ACI, la solution SDN du datacentre de la Communauté Université Grenoble Alpes

Julien Tomassone, Rémi Cailletaud, Patrick Fulconis, Patrice Navarro, Christian Seguy

ACI, la solution SDN du datacentre de la Communauté Université Grenoble Alpes

Julien Tomassone
DGDSI UGA
41 rue des Mathématiques
38401 Saint Martin d'Hères

Rémi Cailletaud
3SR - Laboratoire Sols, Solides, Structures - Risques
BP53, 38041 Grenoble CEDEX 9

Patrick Fulconis
Laboratoire VERIMAG
Université Grenoble Alpes
700, avenue centrale
38401 Saint Martin d'Hères

Patrice Navarro
Laboratoire LJK
Université Grenoble Alpes
700, avenue centrale
38401 Saint Martin d'Hères

Christian Seguy
Laboratoire LIG
Université Grenoble Alpes
700, avenue centrale
38401 Saint Martin d'Hères

Résumé

Depuis quelques années le, Sofware-Defined Networking (SDN), fait parler de lui. On le présente comme la solution d'avenir pour le déploiement des infrastructures réseaux, plus souples, évolutives,…

Cette technologie permet de gérer les équipements de façon centralisée et automatisée en ajoutant un niveau d'abstraction à l'infrastructure. La séparation des parties décisionnelle et opérationnelle autorise une grande souplesse dans le déploiement, l'évolution et l'automatisation au travers d'API.

C'est le choix que l'Université de Grenoble Alpes (UGA) a fait. Dans un contexte de fusion des universités où l'agrégation des salles informatiques au sein d'un seul datacentre était une condition de réussite, l'UGA a déployé une nouvelle infrastructure en s'appuyant sur l'offre CISCO ACI (Application Centric Infrastructure).

Cette présentation s'attachera à développer les grands concepts du SDN et détaillera la solution que nous avons retenue.

Nous expliquerons pourquoi nous avons fait le choix de CISCO ACI plutôt qu'une autre solution SDN et nous aborderons les difficultés rencontrées ainsi que les bénéfices liés à son adoption : nous nous attarderons sur les défauts de jeunesse de la solution au moment du déploiement, les bugs rencontrés, le manque de ressources documentaires mais aussi sur la souplesse d'utilisation, les performances, etc.

Mots-clefs

SDN, ACI, CISCO, Fabric IP, Leaf, Spine, SPRING, ...

1 Introduction

Depuis plusieurs années, les solutions SDN (*Software-Defined Networking*) sont présentées comme l'avenir du déploiement d'infrastructures réseaux, plus souples, plus évolutives. Cependant ces déploiements restent rares et souvent peu présentés. Si la mise en place d'une telle infrastructure sur un nouveau site s'appréhende relativement facilement, migrer d'une solution classique vers du SDN pose plus de problèmes, notamment afin d'éviter les perturbations et les coupures de services. Il faut donc que le contexte s'y prête, mais aussi que les gains obtenus soient à la hauteur de l'investissement financier et humain.

C'était notre cas : une fusion des universités, la volonté de regrouper les salles informatiques ainsi que la mise en service d'un nouveau datacentre ont été les conditions de réussite de ce projet ambitieux. La Communauté d'Universités et Établissements Université Grenoble Alpes (COMUE UGA) a déployé sa nouvelle infrastructure en s'appuyant sur l'offre CISCO *Application Centric Infrastructure* (ACI) ainsi qu'une partie d'infrastructures classiques (*legacy*).

2 Contexte

La COMUE UGA couvre un large périmètre, elle regroupe des entités du CNRS, d'Inria, de l'Institut Polytechnique de Grenoble (Grenoble INP) et de l'Université Grenoble Alpes (UGA), elle-même issue de la fusion de 3 universités au 1er janvier 2016. La COMUE UGA est forte de 62 000 étudiants, 3 700 doctorants, 7 000 personnels, 30 facultés, écoles ou instituts et plus de 80 laboratoires. Elle s'étend sur un périmètre géographique très large puisqu'elle comprend 12 sites répartis sur 6 départements.

Dans ce cadre, une Unité Mixte de Service, l'UMS Grenoble Infrastructure de CAlcul et de Données (GRICAD) a été créée au 1er janvier 2016 pour répondre aux enjeux des besoins scientifiques actuels en matière de calcul intensif et de données. Elle a pour tutelles le CNRS, l'UGA, et Grenoble-INP. Elle est porteuse des projets d'infrastructures numériques mutualisées du site grenoblois, en terme de coordination et de pilotage.

Les différentes unités de recherche mais aussi les UFRs, écoles et services centraux sont associés en vue de promouvoir, développer et gérer les datacentres et plateformes mutualisés du site.

GRICAD mutualise les compétences et les expertises présentes selon un modèle organisationnel original : des personnels rattachés à l'UMS, aux unités de recherche, aux services centraux ou aux composantes d'enseignement, investissent une partie de leur temps dans les projets soutenus par GRICAD. Le comité technique SPRING (Service Partagé des Réseaux Innovants de Grenoble) s'inscrit dans ce cadre, fort de 15 personnes issues de différentes entités (laboratoires, UFR et services centraux, ...), SPRING s'occupe de la conception, du déploiement et de la gestion de la solution d'infrastructure réseau du DC Grenoblois.

Cette organisation favorise la proximité entre les différentes équipes et permet une acquisition de compétences plus rapide des personnels impliqués dans les projets.

3 Périmètre technique

Précédemment, chaque établissement opérait sa propre infrastructure réseau et/ou sécurité séparée de celles des autres établissements. Certains services étaient hébergés en commun, comme les applicatifs métiers.

La fusion de plusieurs établissements peut être l'occasion de repenser l'organisation de l'infrastructure informatique. Les conséquences sont certes importantes, mais le bénéfice peut l'être tout autant, que ce soit en termes humain, organisationnel ou technique.

La COMUE UGA n'a pas hésité, et même si un investissement non négligeable est demandé, cela a permis la mise en place de plateformes communes, adaptées aux besoins des établissements et de leurs

services.

Une démarche de rationalisation des salles machines a été lancée avant la fusion et a permis de définir une nouvelle stratégie basée sur la création d'un datacentre composé de plusieurs salles. Fort de ces choix, la DSI UGA et l'UMS GRICAD ont imaginé et réfléchi à ce que pourrait être une nouvelle infrastructure informatique mutualisée et innovante sur le pôle grenoblois.

4 Choix de la solution

La technologie du SDN permet de gérer les équipements de façon centralisée et automatisée en ajoutant un niveau d'abstraction à l'infrastructure. La séparation des parties décisionnelle et opérationnelle autorise une grande souplesse dans le déploiement, l'évolution et l'automatisation au travers d'API. L'émergence de l'offre de services du type des réseaux SDN, les possibilités d'automatisation avancées offertes par ces services, mais aussi les infrastructures de type « Fabric IP » sur lesquelles elles reposent, nous ont fortement orientés dans notre choix.

Les principaux critères de choix de la nouvelle solution reposaient sur :

- une infrastructure réseau répartie (au moins sur trois salles), mais pilotable depuis un point central ;
- des équipements performants (10G/40G avec une possibilité d'évolution en 100G), redondants et permettant des mises à jour sans coupure ;
- la capacité d'héberger des machines virtuelles aussi bien que des machines physiques ;
- des possibilités d'automatisation poussées ;
- la mise en place d'une délégation fine sur l'accès à l'interface (ou aux interfaces) de gestion.

À l'été 2015 on ne trouvait sur le marché du SDN que 3 acteurs principaux : VMware, Juniper et CISCO. Nous avons organisé une présentation par un partenaire des solutions proposées par ces trois fabricants afin de déterminer laquelle pourrait rendre au mieux le service attendu.

- NSX, la solution de VMware, était trop limitative. En effet il n'était possible de gérer que les environnements de virtualisation VMware. Notre communauté utilisant des serveurs de type *baremetal* et d'autres types d'hyperviseur comme Xen ou Proxmox, nous n'avons donc pas retenu cette solution ;
- Contrail de Juniper ne proposait pas de solution unique pour la gestion de serveurs *baremetal* et virtuels au sein d'une même infrastructure ;
- ACI, la solution de CISCO, répondait au plus grand nombre de critères, notamment celui de pouvoir opérer et gérer des machines virtuelles de la même manière que des machines physiques.

Si NSX a été écartée rapidement, le choix entre les deux autres solutions a été plus difficile. Il s'est finalement porté sur la solution de CISCO qui devait nous permettre de déployer notre infrastructure réseau sur le campus Grenoblois à travers l'interconnexion de trois salles machines donnant la vision unique de tous les équipements du datacentre avec une gestion centralisée et sécurisée.

5 Rappel sur ce qu'est le SDN

Un rapide rappel sur le SDN s'impose même si depuis dix ans il fait partie des sujets phares du monde informatique et a été plusieurs fois le sujet de présentations lors des derniers JRES.

La définition la plus communément admise est la suivante : le *Software-Defined Networking* (SDN) se caractérise par la séparation physique du plan de contrôle et du plan de données, ainsi que la centralisation des fonctions de contrôle.

Le SDN désigne un ensemble de technologies innovantes visant à permettre un contrôle centralisé des ressources réseau, une meilleure programmabilité et une orchestration de ces ressources, ainsi que leur virtualisation en les dissociant des éléments physiques du réseau.

La partie décisionnelle des équipements est séparée de leur partie opérationnelle et déportée vers un unique point de contrôle.

Avec le SDN, le logiciel communique avec le matériel et permet de contrôler le réseau et ses terminaux physiques.

Le SDN a pour but de rendre les réseaux programmables par le biais d'un contrôleur centralisé.

C'est essentiellement une solution de gestion de nouvelle génération qui crée un réseau provisionné de manière dynamique, évolutif et programmable.

Le SDN permet :

- de centraliser la gestion du réseau ;
- de contrôler les routeurs et les commutateurs ;
- d'appliquer des règles à grande échelle et transférer les trames et les paquets de manière dynamique.

Cette abstraction à travers une API standard permet un développement de services réseaux à forte valeur ajoutée (équilibrage de charge, automatisation de configuration, planification, routage intelligent,...) affranchis des spécificités des équipementiers.

Avec le SDN, les fonctions d'orchestration, de gestion et d'automatisation relèvent des contrôleurs SDN.

Un contrôleur centralisé est chargé du pilotage de l'intégralité des équipements réseaux, qui sont pour leur part responsables du transfert des paquets d'une interface à une autre en fonction des règles établies.

Pour ce faire, le SDN s'appuie sur divers protocoles dont le plus connu est OpenFlow. Ces protocoles permettent la communication entre les contrôleurs et les équipements réseaux.

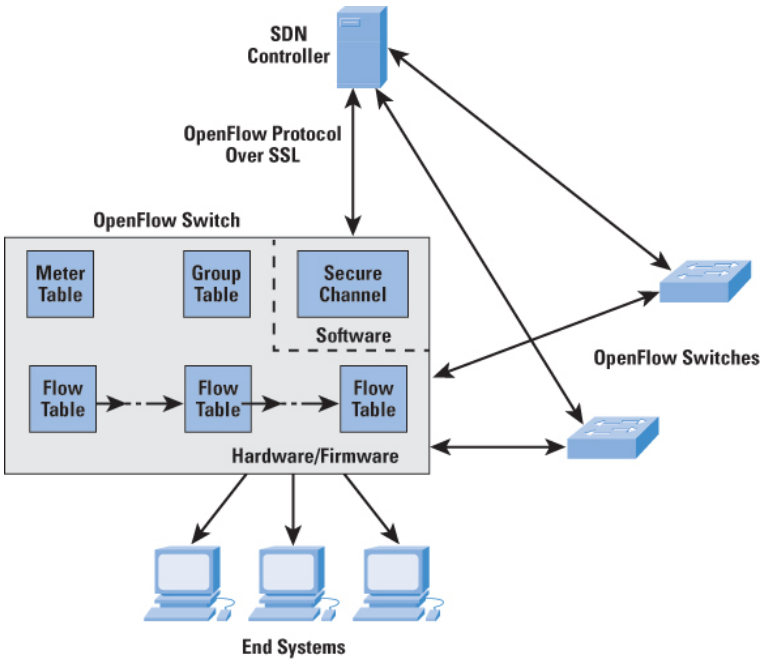

Figure 1 - OpenFlow Switches[1]

La figure ci-dessus illustre le routage de paquets entre deux réseaux distants via une implémentation de SDN avec le protocole OpenFlow.

Dans une approche SDN, la charge de calcul associée au contrôle est en grande partie retirée des routeurs.

Les routeurs discriminent les paquets et déterminent leur interface de sortie, mais ce comportement découle des règles émises par le contrôleur.

Dans le mode réactif, le routeur qui réceptionne le paquet de l'envoi signale l'événement au contrôleur et

1. www.cisco.com - Software-Defined Networks and OpenFlow - The Internet Protocol Journal, Volume 16, No. 1 - by William Stallings

reçoit en retour des règles au cours d'un échange, afin de décider du transfert vers le routeur suivant approprié.

Un comportement proactif consisterait en la transmission de règles avant que le routeur ne reçoive les paquets associés.

Les routeurs effectuent ainsi essentiellement des fonctions de commutation d'où la désignation de « switches OpenFlow ».

L'OpenFlow identifie les flux spécifiques en utilisant une variété de critères (adresse MAC, adresse IP de destination, etc), puis effectue des actions sur ces flux (forwarding via le port X ou Y, abandon du trafic, etc).

Un contrôleur OpenFlow centralisé ayant connaissance de l'ensemble de la topologie du réseau peut programmer ces politiques pour tous les commutateurs de réseau, quel que soit le constructeur.

6 Présentation technique

Le principal objectif de l'ACI est de fournir une solution souple et performante pour les datacentres, en particulier pour le trafic est-ouest (trafic interne à la *Fabric,* à la différence du trafic nord-sud qui va de l'extérieur à l'intérieur). Dans cette partie, nous présentons rapidement l'architecture de la solution, et le modèle de configuration logique. Nous n'entrerons pas dans les détails de la configuration matérielle, car elle n'apporte rien à la connaissance d'une architecture SDN.

6.1 Architecture matérielle

La solution ACI se base d'une part sur des contrôleurs *Application Policy Infrastructure Controller* (APIC) redondants, qui gèrent un point unique de provisionnement, configuration, déploiement et surveillance, et d'autre part sur une infrastructure de type *Fabric leaf/spine* (voir partie 2) constituée de switches Nexus 9000 Series. L'ensemble des équipements de la Fabric sont raccordés par des liens à 40G pouvant évoluer à 100G.

À la différence d'une solution SDN classique, les contrôleurs APIC de l'ACI ne traitent aucun trafic, et ne servent qu'à pousser les configurations sur les équipements. Ils peuvent donc être temporairement déconnectés.

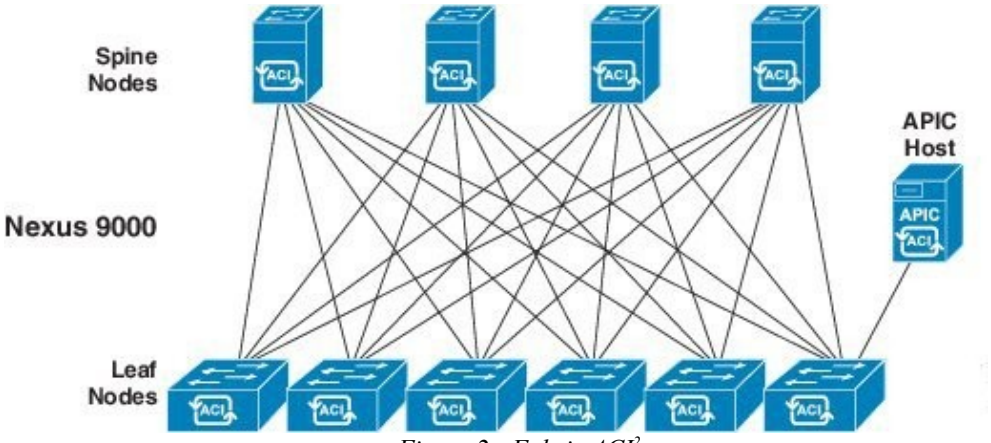

Figure 2 - Fabric ACI[2]

6.2 Déploiement et auto-configuration

Le déploiement de l'infrastructure est relativement aisé, grâce à ses *capacités* d'auto-configuration : les nouveaux équipements sont détectés et ajoutés automatiquement à l'infrastructure. Pour cela, l'ACI utilise

2. www.cisco.com – ACI Fundamentals - Cisco Application Centric Infrastructure Fabric Overview

les protocoles suivants en interne : *Dynamic Host Configuration Protocol* (DHCP), *Intermediate System-to-Intermediate System Protocol* (IS-IS) et *Link Layer Discovery Protocol* (LLDP). Pour les mises à jours, l'APIC fournit une solution centralisée permettant de gérer les images systèmes des équipements, ainsi que leur déploiement.

6.3 Traffic forwarding

Dans une architecture classique, on utilise les VLAN pour segmenter le réseau. Cette solution est généralement contraignante, et la limitation à 4094 VLAN pose rapidement problème.

L'ACI, grâce à son architecture *spine/leaf* et à l'utilisation des *VxLAN tunnel endpoints* (VTEP), apparaît au monde extérieur comme un seul et unique switch. Quand le trafic entre dans la Fabric, il est encapsulé. Il est ensuite transmis au travers de la Fabric en deux sauts au maximum (stratégie *Equal Cost Multi Path* ou ECMP). Il est finalement désencapsulé à sa sortie.

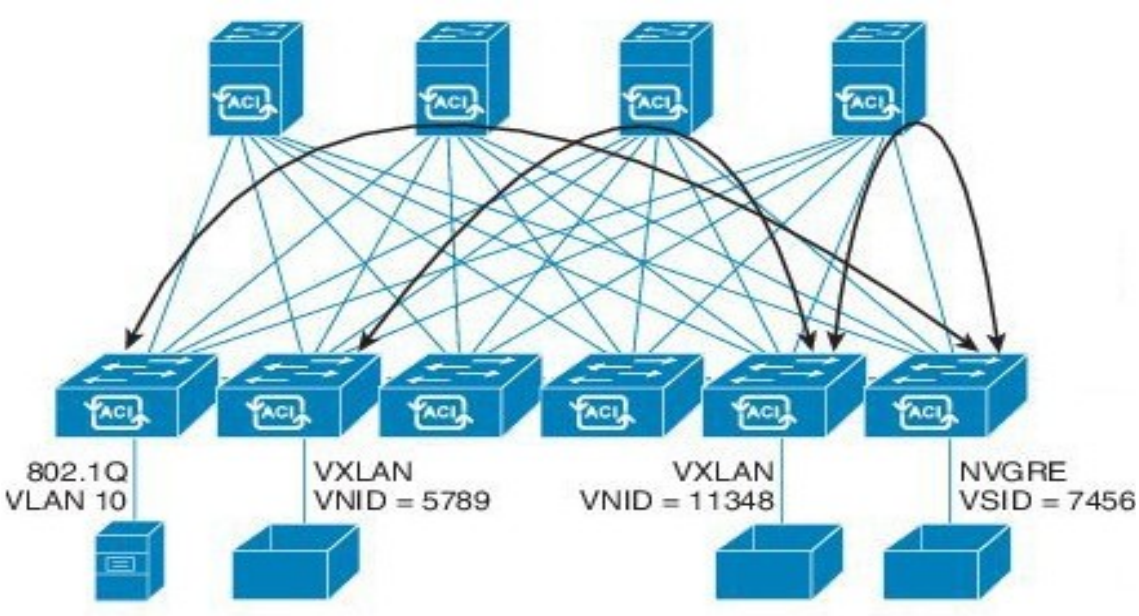

Figure 3 - Encapsulation VTEP[3]

L'ACI utilise pour la communication point-à-point IS-IS et *Council of Oracle Protocol* (COOP ; un protocole qui utilise Zero Message Queue pour maintenir les informations de cartographie du réseau (identité et localisation) au niveau des spines), et *Multiprotocol Border Gateway Protocol* (MP-BGP) pour la gestion des informations de routage externe.

3. www.cisco.com – ACI Fundamentals – Forwarding within the ACI Fabric

6.4 Modèle de configuration

Toute la configuration se fait au niveau de l'APIC par son API REST, soit en utilisant l'interface web, soit en ligne de commandes. La figure ci-dessous présente le modèle de configuration.

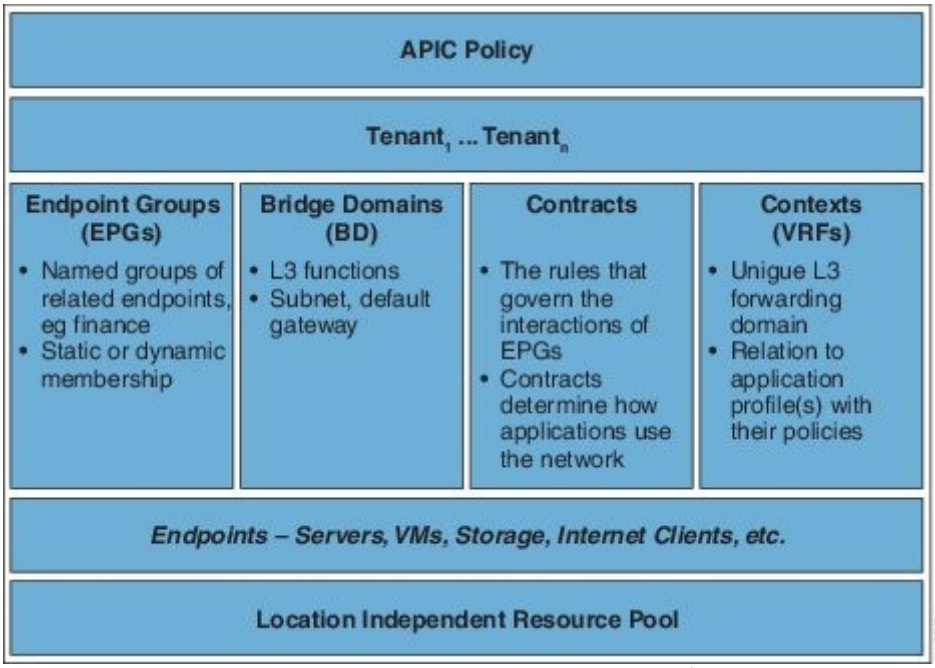

Figure 4 - Modèle de configuration[4]

Les administrateurs de la Fabric configurent l'aspect infrastructure et créent les tenants dont ils délèguent la gestion. De leur coté, les administrateurs de tenants gèrent l'aspect applicatif, ce qui leur permet de penser en terme d'application, et non en terme de réseau.

6.5 Tenants

Ils représentent une unité de contrôle d'accès logique (*domain-based access control*), typiquement une entité (service, laboratoire), un projet, ou une configuration particulière. Les tenants sont par défaut isolés les uns des autres, mais peuvent partager leurs ressources. Ils contiennent des filtres (ports/protocoles), des contrats, des réseaux externes, des *Bridges Domains*, des *Virtual Routing and Forwarding* (VRF, ou contextes) et des profils d'applications, qui eux-même contiennent les *EndPoint Groups* (EPG).

4. www.cisco.com – ACI Fundamentals – ACI Policy Model

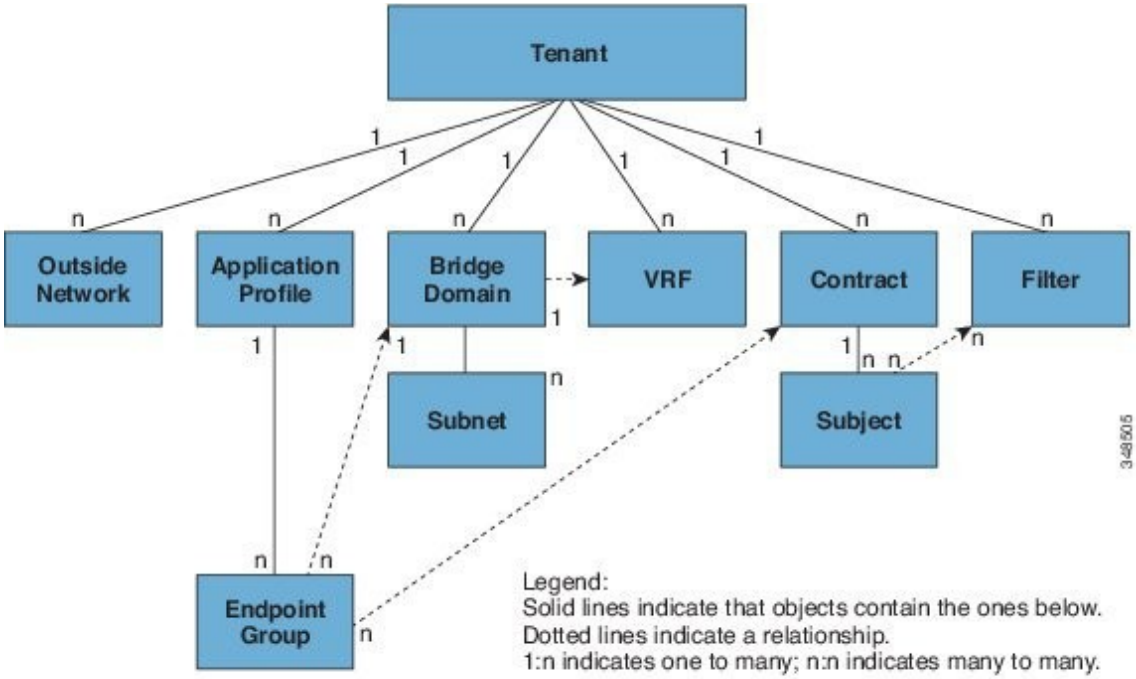

Figure 5 - Tenants[5]

6.6 Virtual Routing and Forwarding, Bridge Domain, Subnet

Un tenant peut avoir plusieurs VRF. Les VRF sont des «contextes réseaux», ou plus schématiquement des routeurs virtuels au sein de la « Fabric IP ».

Les Bridges Domain (BD), qui représentent la couche de transfert L2, sont liés aux VRF. Ils définissent un domaine de diffusion et un espace d'adressage (à la manière d'un vlan, cet espace peut contenir un ou plusieurs sous-réseaux). Les sous-réseaux peuvent être déclarés dans un ou plusieurs BD, et être publics (visibles depuis l'extérieur), privés (limités au tenant), ou partagés (entre les VRF, même de différents tenants). Évidemment, ils ne peuvent pas être annoncés publics et partagés dans différents VRF.

[Voir figure 4 *Modèle de configuration*]

6.7 Applications Profiles, EndPoint Group

Les *Applications Profiles* (AP) permettent une organisation logique des *EndPoint Groups* (EPG).

Les EPG sont les objets les plus importants du modèle de configuration : ils contiennent un ensemble d'EndPoints, qui sont en fait les périphériques connectées. Cela peut être un serveur, une machine virtuelle, un NAS… L'appartenance d'un EndPoint à un EPG peut être définie statiquement ou dynamiquement. L'ensemble des EndPoints d'un EPG partage la même configuration, en particulier en terme de sécurité.

Typiquement, une application web simple se compose de deux services (web et base de données). L'AP pour cette application contiendra alors deux EPG, l'un pour le serveur web et l'autre pour le serveur de base de données.

Un des types particuliers d'EPG est « l'external network » qui est composé d'un ou plusieurs subnets. Ces « external network » sont associés à une « L3out » qui établit les communications entre l'ACI et l'extérieur au travers de protocoles de routage (OSPF, BGP, EIGRP, …).

5. www.cisco.com – ACI Fundamentals – ACI Policy Model

6.8 Contrats

Le deuxième objet clé du modèle de configuration est le contrat. Les EPG ne peuvent communiquer les uns avec les autres que si au moins un contrat les y autorise. Le contrat spécifie le type de trafic (protocoles, ports). En l'absence de contrat, la communication inter-EPG est impossible. En revanche, la communication intra-EPG est autorisée par défaut.

6.9 Gestion des systèmes de virtualisation

Le gestionnaire de système de virtualisation de l'ACI permet de configurer facilement les politiques réseaux des machines virtuelles. Hyper-V, Openstack et VMWare sont actuellement supportés. Dans le cas d'Openstack par exemple, c'est le plugin ACI du système de virtualisation qui utilise l'API de l'ACI pour la configurer automatiquement. Dans le cas de VMWare, c'est l'ACI qui pilote le réseau côté VCenter. Cette intégration permet la mobilité des VM et leur placement automatique dans les bons EPG.

7 Notre infrastructure et nos choix techniques

7.1 Architecture matérielle

Avec un DC composé de trois salles dans un premier temps et une évolution probable vers quatre salles ou plus encore, il nous fallait prévoir une architecture suffisamment souple pour permettre une évolution, tout en étant la plus simple et robuste possible.

Lors du maquettage de l'ACI un seul mode de distribution multi-DC était disponible, le mode *Stretched Fabric*. Ce mode permet d'opérer une Fabric déployée sur une ou plusieurs salles (trois maximum) et de la gérer via un seul point d'entrée : le contrôleur.

S'il permettait de répondre à nos besoins à un instant T, d'autres solutions sont à l'étude, notamment via une évolution de l'ACI (ex : une évolution du nombre de salles est envisagée).

Une fois le contexte, le périmètre et la solution définis, il nous restait l'étape de la mise en application.

7.2 La mise en place de notre première salle

Dans un premier temps, afin de nous familiariser avec les concepts du SDN en général et d'ACI en particulier, nous avons commencé à déployer l'architecture sur une salle (DC3).

Celle-ci est composée de 34 baies d'hébergement standard, trois baies dédiées à la gestion du bâtiment et une baie séparée dans un local à hygrométrie contrôlée.

Après étude des équipements à héberger, nous nous sommes orientés vers la solution suivante :

- deux *spines* 36 ports 40G QSFP+ sur 2U pour la concentration ;
- deux *leaves* dédiées à la connexion au reste du monde pourvues de 32 ports 40G QSFP+ de desserte et 6 ports 40G QSFP+ de remontée vers les *spines* ;
- une *leaf* dédiée à l'hébergement pour chaque paire de baies, dotée de 48 ports 10G SFP+ de desserte et 6 ports 40G QSFP+ de remontée vers les *spines* ;
- chaque *leaf* d'hébergement est connectée à un *switch* 48 ports 1G cuivre de desserte et 2 ports 10G SFP+ ;
- un des ports 10G SFP+ des *switches* est utilisé pour se connecter à sa *leaf*, le second est utilisé pour véhiculer un réseau de gestion « out of band » géré par un firewall dédié autonome.

Dans un second temps, nous avons étendu notre « Fabric IP » à deux autres salles (DC1 et DC2). Dans chacune de ces dernières, nous avons mis en place deux *spines*, deux *leaves* et deux *switches*.

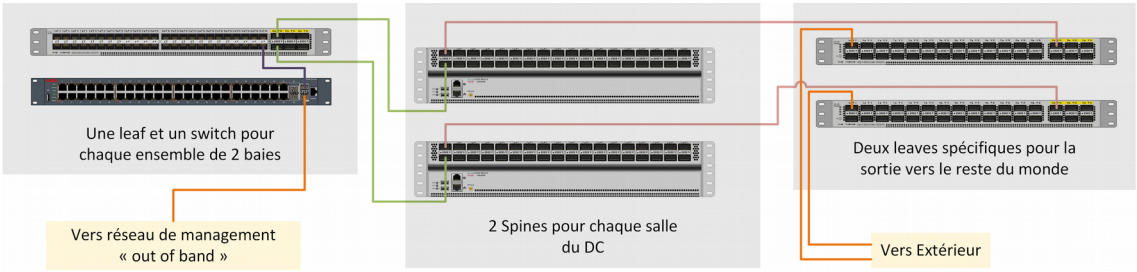

Figure 6 - Exemple d'infra de salle du DC

Nous avons ensuite réalisé l'extension - stretch - de la « Fabric IP » originale selon le schéma suivant :

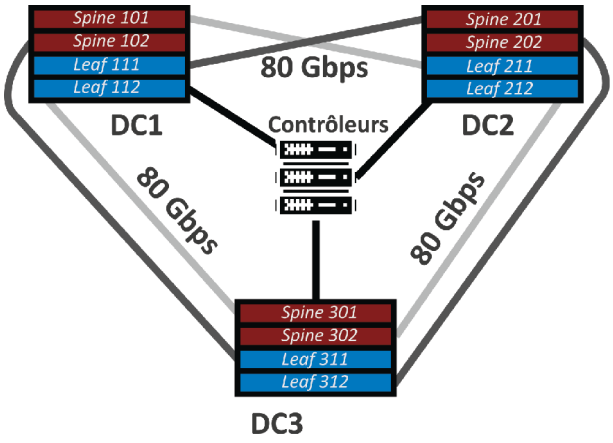

Figure 7 - Schéma des 3 salles du datacentre

Enfin, pour sécuriser les infrastructures et améliorer les performances, nous avons mis en place différentes sorties vers le reste du monde dans chaque salle.

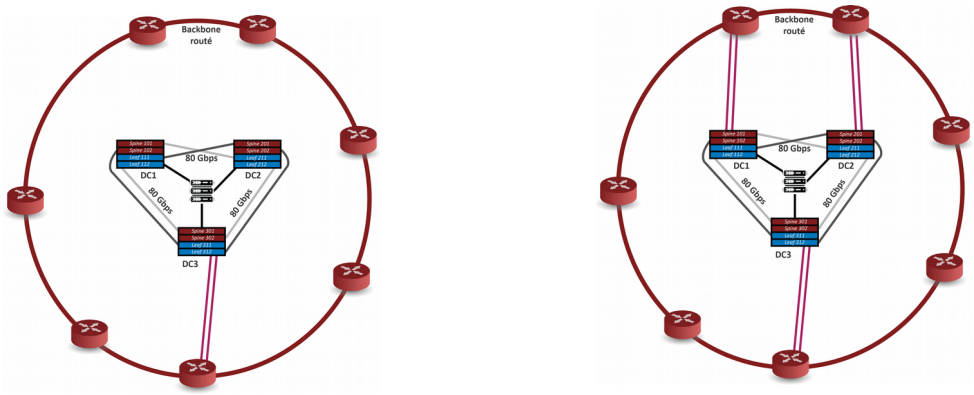

Figure 8 - Une seule connexion initialement, puis redondance avec une connexion par salle

En terme d'interconnexions, l'infrastructure ACI de SPRING est donc reliée au backbone de l'UGA en différents points et utilise divers protocoles pour échanger les routes de ce dernier.

Si le protocole utilisé par le backbone est OSPF et que nous l'avons utilisé directement au départ en intégrant l'ACI à nos aires existantes, l'utilisation de multiples routeurs virtuels à l'intérieur de la Fabric ainsi que le nombre conséquent de routes gérées nous a contraint à revoir notre copie. En effet, nous

avons rapidement saturé les tables de routages des *leaves* d'interconnexion, nous préparons donc une évolution dans les mois qui viennent qui nous permettra de mieux gérer nos échanges vers le reste du monde.

À plusieurs reprises, nous avons modifié la structure de nos interconnexions, l'ACI proposant à ce niveau une souplesse très appréciable.

7.3 Architecture logicielle

Il s'agit sans doute de la partie qui a le plus évolué, autant dans la période de pré-production que depuis le démarrage de l'exploitation.

Si la définition d'une structure logique pour une entité (service, UFR, laboratoire,..) est relativement simple, il en va tout autrement si l'on souhaite gérer finement plusieurs dizaines d'entre elles.

Nous avons pu échanger sur ce sujet à plusieurs reprises avec Cisco et avons affiné nos choix progressivement.

Au départ nous avons choisi d'utiliser un seul tenant contenant différents routeurs virtuels (VRF) communiquant tous vers l'extérieur par une seule et même sortie (*common L3out*).

Afin de limiter l'impact de modifications d'une entité sur les autres, la première évolution a consisté à déployer un tenant par entité hébergée contenant une ou plusieurs VRF communiquant toujours avec l'extérieur par une seule et même sortie (common L3out).

Toujours pour la même raison, une seconde évolution a concerné la communication externe des tenants en leur affectant une L3out dédiée.

La suivante nous a amené à n'utiliser qu'une seule VRF par tenant, tout en conservant ces L3out dédiées ce qui simplifie le routage interne.

Et enfin nous envisageons de changer le protocole de routage vers l'extérieur en basculant d'OSPF vers BGP afin de mieux maîtriser les imports/exports de routes entre la Fabric et le backbone.

En résumé :

Un tenant avec plusieurs VRF et sorties par L3out common

puis

Un tenant par entité avec une ou plusieurs VRF et une sortie par L3out common

puis

Un tenant par entité avec une ou plusieurs VRF et une sortie par L3out dédiée

puis

Un tenant par entité avec une seule VRF et une sortie par L3out dédiée

Toutes ces évolutions ont pu être réalisées assez simplement sans perturbation majeure sur la production.

8 Bilan

Après le choix de la solution CISCO ACI, essentiellement pour l'intégration de nos serveurs *baremetal* et de la virtualisation non VMware, les débuts ont été difficiles. Il a fallu assimiler les notions de SDN, tant en termes de technique, de vocabulaire que de philosophie, car les concepts sont très différents des réseaux classiques et peuvent paraître complexes au premier abord. Si cela est particulièrement motivant, l'investissement humain nécessaire n'est pas à négliger. Ceci a pris beaucoup de temps, et surtout demandé de la régularité dans l'utilisation de la solution (on oublie beaucoup et vite si on ne pratique pas régulièrement).

De plus, nous avons commencé (fin 2015) à expérimenter puis exploiter une solution qui présentait encore quelques défauts de jeunesse : quelques bugs handicapants, documentation pas toujours aboutie,

interface graphique lente. Certaines mises à jour ont aussi pu être la source de problèmes.

Malgré tout, l'interface graphique reste peu conviviale et présente des limitations de fonctionnalités (exemple : renommage d'objets impossible).

Actuellement notre solution comprend 3 salles, 10 tenants, 6 *Spines,* 30+ *Leaves*, 30 *Switches* L2, 1500+ ports 10G SFP+, 64 ports 40G QSFP+, 1300+ ports 1G cuivre, 1000+ EndPoints, 150+ EPG, 60+ BD.

L'exploitation d'une telle solution est parfois frustrante, car on a le sentiment de n'utiliser qu'une petite partie de son potentiel. Nos applications sont souvent monolithiques, et nous ne pouvons pas tirer avantage de la programmabilité et de l'automatisation qu'offre ce type de solution.

La mise en place et l'utilisation d'une telle solution SDN incite à réfléchir en termes d'applications et services.

Une fois la complexité initiale appréhendée, l'ACI apporte de nombreux avantages :

- deux *spines* 36 ports 40G QSFP+ sur 2U pour la concentration ;
- performances (temps de latence, débits observés) excellentes ;
- intégration de nouveaux équipements facile et quasi automatique ;
- association simple avec les plateformes de virtualisation et mobilité des VM ;
- souplesse de gestion des services (web, bases de données, ...).

Pour une question de budget (intégration de firewall L4-L7 CISCO dans l'ACI) nous n'avons pas pu atteindre le niveau de sécurité de notre infrastructure existante. En effet, d'une part l'ACI permet d'appliquer une politique de filtrage uniquement sur un EPG et non sur les éléments le composant ; d'autre part le contrôle des flux se fait en mode « *stateless* » et non « *stateful* » comme le ferait un *firewall*. Nous avons donc dû ajouter un équipement externe pour permettre la mise en place de cette politique de filtrage.

Un des constats que nous avons pu faire après la mise en production de l'ACI est l'importance de la prise en compte de l'infrastructure existante dès la phase de définition de l'architecture.

En effet, dans le cas d'une imbrication forte des deux, il est préférable d'adapter l'ACI à l'existant plutôt que l'inverse.

Le fonctionnement du comité technique SPRING, composé d'une douzaine de personnes issues de structures différentes et travaillant entre 10 et 20% de leur temps sur ce projet, est très bénéfique. Il permet la montée en compétence des personnes, apporte une grande motivation et crée une dynamique au sein du groupe.

9 PERSPECTIVE

Maintenant que la mise en production est terminée, nous souhaitons améliorer notre solution en développant des APIs permettant d'automatiser les créations/modifications d'objets.

Nous souhaitons aussi développer des interfaces et configurer les droits afin de donner accès aux utilisateurs, au moins en lecture.

L'utilisation plus fréquente du mode CLI nous permettra de nous soustraire des contraintes de l'interface graphique.

Enfin, notre infrastructure va évoluer avec l'intégration prochaine d'une quatrième salle, ce qui nous demandera de passer dans le nouveau mode ACI « multi-POD » (au lieu du « Stretched Fabric » actuel).

Annexe

LEXIQUE :

Fabric/fabric/Fabric IP : Ensemble de leaves et de spines ou toutes les spines sont connectées à toutes les leaves.

Leaves : Pluriel de leaf

Spine : Equipement réseau composant une Fabric IP et agglomérant les flux. Connecté directement aux leaves.

Leaf : Equipement réseau composant une Fabric IP. Connecté directement aux spines, aux switches ou routeurs externes et aux équipements offrant des services.

Switches : pluriel de switch