Software-defined network dans Proxmox

Ryan Lahfa, École normale supérieure de la rue d'Ulm

Qui suis-je ?

Qui suis-je ?

Ryan Lahfa (ou Raito Bezarius sur Internet), responsable du club réseau à l'ENS Ulm.

- 8 ans d'expérience professionnelle en informatique[1]
- Étudiant en 2A à l'ENS Ulm au département d'informatique et de mathématiques
- « Entrepreneur d'Intérêt Général » au ministère de la Justice sur la numérisation de la chaîne pénale
- Professionnel du logiciel libre
- Passe son temps libre à numériser les mathématiques dans l'assistant de preuve Lean

Et plus encore, sur https://ryan.lahfa.xyz/about-me.html :-).

[1] Des FPGAs au développement web avec React.js, en passant par l'administration système avec NixOS ou la sécurité informatique bas-niveau.

SDN c'est quoi ?

Un détour par la question de reconfiguration

Les ordinateurs sont connus pour être très reconfigurables, on peut:

- « reconfigurer » son navigateur, ses applications, son système d'exploitation ;
- « reconfigurer » la disposition de nos périphériques, e.g. PCIe, batteries, RAM, CPU ;
- « reconfigurer » une partie du CPU, e.g. activer les extensions de virtualisation.

Visuellement: https://appro.mit.jyu.fi/tools/biossimu/simu.html

Qu'est ce qu'on ne peut pas reconfigurer alors ?

Ce qui est difficilement reconfigurable sur une machine lambda

- « reconfigurer » **intégralement** le jeu d'instructions de son processeur, e.g. passer d'un processeur x86_64 à RISC-V sans le changer[2] ;
- « reconfigurer » un périphérique PCIe qui est utilisé comme disque NVMe en modem 4G[3] ;
- « reconfigurer » Intel Management Engine pour le désactiver[4] ;
- « reconfigurer » Intel BootGuard pour avoir ses propres clefs

[2] Mais on peut utiliser un FPGA pour ça !
[3] On peut mais seulement avec un BIOS en état manufacturer!
[4] Bon, il y a le HAP bit... Mais c'est compliqué™.

Mézalor, quel rapport avec la choucroute ?

Les switch, les routeurs, les serveurs, ce sont aussi des machines qui ont des contraintes de reconfiguration.

- Reconfigurer les VLANs, les ACL de ports, les rate-limits, c'est souhaitable!
- Reconfigurer ce que le switch est capable de faire ou non, e.g. IPv6, MPLS, etc. — c'est souhaitable aussi!

C'est l'idée derrière la notion de « Software-defined network », traiter le problème de reconfiguration en faisant remonter les choses au niveau logiciel, plutôt qu'hardware.

Quelques exemples de « Software-defined »

- « Software-defined radio », faire de la radio sans encoder toute la logique au niveau ASIC directement, e.g. LimeSDR, WebSDR, etc. ;
- « Software-defined WAN », abstraire les technologies WAN (LTE, MPLS, etc.) et penser l'usage d'Internet comme partie intégrale (e.g. SaaS, O365, Salesforce) ;
- « Software-defined storage », centraliser la ressource de stockage et l'abstraire, e.g. Amazon S3 ;
- « Software-defined data center », on combine SDN, SDS, la virtualisation et le management de tout ça, e.g. AWS !

Retour au SDN

- Control plane: tables de routage, etc.
- Management plane[5]: monitoring, centralisation, etc.
- Data plane: tables « ARP », plan de commutation, etc.

[5] Considéré parfois comme un sous ensemble du control plane.

NETCONF, YANG, SNMP, etc.

Plusieurs protocoles existent pour observer et changer un parc d'équipements réseaux, ils ont plusieurs défauts et particularités :

- Très vendor-specific (suffit de lire le nombre de RFC) ;
- Basé sur XML, sur des fichiers à trimbaler (MIBs) ;
- La sécurité a été une arrière-pensée sur certains ;
- Requiert de l'équipement dans chaque rack pour contrôler l'ensemble correctement ;
- Ils sont utilisés et existent, je ne considère pas ça comme du « vrai SDN », même si on peut en construire par dessus

OpenFlow

OpenFlow est un protocole de configuration dynamique des switch, l'on peut :

- Configurer des « flows »
- Prendre des décisions dynamiques sur des paquets inconnus

Les « flows »

Un « flow »[6] c'est une règle en simple, elle décide :

- Conditions de matching, e.g. L2 ou L3, VLAN, etc.
- Ensemble d'actions à prendre, e.g. dropper, copier vers un port, etc.
- Priorités, timeouts, « cookies »[7]

[6] Dans le monde OpenFlow, mais c'est un vocabulaire devenu générique.
[7] L'équivalent d'un packet marker dans le monde Linux.

Bibliographie sur le SDN

- https://networkingnerd.net/2016/11/29/openflow-is-dead-long-live-openflow/
- https://www.bitsinflight.com/sdns-promise-lives-on/
- http://www.forwardingplane.net/2018/11/faucet-enterprise-openflow-in-production/
- https://lightbytes.es.net/2018/11/05/__trashed/

Mininet

Mise en bouche

Mininet permet de créer des réseaux réalistes virtuels, qui font tourner un noyau Linux, du code de switch et applicatif sur la même machine, en quelque secondes.

Outil de développement classique pour travailler avec OpenFlow ou P4[8]

[8] Que nous verrons plus tard dans cette présentation.

Demo (ou fail) time :-)

- Aller plus loin: https://ce.sc.edu/cyberinfra/workshops/Material/SDN/SDN_Labs.pdf

Faucet

Un contrôleur SDN à l'échelle locale

- Écrit en Python ;
- Configuration en YAML[9]
- S'intègre avec Prometheus, InfluxDB et autres ;
- Supporte TLS et les configurations `fall-secure` / `fall-standalone` ;
- OpenFlow $\geq$ 1.3.x avec plusieurs tables de groupe ;
- Haute disponibilité par idempotence.

[9] Beurk, mais... :-)

Un contrôleur SDN petit... mais puissant

- IPv6 supporté!! ;
- Apprentissage sécurisé: unicast flooding ;
- ACLs ;
- Policy-based forwarding avec offloading possible (e.g. 802.1x) ;
- Stacking de switch ;
- Conçu pour le NFV: Network Function Virtualization
- Routage BGP.

Interlude de normalien : Dhall

Si vous aussi, vous n'aimez pas YAML ou JSON pour configurer les choses, vous aimerez peut-être Dhall.

Autrement, vous serez sensible à l'importance des problématiques induites par les différents langages « de configuration »[10]

[10] Qui n'en sont pas.

Configurer Faucet

C'est du YAML que l'on va voir.

Configurer Faucet avec Dhall

On va voir ça aussi sur le site.[11]

[11] Cette slide a été écrit le matin de Federez, donc, avec le niveau de fatigue :>

Démonstration avec Mininet

Bon, il faut montrer que ça marche maintenant. :-)

Pas de vrai switch sous la main, pas de vraie démo, mais on a mininet et OVS au moins !

Monitoring avec Faucet

Je triche, puisque je n'ai rien sous la main.

- https://grafana.redcables.wand.nz/d/000000008/redcables-openflow-statistics?orgId=1&refresh=10s&search=open&folder=current

Proxmox

Un petit tour des capacités SDN

- BGP-EVPN / VXLAN: contrôleurs et zones (QinQ / VLAN supportés)
- IPAM: phpIPAM / Netbox IPAM
- DNS: PowerDNS

Capacité SDN inattendue ?

On peut utiliser des switch OpenVSwitch dans Proxmox!

Donc, on peut utiliser Faucet et OpenFlow :-).

Démonstration laissée en exercice aux spectateurs[12]

[12]Non, je n'ai pas eu le temps de reconfigurer le lab du club réseau avec OVS, la vie est difficile.

Retour sur la reconfiguration: P4

Introduction rapide

On peut aller plus loin et coordonner OpenFlow avec une reconfiguration dynamique des dissecteurs de protocole des switch.

- On peut inventer un nouveau protocole de tunnel
- On peut implémenter des protocoles non disponibles dans le switch et les accélérer
- On peut ré-attribuer les ressources du switch pour les concentrer sur des fonctionnalités, e.g. supprimer le support IPv4, rediriger la TCAM, etc.

Un petit mot sur le club réseau de l'ENS Ulm

KlubRZ

Le club réseau de l'ENS Ulm est un club d'expérimentation autour des réseaux informatiques et des nouvelles technologies, on fait joujou sur tout ce qu'on a sous la main :

- NixOS : demandez moi ;)
- Proxmox : SDN, automatisation des ressources, ZFS
- DN42 et BGP : BGP hijacking, RPKI
- Techniques de sécurité avancés: TPM, Secure Boot, remote unlock of full disk encryption[13]
- IPv6 : link-local, multicast
- Autres : WireGuard, Kubernetes, eBPF, etc.

[13] Avec Mandos par exemple.

Équipement

Nous avons un petit peu de budget tous les ans ($< 1\,000$ €) pour acheter le matériel critique (e.g. switch, un serveur R720, de l'espace, des tunnels chez un provider — MilkyWAN).

Si vous aimez ce qu'on fait et vous souhaiteriez voir plus de contenu, n'hésitez pas à nous le dire et si vous avez du matériel dont vous voulez vous débarrasser, nous sommes preneurs !

- En serveurs: PowerEdge RX20 ou de génération supérieures, RAM DDR3 adaptée, disques SAS (600GB et plus), SSDs ;
- En switch: OpenFlow ≥ 1.3, P4 idéalement ;
- En borne WiFi: compatible openwrt, WiFi 5 2nd wave préférablement, ARM/x86 en priorité, sinon MIPS et autres exotismes OK ;
- Autres: de la fibre, des SFP+ (ou des DAC), des cartes réseaux 10Gbps et plus, des ressources IPs, baies de stockage

Questions ?