aux tenants et la **scalabilité**.

Comme pour SDN, NFV est fondée sur un environnement ouvert et multi-vendeur pour maximiser les choix et réduire le **Capital Expenditure, Coûts d'Investissement (CapEx)**. Une plateforme commune d'automatisation doit être capable d'approvisionner l'infrastructure (physique et virtuelle) et être accompagnée d'un modèle de déploiement qui élargit les limites géographiques du fournisseur de services. Tout cela, ne pouvant pas être accompli sans un contrôle automatisé du réseau, qui sera implémenté de manière plus efficace par SDN. [39] [41]

3.4 Scénarios d'utilisation

Un système cloud qui s'intègre de façon transparente et dynamique avec un réseau programmable grâce à SDN peut fournir une importante plus-value à ses opérateurs et à leurs abonnés (consommateurs finaux et entreprises). Aujourd'hui la connectivité seule ne suffit pas, les utilisateurs réclament une variété de services hébergés dans le Cloud, et cela exige des réseaux la capacité de fournir la connectivité correcte à l'application souhaitée. C'est dans ce cadre que la réelle valeur d'un Cloud à réseau dynamiquement programmable devient visible.

Cette capacité permet de découper le réseau en tranches et offrir aux clients leurs morceaux dédiés et personnalisés. Il existe une variété de scénarios imaginables à partir de ce concept de diviser le réseau et services pour convenir à différents applications et besoins.

Un des cas d'utilisation est l'infrastructure virtuelle d'entreprise, dans laquelle un portail basé sur SDN peut être étendu selon les particularités de l'organisation. La solution associe la coordination riche d'un contrôleur cloud et d'un contrôleur SDN. Cela permet l'instanciation, la réplication et la migration du réseau et services basés cloud dans la meilleure localisation disponible, en fonction des pré-requis des tenants, de la congestion globale du réseau et de la disponibilité de ressources. Cette solution est conforme à l'idéal de ne pas limiter le Cloud avec les contraintes physiques du data centre, en implémentant un suivi de flux et un renforcement de politiques dans un niveau logique pour le Cloud. Cela peut englober plusieurs data centres, quelles que soient leurs localisations géographiques dans l'infrastructure physique du réseau.

Un des scénarios les plus traditionnels de l'intégration des services dynamiques avec SDN consiste à en resserrer l'interaction entre le réseau et le Cloud. Pour les services "inline" tels que filtrage, modification des entêtes et **Network Address Translation, Traduction d'adresse réseau (NAT)**, les opérateurs utilisent diverses "appliances", ou d'autres services pour gérer le trafic utilisateur. Ces services sont hébergés dans du matériel physique ou en machines virtuelles. L'enchaînement de services est nécessaire pour router le trafic client à travers ces services. Les solutions traditionnelles sont soit statiques ou très limitées en flexibilité et **scalabilité**.

Dans le chapitre précédent, les principaux défis des réseaux traditionnels pour fonctionner en mode cloud ont été détaillés. Un résumé simplifié des principales difficultés

a été proposé. Les prochaines sections, traitent ces points et illustrent comment ces objectifs peuvent être atteints en utilisant SDN. [42]

3.5 Complexité et Agilité

La complexité et l'agilité sont des éléments qui sont fortement associés et doivent donc être étudiés ensemble. En effet, la complexité réduit considérablement l'atteinte des objectifs d'agilité, et l'agilité augmente la complexité dans l'approche réseau traditionnelle. Cette impasse a pu être observée dans les scénarios présentés précédemment.

Les architectures réseaux implémentant SDN proposent d'aplatir physiquement la topologie avec l'interconnexion de tous les élément à un **fabric** pour à la fois simplifier le réseau et gagner en agilité. L'image ci-dessous visualise cette proposition. Le contrôleur SDN peut être programmé pour reproduire logiquement l'infrastructure souhaitée par chaque tenant.

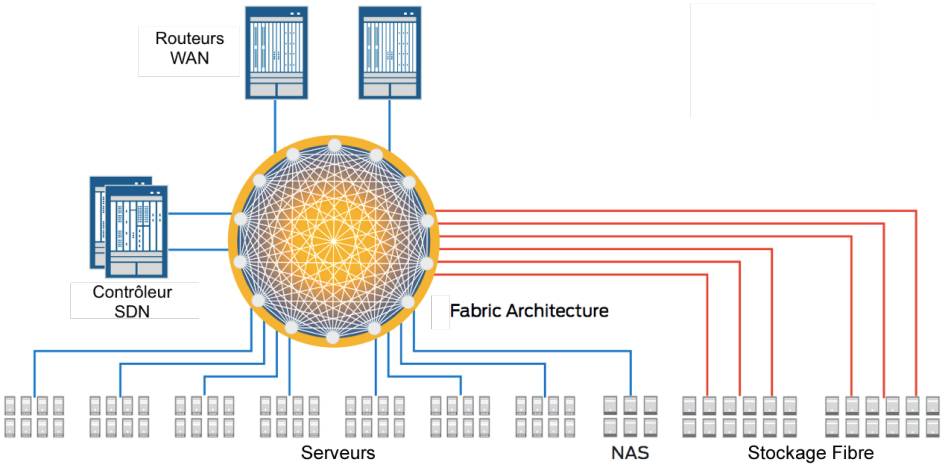

FIGURE 3.4 – Topologie réseau simplifiée. [19]

Avec les solutions SDN, les changements sur le réseau sont traités par un processus complètement automatisé qui peut réagir instantanément à divers événements. Comme effectué pour les VMs, SDN permet de créer des templates de réseaux logiques des tenants qui peuvent être facilement instanciés pour implémenter tous les aspects réseaux requis de manière automatique pour de nouveaux besoins.

Les templates peuvent être réutilisés plusieurs fois, si l'opération doit être répétée.

Les solutions SDN peuvent créer de manière automatique le réseau nécessaire et fournir l'isolation complète entre les tenants, ce qui simplifie considérablement les opérations réseaux au sein des data centres.

Un autre avantage est une réponse plus rapide aux demandes des clients avec une simplification opérationnelle. Grâce à sa capacité de programmation et d'abstraction, SDN élimine les opérations manuelles très sensibles aux erreurs, avec une meilleure efficacité du réseau, réduction de coûts opérationnels et d'investissement. [25] [27]

3.6 Sécurité

La sécurité dans un data centre cloud doit protéger le trafic entre les clients et les serveurs, le trafic entre les machines virtuelles dans les serveurs ainsi que le trafic entre serveurs (physiques ou virtuels) et les applications ou systèmes dans d'autres data centres. La capacité d'extension représente un requis de sécurité essentiel dans ces environnements.

L'importante croissance des accès utilisateurs ainsi que de la sophistication des menaces de sécurité dans un data centre cloud exigent une visibilité étendue du réseau et la mise en place des protections associées. Les solutions de sécurité doivent à la fois renforcer les politiques de manière complète et rester flexibles pour assurer l'adaptation du réseau aux divers usages.

Pour aborder ces défis, la sécurité d'un Cloud doit réunir des capacités telles que **scalabilité**, monitoring et contrôles renforcés. Les services de sécurité doivent être consolidés et coordonnés en complément de la simplification et la mutualisation agile du réseau. Cette approche améliore la flexibilité et l'efficacité du système entier.

SDN propose des moyens pour fournir des services dynamiques de sécurité et pour répondre aux pré-requis de performances tout en accommodant les évolutions futures pour répondre aux nouvelles demandes. Des services tels que supervision, filtrage, détection et prévention d'intrusion et VPNs sont consolidés dans une plateforme extensible avec des ressources affectées dynamiquement. La solution améliore également chaque service de sécurité en augmentant de manière dynamique la capacité d'accès à tout flux de trafic au sein du **fabric** réseau.

Les services de sécurité doivent pouvoir reconnaître les applications, tout en étant mobiles. L'interface de programmation fournie avec SDN permet l'interaction avec les hyperviseurs en vue d'assurer la sécurité inter-VMs. Les applications peuvent déployer un ensemble de politiques de sécurité à partir des hyperviseurs et à travers le **fabric** réseau, grâce à l'établissement d'un lien entre les couches virtuelles et physiques.

Il en résulte pour les consommateurs la possibilité de profiter des bénéfices de la virtualisation, tout en étant protégés contre les risques de sécurité associés. SDN surmonte les deux principales préoccupations concernant les plateformes cloud : la sécurité et le contrôle. Le réseau, étant en contact avec tous les éléments de l'infrastructure, est le meilleur endroit pour placer la sécurité et la gestion. Avec SDN, ces défis peuvent être abordés et adaptés aux besoins du Cloud Computing. [32] [31] [28]

Conclusion

Au sein des data centres, les technologies de virtualisation des serveurs ont évolué considérablement pour accompagner les nouveaux besoins clients. De nouvelles technologies pour la virtualisation du stockage se trouvent également disponibles sur le marché. Toutefois, la réelle valeur du Cloud Computing ne pourra pas être atteinte sans une évolution similaire dans les aspects réseau.

On constate donc qu'il est impératif de faire évoluer les technologies réseau, mais des résistances s'opposent à cette évolution en raison de la complexité et d'une possible saturation du système. Les opérateurs cloud ont des difficultés à adapter l'architecture réseau traditionnelle au rythme actuel des demandes pour assurer le niveau de sécurité exigé.

Des contraintes de complexité opérationnelle et de sécurité sur les réseaux Cloud empêchent le déploiement agile de nouveaux services et applications. Les réseaux deviennent donc une cible de critiques, dont le principale reproche est de freiner le rythme d'innovation espéré aujourd'hui.

En réponse à ces difficultés, les réseaux programmables ont été un objet intensif de recherche par la communauté. Les travaux dans ce domaine s'orientent vers la virtualisation du réseau (NFV) facilitée par l'offre SDN, nouveau paradigme qui transforme l'architecture traditionnelle.

L'approche SDN sépare le plan de contrôle et le plan de données, offrant ainsi un contrôle et une vision centralisés du réseau. Cela peut apporter certains bénéfices comme le contrôle directement programmable, la simplification des équipements et l'ingénierie du trafic.

Le contrôleur SDN, en association avec l'orchestration cloud, permet d'approvisionner dynamiquement les réseaux et de manière simplifié, tout en assurant la sécurité et la qualité de services nécessaires. Même si l'approche est encore récente, elle est suivi

attentivement par le marché qui accompagne la parution des premières offres SDN, proposées par de grandes sociétés ainsi que par des startups.

Bien que la technologie se développe à un rythme accéléré, pour que son adoption soit plus répondue il faut que sa consolidation soit bien établie. Le marché bouge, avec des acteurs proposant des options stratégiquement différentes selon leurs produits de base, et le consommateur final redoute de ne pas choisir la bonne solution.

Parallèlement, le besoin d'une infrastructure plus agile, intégrée par le Cloud, se fait toujours sentir et les fournisseurs ont intérêt à réagir rapidement dans la bataille pour les parts de marché. Ainsi, ceux qui les premiers auront osé saisir cette opportunité, seront ceux qui dessineront le futur de la technologie des réseaux informatiques pour les prochaines années.

Index

Acronymes

ACI Application Centric Infrastructure, Infrastructure centrée sur les applications

API Application Programming Interface, Interface de Programmation

ASIC Application Specific Integrated Circuit, Circuit intégré pour application spécifique

CapEx Capital Expenditure, Coûts d'Investissement

DHCP Dynamic Host Control Protocol, Protocole pour la configuration automatique d'hôte

DNS Domain Name System, Système de noms de domaine

DPI Deep Packet Inspection, Inspection de Paquets en Pronfondeur

ETSI European Telecommunications Standards Institute

GRE Generic Routing Encapsulation

HTTP HyperText Transfer Protcol, Protocole de transfert de hypertexte

IaaS Infrastructure as a Service, Infrastructure en tant que service

IDS Intrusion Detection System, Système de Détection d'Intrusion

IETF Internet Engineering Task Force, Détachement d'ingénierie d'internet

IP Internet Protocol, Protocole d'Internet

IPS Intrusion Prevention System, Système de Prévention d'Intrusion

IRTF Internet Research Task Force, Détachement de recherche d'internet

LAN Local Area Network, Réseau local

MPLS MultiProtocol Label Switching, Commutation multi-protocoles par étiquettes

NAT Network Address Translation, Traduction d'adresse réseau

NFV Network Functions Virtualization, Virtualisation des fonctions réseau

NOS Network Operating System, Système d'exploitation réseau

NPB Network Packet Broker

NVGRE Network Virtualization using Generic Routing Encapsulation

ONE Open Network Environment, Environnement Réseau Ouvert

ONF Open Networking Foundation

OpEx Operational Expenditure, Coûts Opérationnels

QoS Quality of Service, Qualité de service

ROI Return of Investiments, Retour sur Investissement

SDN Software-Defined Networking : Réseau Informatique Défini par Logiciel

SDS Software Defined-Storage, Stockage Défini par Logiciel

SI Système d'Information

TI Technologie de l'Information

VLAN Virtual Local Area Network, Virtual LAN

VM Virtual Machine, Machine Virtuelle

VNF Virtual Network Function, Fonction Réseau Virtuelle

VPN Virtual Private Network

VXLAN Virtual eXtensible LAN

WAN Wide Area Network, Réseau étendu

Glossaire

Abstraction En informatique, l'abstraction est un terme souvent employé pour désigner le mécanisme et la pratique qui réduisent et factorisent les détails négligeables de l'idée exprimée afin de se focaliser sur moins de concepts à la fois. C'est aussi la notion de couches d'abstraction utilisée comme moyen pour gérer la complexité des systèmes informatiques lorsque les couches correspondent à des niveaux de détails appliqués. [43]

Big Data Big Data est un terme appliqué aux ensembles de données dont la taille (ou le format) est supérieure à la capacité des outils logiciels communs, qui ne peuvent plus les capturer, les gérer et les traiter. Une nouvelle classe de technologies et outils a été développée pour attribuer une valeur commerciale à ces données grâce à une analyse complexe. Le terme est employé en référence à ce type de données ainsi qu'aux technologies utilisées pour les stocker et les traiter. [44]

Cloud Computing Cloud Computing, ou informatique dans les nuages, est une évolution de la livraison de services **TI**. Le Cloud Computing offre un moyen d'optimiser l'usage et le déploiement rapide de ressources. Cela se fait par des systèmes et solutions plus efficaces et **scalables**, en fournissant un niveau plus haut d'automatisation. Diverses entreprises ont adopté le Cloud Computing et obtiennent des avantages significatifs en agilité, réduction de coûts et croissance du business. [45]

Cluster En réseaux informatiques, un cluster désigne un groupe de machines reliées entre elles à l'aide d'un réseau de communication. Cette configuration est souvent utilisée pour réaliser des calculs de haute performance. [46]

Compute Le terme compute a été utilisé tout au long du document pour faire référence à la capacité de calcul et traitement de l'information des serveurs physiques. C'est l'élément qui est virtualisé au moyen des hyperviseurs qui permettent la mutualisation de ces ressources entre différentes applications.

Data Centre Centre de traitement de données. Il s'agit d'une installation utilisée pour héberger des systèmes informatiques et les composants associés, comme les systèmes de télécommunication et de stockage. En général, un data centre inclut alimentation et connexions des données redondantes, contrôles d'environnements comme la climatisation ainsi que divers dispositifs de sécurité. [47]

DPI DPI est une technologie de filtrage de paquets conçue pour les examiner au moment où ils passent par le point d'inspection pour rechercher une non conformité de protocole, virus, spam, intrusion ou autre critère défini. L'outil décide alors si le paquet doit être transmis à la destination, re-routé ou bloqué. DPI fournit des données pour permettre une automatisation du réseau, une conformité aux règles définies et d'autres fonctions de sécurité. [48]

Fabric En informatique, fabric (qui signifie tissu en anglais) est un synonyme de plateforme ou structure. En général, le terme fabric décrit la façon dont différents composants travaillent ensemble pour former une entité unique. Dans ces systèmes la liaison entre les composants est tellement dense qu'un schéma représentant leurs relations ressemblerait à une pièce de tissu tricotée. Sous ce terme généralement admis par l'industrie réseau, un fabric est une topologie réseau dans laquelle les composants se transmettent des données l'un à l'autre à travers les switches d'interconnexion. [49] [50]

Middlebox Boîtier intermédiaire. Un middlebox est un serveur conservant les états de la communication entre deux hôtes. Ils se différencient des hôtes qui représentent les extrémités de la communication. Ils sont encore différents des routeurs qui ne gardent pas d'états sur les sessions de communications. [51]

MPLS MPLS définit une distinction explicite entre la périphérie et le cœur du réseau. Les routeurs dans la périphérie inspectent les entêtes des paquets entrants et rajoutent une étiquette utilisée pour les transmissions dans le cœur du réseau. Cette commutation basée sur les étiquettes qui sont utilisées pour délivrer les paquets et répondre aux pré-requis des opérateurs, comme tunnel **Virtual Private Network (VPN)**. ou l'ingénierie du trafic. Les étiquettes MPLS ont du sens uniquement dans le cœur du réseau et sont complètement indépendantes des protocoles spécifiques aux hôtes (IPv4, IPv6 ou autre). MPLS ne formalise pas l'interface de contrôle pour les opérateurs. MPLS distingue ainsi les réseaux hôtes des interfaces de commutation des paquets, mais ne développe pas une interface générale pour les opérateurs réseau. [52]

Open Daylight Association initiée par Linux Foundation pour l'union des géants du marché réseau dans le but de développer un contrôleur SDN **open source**, pour l'innover, l'encourager et pour permettre son adoption accélérée. [53]

Open Source Logiciel avec code source ouvert, qui peut donc être utilisé librement, modifié et partagé par quiconque. Un logiciel open source est développé par plusieurs personnes et est distribué sous des licences qui se conforment à la définition d'open source. [54]

OpenFlow Le protocole OpenFlow vise à standardiser l'interface entre les applications et le contrôleur ainsi que l'interface entre le contrôleur et les éléments de commutation. [33] [55]

OpenStack OpenStack est un projet pour le développement d'une plateforme Cloud Computing **open source**. Le projet a été initié par la NASA et Rackspace en 2010, et est rejoint actuellement par diverses entreprises telles que HP, Dell, Cisco etc. [56]

Paradigme Un paradigme consiste en une collection de règles, de standards et exemples de pratiques scientifiques, partagés par un groupe de scientifiques. Sa genèse et poursuite en tant que tradition de recherche sont conditionnées à un fort engagement et consensus des personnes impliquées. [57] D'après Dosi [58], quand un nouveau paradigme technologique apparaît, il représente une discontinuité ou un changement dans la manière de penser. Ce changement apporté par le paradigme est souvent lié à une tentative d'innovation radicale qui implique une nouvelle technologie. Dans ce document, le terme paradigme sera employé dans ce sens d'innovation et application de nouvelle technologie.

Plan de Contrôle Intelligence du réseau, ensemble des données locales utilisées pour établir les entrées des tableaux de commutation, qui sont utilisés par le plan de données pour effectuer la transmission du trafic entre les ports d'entrée et de sortie du dispositif. [59]

Plan de Données Le plan de donnés traite les data-grammes entrant dans le média à travers une série d'opérations au niveau des liens qui collectent ces data-grammes et réalisent divers tests basiques de cohérence. Ensuite les data-grammes sont transférés en accord avec des tableaux pré-remplis par le **plan de contrôle**. [59]

Scalabilité Terme provenant de l'anglais *scalability* qui exprime la capacité à être mis à échelle. En informatique cela désigne la capacité d'un système, d'un réseau ou un processus à gérer l'augmentation ou la réduction de la charge de manière à pouvoir la gérer. [60]. Le terme est souvent employé pour exprimer une extensibilité,

évolutivité ou passage à l'échelle, mais il n'y a « pas d'équivalent communément admis en français ». [61]

SNMP Simple Network Management Protocol (SNMP) est un protocole largement utilisé pour la supervision d'équipements réseau, serveurs et autres dispositifs connectés. [62]

Virtualisation Pour diverses entreprises, l'infrastructure serveur virtualisée est la base sur laquelle le **cloud** est construit. Initialement, les technologies de virtualisation ont permis aux data centers de consolider leurs infrastructures pour réduire les coûts. Avec le temps, l'intégration des technologies pour le management flexible de ressources a facilité une allocation plus dynamique. Cela a permis la réduction des coûts et a également augmenté la flexibilité et la performance. [45]

Bibliographie

[1] *Creating the Cloud-Ready Data Center, section 2. The challenges of the virtualized data center*. Technology White Paper. Alcatel-Lucent, HP. 2010.

[2] Albert Greenberg et al. « The cost of a cloud: research problems in data center networks ». In : *ACM SIGCOMM Computer Communication Review* 39.1 (2008), p. 68–73.

[3] Albert Greenberg et al. « The cost of a cloud: research problems in data center networks, section 3. Agility ». In : *ACM SIGCOMM Computer Communication Review* 39.1 (2008), p. 68–73.

[4] *Examining the Impact of Security Management on the Business*. Executive Summary. An AlgoSec Survey. 2013.

[5] *Solving Critical Challeges of the Virtualized Data Center*. Executive Summary. Market Pulse. 2011.

[6] *Understanding Data Centers and Cloud Computing, Section What Is a Data Center?* White Paper. Global Knowledge Training LLC. 2010.

[7] Luiz Andre Barroso et Urs Holzle. « The datacenter as a computer: An introduction to the design of warehouse-scale machines, Chapitre 1 : Introduction ». In : *Synthesis lectures on computer architecture* 4.1 (2009), p. 1–108.

[8] Krishna Kant. « Data center evolution: A tutorial on state of the art, issues, and challenges, section 2. Data center organization and issues ». In : *Computer Networks* 53.17 (2009), p. 2939–2965.

[9] Sandeep Raghuraman. *The Journey Toward the Software-Defined Data Center*. White Paper. Cognizant (NASDAQ: CTSH). Sept. 2013.

[10] M. Girola et al. *IBM Data Center Networking: Planning for Virtualization and Cloud Computing, chapitre 2 : Servers, storage, and software components*. IBM redbooks. IBM Redbooks, 2011. isbn : 9780738435398. url : `http://books.google.fr/books?id=EKjEAgAAQBAJ`.

[11] Harish Ganesan. *AWS Cost Saving Tip 5: How Amazon Auto Scaling can save costs*. Web Site. http://harish11g.blogspot.fr/2013/04/Amazon-Web-Services-AWS-Cost-Saving-Tips-how-Amazon-AutoScaling-can-reduce-leakage-save-costs.html. Avr. 2013.

[12] Kapil Bakshi. *Cisco Cloud Computing - Data Center Strategy, Architecture, and Solutions*. White Paper. Cisco Systems, Inc. 2009.

[13] *Virtualization & TCO: Linux vs. Microsoft*. Sondage. Gabriel Consulting Group, Inc. 2009.

[14] M. Girola et al. *IBM Data Center Networking: Planning for Virtualization and Cloud Computing, chapitre 1 : Drivers for a dynamic infrastructure*. IBM redbooks. IBM Redbooks, 2011. isbn : 9780738435398. url : `http://books.google.fr/books?id=EKjEAgAAQBAJ`.

[15] *Cloud-Ready Data Center Reference Architecture, Section : Framework*. White Paper. Juniper Networks, Inc. Mai 2010.

[16] M. Girola et al. *IBM Data Center Networking: Planning for Virtualization and Cloud Computing, chapitre 4 : The new data center design landscape*. IBM redbooks. IBM Redbooks, 2011. isbn : 9780738435398. url : `http://books.google.fr/books?id=EKjEAgAAQBAJ`.

[17] V. Josyula, M. Orr et G. Page. *Cloud Computing: Automating the Virtualized Data Center, chpaitre 3 : Data Center Architecture and Technologies*. Networking Technology. Pearson Education, 2011. isbn : 9780132604048.

[18] *Effects of virtualization and cloud computing on data center networks, Section : Introduction*. Technology Brief. Hewlett-Packard Development Company, L.P. Oct. 2011.

[19] *The Cloud-Ready Data Center Network, applying the lessons of cloud computing to vastly improve economics of networking and the user experience*. White Paper. Juniper Networks, Inc. Juin 2012.

[20] *Effects of virtualization and cloud computing on data center networks, Section : Changing business applications*. Technology Brief. Hewlett-Packard Development Company, L.P. Oct. 2011.

[21] *Cloud-Ready Data Center Reference Architecture, Section : Network Infrastructure*. White Paper. Juniper Networks, Inc. Mai 2010.

[22] P. Raj et G.C. Deka. *Handbook of Research on Cloud Infrastructures for Big Data Analytics, chapitre : The Network Infrastructures for Big Data Analytics*. Advances in data mining and database management (ADMDM) book series. IGI Global, 2014. isbn : 9781466658653. url : `http://books.google.fr/books?id=m95GAwAAQBAJ`.

[23] Virtualization Alan Murphy Technical Marketing Manager. *Keeping Your Head Above the Cloud: Seven Data Center Challenges to Consider Before Going Virtual, section Virtual Machine Deployment Challenges*. White Paper. F5 Networks, Inc. 2008.

[24] Zeus Kerravala. *Why Cloud Computing Needs a Cloud-Intelligent Network, section Cloud Computing has Arrived*. White Paper. ZK Research. Avr. 2012.

[25] *Leveraging SDN to Create Consumable, Programmable and Scalable Cloud Networks, section : Network service example*. Strategic White Paper. Alcatel-Lucent. Déc. 2013.

[26] *Effects of virtualization and cloud computing on data center networks, Section : Limitations of a hierarchical networking structure*. Technology Brief. Hewlett-Packard Development Company, L.P. Oct. 2011.

[27] *Leveraging SDN to Create Consumable, Programmable and Scalable Cloud Networks, section Data center connectivity to the wan*. Strategic White Paper. Alcatel-Lucent. Déc. 2013.

[28] Zeus Kerravala. *Why Cloud Computing Needs a Cloud-Intelligent Network, section A Cloud-Intelligent Network*. White Paper. ZK Research. Avr. 2012.

[29] R.D. Kapoor et al. *Sample netflow for network traffic data collection*. US Patent 7,193,968. Mar. 2007. url : `http://www.google.com/patents/US7193968`.

[30] Mea Wang, Baochun Li et Zongpeng Li. « sFlow: Towards resource-efficient and agile service federation in service overlay networks ». In : *Distributed Computing Systems, 2004. Proceedings. 24th International Conference on*. IEEE. 2004, p. 628–635.

[31] *Open SDN for Network Visibility, Section Monitoring in a Modern Data Center*. Solution Guide. Big Switch Networks. Juil. 2013.

[32] *Security threats to evolving data centers*. White Paper. Trend Micro. Déc. 2011.

[33] Bruno Nunes Astuto et al. *A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks*. Anglais. Section 3. SOFTWARE-DEFINED NETWORKING ARCHITECTURE. Jan. 2014.

[34] *Software-Defined Networking: The New Norm for Networks, Section Executive Summary*. White Paper. Open Networking Foundation. Avr. 2012.

[35] S. Sezer et al. « Are we ready for SDN? Implementation challenges for software-defined networks, Section 2 BACKGROUND - Why SDN? » In : *Communications Magazine, IEEE* 51.7 (juil. 2013), p. 36–43. ISSN : 0163-6804. DOI : 10.1109/MCOM.2013.6553676.

[36] *OpenFlow-enabled SDN and Network Functions Virtualization, section Introduction to NFV*. ONF Solution Brief. Open Networking Foundation. 2014.

[37] *The real-time cloud, section NETWORK FUNCTIONS VIRTUALIZATION*. White Paper. Ericsson. 2014.

[38] Tieto Intel. *Carrier Cloud Telecoms – Exploring the Challenges of Deploying Virtualisation and SDN in Telecoms Networks*. White Paper. Section Summary and Conclusions. 2013.

[39] *OpenFlow-enabled SDN and Network Functions Virtualization, section NFV AND SDN*. ONF Solution Brief. Open Networking Foundation. 2014.

[40] *The real-time cloud, section NETWORK-ENABLED CLOUD*. White Paper. Ericsson. 2014.

[41] *Network Functions Virtualisation (NFV); Use Cases, section 5 Use Case #1: Network Functions Virtualisation Infrastructure as a Service*. Group Specification. ETSI. Oct. 2013.

[42] *The real-time cloud, section USE CASES*. White Paper. Ericsson. 2014.

[43] Robert M. Keller. *Computer Science: Abstraction to Implementation, Section 1.1 The Purpose of Abstraction*. Article. Harvey Mudd College. Sept. 2001.

[44] *Information Management and Big Data A Reference Architecture*. An Oracle White Paper. Fév. 2013.

[45] *Intel's Vision of Open Cloud Computing, section Speeding Agility, Reducing Costs, and Accelerating Innovation via Cloud*. White Paper. Intel IT Center. Août 2013.

[46] Qingkui Chen, Haifeng Wang et Wei Wang. « Continuance Parallel Computation Grid Composed of Multi-Clusters. » In : *Journal of Networks* 5.1 (2010).

[47] William Tschudi et al. « High-performance data centers: A research roadmap ». In : (2004).

[48] Lee Doyle. *Where SDN and DPI technology meet: Centralized control and automation*. Web Site. 2014. URL : http://searchsdn.techtarget.com/tip/Where-SDN-and-DPI-technology-meet-Centralized-control-and-automation.

[49] Margaret Rouse. *Network Fabric*. Web Site. Mar. 2014. URL : http://searchsdn.techtarget.com/definition/network-fabric.

[50] Martin Casado et al. « Fabric: A Retrospective on Evolving SDN, section 3 Extending SDN ». In : *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*. HotSDN '12. Helsinki, Finland : ACM, 2012, p. 85–90. ISBN : 978-1-4503-1477-0. DOI : 10.1145/2342441.2342459. URL : http://doi.acm.org/10.1145/2342441.2342459.

[51] Pamela Zave. « Internet Evolution and the Role of Software Engineering ». English. In : *The Future of Software Engineering*. Sous la dir. de Sebastian Nanz. Section 3 The Real Internet et 4 Internet trends and evolution. Springer Berlin Heidelberg, 2011, p. 152–172. ISBN : 978-3-642-15186-6. DOI : 10.1007/978-3-642-15187-3_12. URL : http://dx.doi.org/10.1007/978-3-642-15187-3_12.

[52] Martin Casado et al. « Fabric: A Retrospective on Evolving SDN ». In : *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*. HotSDN '12. Helsinki, Finland : ACM, 2012, p. 85–90. ISBN : 978-1-4503-1477-0. DOI : 10.1145/2342441.2342459. URL : http://doi.acm.org/10.1145/2342441.2342459.

[53] *Open Daylight Project*. Web site. http://www.opendaylight.org/project. Avr. 2014.

[54] *Open Source Initiative*. Web site. http://opensource.org/. Avr. 2014.

[55] Nick McKeown et al. « OpenFlow: Enabling Innovation in Campus Networks ». In : *SIGCOMM Comput. Commun. Rev.* 38.2 (mar. 2008). Section 2. THE OPENFLOW SWITCH, p. 69–74. ISSN : 0146-4833. DOI : 10.1145/1355734.1355746. URL : http://doi.acm.org/10.1145/1355734.1355746.

[56] Konstantinos Kostantos et al. « Open-Source Iaas Fit For Purpose: A Comparison Between Opennebula and Openstack ». In : *International Journal of Electronic Business* 11.3 (2013), p. 191–201.

[57] D. Despotovi S. Cvetanovi et I. Mladenovi. « The concept of technological paradigm and the cyclical movements of the economy ». Anglais. In : *Facta universitatis - series: Economics and Organization* 9.2 (2012), p. 149–159. ISSN : 330.342.143.

[58] G. Dosi. *Technological paradigms and technological trajectories, Research Policy*. Anglais. 1982.

[59] Thomas Nadeau D. et Ken Gray. *SDN: Software Defined Networks*. 1st. Chapitre 2 - Centralized and Distributed Control and Data Planes. O'Reilly Media, Inc., 2013. isbn : 1449342302, 9781449342302.

[60] André B. Bondi. « Characteristics of Scalability and Their Impact on Performance ». In : *Proceedings of the 2Nd International Workshop on Software and Performance*. WOSP '00. Ottawa, Ontario, Canada : ACM, 2000, p. 195–203. isbn : 1-58113-195-X. doi : 10.1145/350391.350432. url : http://doi.acm.org/10.1145/350391.350432.

[61] René J CHEVANCE. « Serveurs multiprocesseurs et SGBD parallélisés ». In : *Techniques de l'ingénieur. Informatique* H2068 (2001), H2068–1.

[62] *Simple Network Management Protocol (SNMP)*. Web site. Juin 2014. url : http://www.net-snmp.org/.

SDN : Software-Defined Networking

rédigé par Cynthia Lopes do Sacramento

Résumé

De récentes technologies et concepts émergent pour répondre aux nouvelles utilisations des réseaux et internet. Les data centres évoluent vers un nouveau mode de livraison, le Cloud Computing, pour pouvoir s'adapter à cette situation et fournir des services à la demande et de manière agile. La virtualisation des serveurs a été une première étape dans cet objectif, mais elle doit être suivie de la virtualisation du Stockage et du Réseau. Des solutions de stockage virtuel se développent et sont disponibles sur le marché, mais l'aspect réseau reste un facteur bloquant pour le passage au Cloud Computing. Une évolution technologique est donc attendue dans le domaine des réseaux informatiques. Ce qui a mobilisé la communauté dans les projets de recherche sur les réseaux programmables, dont un des sujets est l'objet de cette étude : SDN - Réseaux Informatiques Définis par Logiciel. SDN propose une nouvelle architecture plus agile, facile à gérer, rentable et flexible. Cette architecture sépare le plan de contrôle (intelligence et état du réseau) du plan de données (fonctions de transmission). L'approche permet de rendre le contrôle directement programmable avec l'infrastructure sous-jacente abstraite aux applications réseaux et services. Cette étude présentera les apports de SDN dans le cadre des data centres en réponse aux besoins de sécurité et agilité exigés pour le passage au Cloud Computing.

Mots clés : SDN, Data Centre, Virtualisation, Cloud Computing, Agilité, Sécurité

Abstract

New technologies and concepts appear in response to new network and internet usage requirements. Data Centers are moving torwards a new delivery model to be able to handle this new context and fastly offer on-demand services. Servers virtualization has been a first step in this direction, but the same advancement must be applied to the Storage and to the Networks. While virtual storage solutions have been developped and are already availabe on the market, networks remain blocking the Cloud Computing enablement. A technological evolution is then expected for the computer networks. As result the research community has produced many works on programmable networks. Among them, the subject of this study : SDN - Software Defined Networking. SDN proposes a new architecture plus dynamic, ease to manage, profitable and flexibile. This architecture decouples the control plane (network intelligence and state) from the data plane (transmission functions). This approach makes de the control directly programmable and causes the underlying infrastructure to be abstracted to network applications and services. This study is going to present SDN contributions on data centers so as to respond to the security and agility challenges of the Cloud Computing.

Keywords : SDN, Data Center, Virtualization, Cloud Computing, Agility, Security