

Mémoire

présenté et soutenu le 10 avril 2014 par
Cynthia LOPES DO SACRAMENTO

Applications SDN

du data centre au Cloud

Jury : Romain KOBYLANSKI
 François MILLER
 Véronique PANNE

Tuteur : Claude CASERY
Entreprise : Bull

Table des matières

Introduction	1
1 Les évolutions du business modèle des Data Centres	3
1.1 Data centres et ses objectifs	3
1.2 Organisation d'un data centre et difficultés	4
1.3 Virtualisation et partage de ressources	6
1.4 Le besoin d'un modèle plus dynamique	7
1.5 Tendances des meilleures pratiques	8
1.6 Architecture	8
1.7 Cloud Computing	8
1.8 Définition	9
1.9 Nouveau business model	10
1.10 Principaux avantages du Cloud Computing	11
1.11 Barrières au Cloud Computing	11
2 L'aspect réseau	13
2.1 Le réseau dans un data centre	13
2.2 Challenges réseau	15
2.3 Différents usages	15
2.4 Agilité	15
2.5 Sécurité	15
2.6 Changement des réseaux data centre	15
3 Applications SDN	17
3.1 Redéfinition de SDN	17
3.2 Solutions	17
4 Apports de SDN aux data centres	19
4.1 Différents usages	19

4.2	Agilité	19
4.3	Sécurité	19
	Conclusion	21
	Acronymes	23
	Glossaire	25
	Bibliographie	29

Liste des tableaux

Table des figures

1.1	Organisation de racks. [5]	5
1.2	Modèle d'infrastructure à ressources partagées. [7]	6
1.3	Capacité fixe de ressources vs charge prévisionnelle. [8]	8

Introduction

Les centres de traitement de données évoluent aujourd'hui à un rythme intense pour accompagner l'explosion de l'usage des données. L'accélération de l'innovation dans l'informatique impose une rénovation constante du business. La virtualisation a permis aux centres de données d'améliorer la productivité de ses serveurs, mais pour arriver à l'agilité envisagée, les data centres doivent évoluer leurs réseaux et débloquent le Cloud Computing. Cette étude analyse les applications **Software-Defined Networking : Réseau Informatique Défini par Logiciel (SDN)** pour distinguer ses apports dans le contexte des data centres et habilitier 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 de la capacité de traitement de l'information qu'on observe aujourd'hui. Plusieurs centres de données sont devenus hautement distribués et relativement fragmentés. Comme résultat, ils sont limités dans leur capacité d'évoluer rapidement et de supporter l'intégration des nouveaux types de technologies ou se mettre à l'échelle pour monter le business en puissance selon les besoins de ses consommateurs.

Quand équipée avec des infrastructures assez performantes, partagées et dynamiques ainsi qu'avec des outils nécessaires pour libérer ressources de la demande traditionnelle, les **Système d'Information (SI)** peuvent répondre plus efficacement aux besoins métiers. Par conséquent, les organisations pourraient se focaliser dans l'innovation et l'alignement de ses ressources à leurs priorités stratégiques plus larges. Cela soulagerait la prise de décisions, qui pourrait passer à se baser sur l'information en temps-réel.

Alors que les coûts du réseau dans un data centre est estimé à 15% [1] du total et n'étant pas un des plus larges, il est largement argumenté qu'il représente un élément clé pour la réduction des coûts et augmentation du retour sur l'investissement. Les coûts d'investissement dans le serveurs ont été évalués à 45% des coûts des data centres. Malheureusement l'usage utile des serveurs est remarquablement bas, pouvant arriver à

seulement 10% d'utilisation dans certains exemples.

La technique de la virtualisation a réussi à habiliter les processus d'être déplacés entre machines, mais des contraintes réseau continuent à limiter l'agilité dans les data centres. L'agilité est définie par la capacité d'assigner n'importe quel 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 ; ils fragmentent par nature tout ensemble les réseaux et la capacité des serveurs, limitant et réduisant la croissance dynamique des pools de serveur. [2]

L'agilité est donc un élément clé ; diverses entreprises bataillent péniblement pour déployer des nouvelles applications ou faire évoluer les existantes à l'allure de croissance de leurs 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 se lève à plus de 5 mois). Les résultats affichent également que 59% des organisations nécessitent plus de huit heures pour réaliser un changement de connectivité dans une application. [3]

Cependant, lors du passage au Cloud, les organisations réalisent que la virtualisation des serveurs est sévèrement 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 critiques deviennent pénibles :

- Prévention de la congestion du trafic ;
- Réduction de la complexité des politiques réseau et garanties du niveau service ;
- Élimination des points aveugles qui conduisent à des pannes ;
- Scellage des failles de sécurité pour protéger les données.

Cette étude a pour but de démontrer comment SDN peut être appliqué aux data centres pour débloquer le Cloud Computing et élargir ses limites. Dans le premier chapitre, le contexte des data centres sera définie. En suite, les enjeux dans l'aspect réseau seront démontrés. Le chapitre suivant présentera les applications SDN qui répondent à ses enjeux. Finalement, le quatrième et dernier chapitre présentera les apports de SDN dans ce cadre.

Chapitre 1

Les évolutions du business modèle des Data Centres

Ce chapitre a pour but de définir un data centre afin de pouvoir analyser ses problématiques, enjeux et possibles solutions. En vue de comprendre l'état actuel des data centres et ses limitations par rapport aux nouveaux besoins et challenges business. Un regard sur le nouveau business model apporté avec le Cloud Computing, les bénéfices de son adoption et les enjeux pour les infrastructures qui doivent répondre à ce nouveau paradigme.

1.1 Data centres et ses objectifs

Un data centre (ainsi dénoté ferme de serveurs) est un répertoire centralisé pour le stockage, management et distribution de données et informations. Typiquement, un data centre est une installation utilisée pour loger des systèmes informatiques et ses composants associés, tels que systèmes de télécommunication et stockage.

Les data centres traditionnels hébergent historiquement des nombreuses applications relativement petites ou moyennes, chacune exécutant dans une infrastructure matérielle dédiée qui est isolée et protégée des autres systèmes dans la même installation. 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 en commun en termes de matériel, logiciel ou infrastructure de maintenance, et en général ne se communiquent pas entre eux.

Les tendances vers l'informatique côté serveur et l'explosion en popularité des services sur internet ont changé ce scénario. Des infrastructures data centre entières ont été dédiée à un seul acteur pour faire fonctionner ses services offerts. Dans ce cadre, un data centre appartient à une seule organisation et utilise du matériel et plateforme logicielle relativement homogènes qui partagent une couche commune de systèmes de management. Surtout, ces data centres dédiés exécutent un nombre réduit d'applications (ou services internet) beaucoup plus importants en taille, l'infrastructure commune de management permettant une significative flexibilité de déploiement.

Ces infrastructures sont montées pour gérer la taille des applications déployées et la haute disponibilité exigée pour ces services, visant en général 99,99% de durée de fonctionnement (une heure au maximum de temps d'arrêt par an). Atteindre un fonctionnement libre des failles dans une large collection de systèmes matériel et logiciel est dur et devient encore plus difficile 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 conséquence, des nouvelles approches ont été proposées pour la construction et opération de ces systèmes qui doivent être conçus pour tolérer ce nombre important des failles avec très peu ou aucun impact sur la performance et disponibilité des services offerts. [4] [5]

1.2 Organisation d'un data centre et difficultés

Un data centre est en général organisé en lignes de racks où chaque rack contient des dispositifs modulaires tels que serveurs, switches, briques de stockages ou instruments spécialisés. Des composants essentiels de l'infrastructure qui sont branchés aux racks des data centres d'entreprises tels que compute, stockage et réseau sont la base sur laquelle les applications business sont construites . Un chassis vient complet avec ses propres ventilateurs, source d'alimentation, panier d'interconnexion et module de management. Pour réduire l'espace occupé, des serveurs peuvent être compartimentés dans un chassis qui est glissé dans le rack. Un chassis fournit des slots de taille standard où il est possible d'insérer des éléments actifs modulaires (ainsi connus tant que "blades"). Un seul chassis peut contenir 16 serveurs 1 U, comme les racks supportent 6 chassis, ils ont une capacité théorique de 96 éléments modulaires.

La figure 1.1 montre l'organisation des racks dans un data centre. Un serveur occupe

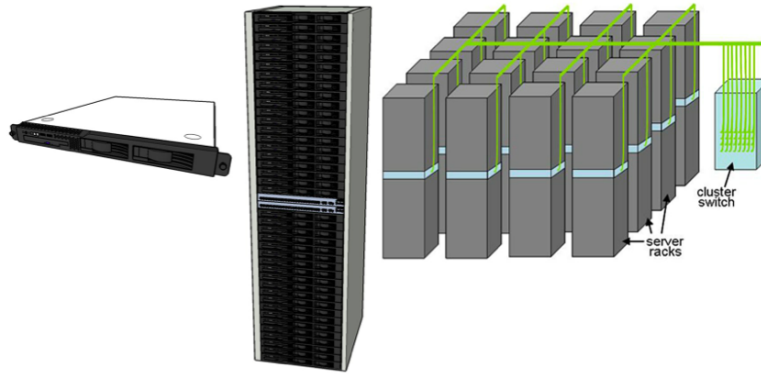


FIGURE 1.1 – Organisation de racks. [5]

1 U du rack est montré à gauche. Au milieu on affiche un rack et à droite un cluster de racks avec un switch/routeur de niveau cluster. En général un ensemble de serveurs 1U sont montés dans un rack et inter-connectés avec commutateur Ethernet local. Ces switches au niveau des racks, qui peuvent utiliser des liens de 1 à 10 Gbps, ont un nombre de connexions uplink vers un ou plus switches de niveau cluster (data centre).

Le stockage dans les data centre peut être fourni en diverses manières. Souvent le stockage de haute performance est logé dans des « tours de stockage » qui permettent un accès distant transparent au stockage indépendamment du nombre et des types des dispositifs de stockage physiques installés. Le stockage peut également être fourni dans plus petit « brique de stockage » localisé dans le rack ou slot de chassis ainsi que directement intégré aux serveurs. Dans tous les cas, un accès réseau efficace au stockage est crucial.

Le problème le plus important dans cette structure est la potentielle insuffisance de bande passante. En général, les connexions uplink sont conçues pour supporter un certain taux de demande excédentaire puisque la fourniture d'une bande passante entière n'est toujours possible. Par exemple, 20 serveurs à 1Gbps chacun doit partagé un uplink Ethernet unique de 10Gbps à un taux de demande excédentaire de 2. Cette situation peut être problématique si la charge réseau non local monte considérablement. Comme le stockage est traditionnellement fourni dans une tour séparée, tout le trafic de stockage traverse le lien uplink 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 essentiel.

La consommation d'énergie constitue également des préoccupations à la conception des data centres car les coûts liés sont devenus un important composant de la totalité des coûts de cette classe de systèmes. Actuellement les CPUs ne sont plus le seul élément cible d'amélioration de l'efficacité énergétique, vu qu'ils ne dominent plus la majorité de la consommation. Des problématiques de ventilation et surconsommation d'énergie sont de plus en plus critiques de la conception de data centres.[5] [6]

1.3 Virtualisation et partage de ressources

Le besoin d'augmenter l'efficacité de l'utilisation des ressources a amené les tendances de conception d'infrastructures avec le partage de ressources et la virtualisation . Virtualisation fait référence à l'abstraction de ressources logiques de leurs couches physiques pour améliorer l'agilité, la flexibilité et la réduction des coûts et ainsi mettre en valeur le 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 significativement à l'économie des coûts.

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

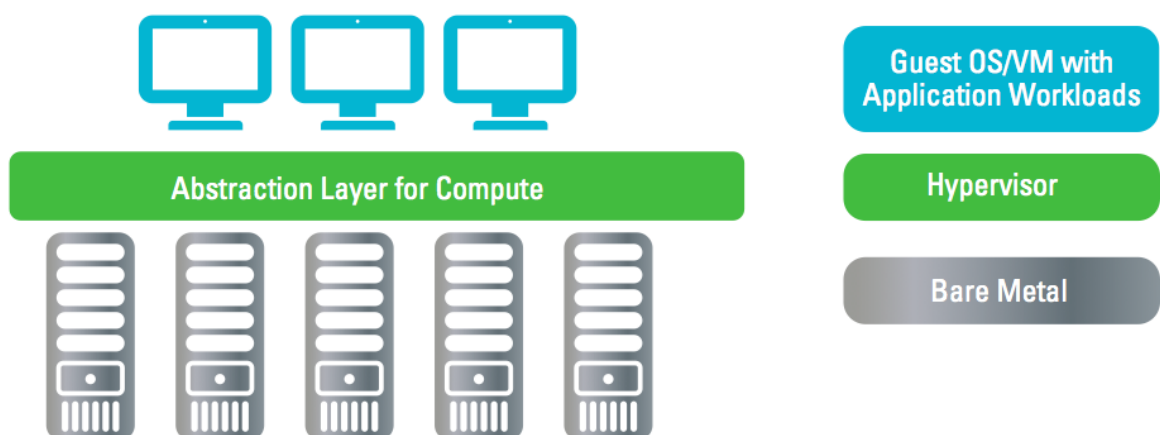


FIGURE 1.2 – Modèle d'infrastructure à ressources partagées. [7]

Un hyperviseur ou moniteur de machines virtuelles 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. L'hyperviseur fournit l'émulation matériel aux systèmes invités et gèrent l'allocation de ressources matérielles.

Ce modèle apporte des avantages en termes de comment les ressources sont efficacement utilisées avec des charges applicatives idéales. Cependant, quand une application commence à consommer plus ressources que l'estimé, il peut arriver des scénarios où les systèmes d'exploitation invités n'ont pas assez de ressources, impactant la qualité du service business offert.

Alors que cette approche a apporté une capacité globale de management, monitoring et outillage, elle a aussi mis en évidence que le composant compute de l'infrastructure ressources serveurs se bénéficient clairement l'amélioration et automatisation de leur utilisation. 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 application restait toujours en manque. [ibmPlanningVirtCCchap2][7]

1.4 Le besoin d'un modèle plus dynamique

Traditionnellement les data centres d'entreprises sont conçus pour durer pour toujours et atteindre les objectifs visibles du business. Cela veut dire que les éléments sous-jacents sont dimensionnés et construits pour supporter le pic de charge projeté en termes de performance, disponibilité et sécurité. Quand la croissance volumétrique projetée ne correspond pas à la réalité, cette méthode de dimensionnement peut conduire à une situation de sous-dimensionnement ou sur-dimensionnement. Ce qui apporte 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 varie continuellement dans les applications sur internet, il reste deux choix aux business : soit sous-dimensionner la provision 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, étant donné que leur capacité est fixe. L'image suivante illustre cette planification et les situations de problèmes de dimensionnement.

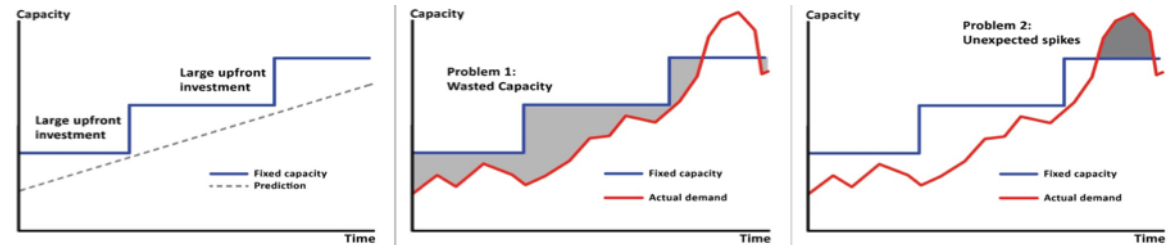


FIGURE 1.3 – Capacité fixe de ressources vs charge prévisionnelle. [8]

En face à cette problématique, un nouveau modèle de livraison pour adresser les défis du traitement des demandes pour la variation dynamique des charges applicatives. Avec la nouvelle tendance du Cloud Computing et l'infrastructure tant que service (IaaS), la conception de clusters hautement disponibles et des solutions extensibles peuvent être architecturés avec les requis non-fonctionnels comme base.

Avec sa nature extensible, le modèle de livraison cloud permet aux ressources d'étendre et de réduire dynamiquement en fonction de la consommation. Une couche logicielle d'abstraction, implémentée par les hyperviseurs, virtualise le traitement des ressources physiques, ainsi permettant le processeur, la mémoire et les disques durs de s'accommoder aux variations de demande. [8] [7]

Chapitre 2

Cloud Computing

2.1 Définition

In very simple terms, cloud computing is a new consumption and delivery model for information technology (IT) and business services and is characterized by : • On-demand self-service • Ubiquitous network access • Location-independent resource pooling • Rapid elasticity and provisioning • Pay-per-use Cloud has evolved from on demand and grid computing, while building on significant advances in virtualization, networking, provisioning, and multitenant architectures. As with any new technology, the exciting impact comes from enabling new service consumption and delivery models that support business model innovation.

As we have seen, data centers have grown to serve a wide range of business needs, and there are many factors to consider when designing a solution that meets different objectives. Within the past several years, a powerful new paradigm has emerged that has important implications for data center architectures and how they meet these varied objectives. This is the paradigm of cloud computing.

Cloud computing delivers services dynamically over networks from an abstracted set of resources. The resources are somewhere in the cloud and available on demand. The types of resources and their location are transparent to end users. End users primarily care that their applications, data and content are secure and available, with a desired level of quality.

From the infrastructure perspective, cloud computing heavily leverages resource pools

in a variety of technologies— compute, storage and network—for dynamic allocation in an automated, orchestrated and logically diversified environment, accommodating a variety of applications. Using orchestration, resources can be pooled within and across multiple data centers to provide an environment that responds dynamically to user needs.

2.2 Besoins qui amènent au Cloud

Information technology (IT) is at a breaking point, and there is a critical need to improve IT's impact on the business.⁹ Consider the following : -As much as 85% of computing capacity sits idle in distributed computing environments. -Seventy percent of IT budgets is typically spent on maintaining current IT infrastructures, and only 30% is typically spent on new capabilities. -Over 30% of consumers notified of a security breach will terminate their relationship with the company that contributed to the breach. Clearly, infrastructures need to be more dynamic to free up budgets for new investments and accelerate deployment of superior capabilities being demanded by the business. Nearly all CEOs are adapting business models; cloud adoption can support these changing business dynamics.

Functional Areas in the Cloud-Ready Data Center

- Network Infrastructure—provides connectivity and transport for applications and services between users and the data center, within the data center and across multiple data centers. The Network infrastructure has three main sub components, namely the access network, the core network and the edge network.
- Compute and Storage—represents the compute and storage infrastructure appropriate for applications (rack-mount and chassis-based, cost-effective and multi-core, with unstructured content and highly structured transaction databases). The compute and storage functional area hosts all business applications such as Enterprise Resource Planning (ERP), SaaS, SOA and Web 2.0 applications (among others).
- Services—supports applications with security, user verification, and entitlement, and application support, including application acceleration, deep packet inspection (DPI), and load balancing
- Management and Orchestration—ties together all of the elements of the cloud-computing infrastructure, enabling efficient and responsive monitoring, management, and planning

2.3 Nouveau business model

Even within the cloud computing space there is a spectrum of offering types. There are five commonly used categories :

- Storage as a Service - SaaS Provisioning of database-like services, billed on a utility computing basis, for example, per gigabyte per month.
- Infrastructure as a Service - IaaS Provisioning of hardware or virtual computers where the client has control over the OS, therefore allowing the execution of arbitrary software.
- Platform as a Service - PaaS Provisioning of hardware and OS, frameworks and databases, for which developers write custom applications. There will be restrictions on the type of software they can write, offset by built-in application scalability.
- Software as a Service - SaaS Provisioning of hardware, OS, and special-purpose software made available through the Internet.
- Desktop as a Service - DaaS Provisioning of the desktop environment, either within a browser or as a Terminal Server.

2.4 Principaux avantages du Cloud Computing

Key benefits of cloud computing :

- Flexibility – There is the ability to update hardware and software quickly to adhere to customer demands and updates in technology.
- Savings – There is a reduction of capital expenditures and IT personnel.
- Location & Hardware Independence – Users can access application from a web browser connected anywhere on the internet.
- Multi-tenancy – Resources and cost are shared among many users, allowing overall cost reduction.
- Reliability – Many cloud providers replicate their server environments in multiply data centers around the globe, which accounts for business continuity and disaster recovery.
- Scalability – Multiply resources load balance peak load capacity and utilization across multiply hardware platforms in different locations
- Security – Centralization of sensitive data improves security by removing data from the users' computers. Cloud providers also have the staff resources to maintain all the latest security features to help protect data.
- Maintenance – Centralized applications are much easier to maintain than their distributed counter parts. All updates and changes are made in one centralized server instead of on each user's computer.

2.5 Barrières au Cloud Computing

IT organizations have identified four major barriers to large-scale adoption of cloud services :

- **Security, particularly data security** Interestingly, the security concerns in a cloud environment are no different from those in a traditional data center and network. However, since most of the information exchange between the organization and the cloud service provider is done over the web or a shared network, and because IT security is handled entirely by an external entity, the overall security risks are perceived as higher for cloud services. Some additional factors cited as contributing to this perception :
 - Limited knowledge of the physical location of stored data
 - A belief that multitenant platforms are inherently less secure than single-tenant platforms
 - Use of virtualization as the underlying technology, where virtualization is seen as a relatively new technology
 - Limited capabilities for monitoring access to applications hosted in the cloud
- **Governance and regulatory compliance** Large enterprises are still trying to sort out the appropriate data governance model for cloud services, and ensuring data privacy. This is particularly significant when there is a regulatory compliance requirement such as SOX or the European Data Protection Laws.
- **Service level agreements and quality of service** Quality of service (availability, reliability, and performance) is still cited as a major concern for large organizations :
 - Not all cloud service providers have well-defined SLAs, or SLAs that meet stricter corporate standards. Recovery times may be stated as “as soon as possible” rather than a guaranteed number of hours. Corrective measures specified in the cloud provider’s SLAs are often fairly minimal and do not cover the potential consequent losses to the client’s business in the event of an outage.
 - Inability to influence the SLA contracts. From the cloud service provider’s point of view it is impractical to tailor individual SLAs for every client they support.
 - The risk of poor performance is perceived higher for a complex cloud-delivered application than for a relatively simpler on-site service delivery model. Overall performance of a cloud service is dependent on the performance of components outside the direct control of both the client and the cloud service provider, such as the network connection.
- 1. **Integration and interoperability** Identifying and migrating appropriate applications to the cloud is made complicated by the interdependencies typically associated with business applications. Integration and interoperability issues include :
 - A lack of standard interfaces or APIs for integrating legacy applications with cloud services. This is worse if services from multiple vendors are involved.
 - Software dependencies that must also reside in the cloud for performance reasons, but which may not be ready for licensing on the cloud.
 - Interoperability issues between cloud providers. There are worries about how disparate

applications on multiple platforms, deployed in geographically dispersed locations, can interact flawlessly and can provide the expected levels of service.

Chapitre 3

L'aspect réseau

Dans ce chapitre, les principales problématiques data centre dans un aspect réseau seront présentées et analysées. Comment on fait aujourd'hui ? Quels sont les limites ?

3.1 Le réseau dans un data centre

As organizations undertake information technology (IT) optimization projects, such as data center consolidation and server virtualization, they need to ensure that the proper level of focus is given to the critical role of the network in terms of planning, execution, and overall project success. While many consider the network early in the planning stages of these projects and spend time considering this aspect of these initiatives, many more feel that additional network planning could have helped their projects be more successful.

The most common types of network changes in IT optimization projects include implementing new network equipment, adding greater redundancy, increasing capacity by upgrading switches, improving network security, and adding network bandwidth. However, many network requirements associated with these changes and the overall initiative are typically not identified until after the initial stages of the project and often require rework and add unanticipated costs. Regardless of project type, network challenges run the risk of contributing to increased project time lines and/or costs.

The networking aspects of projects can be challenging and user complaints about the network are frequently heard. Important challenges include the inability to perform

accurate and timely root-cause analysis, understand application level responsiveness, and address network performance issues. Simply buying more network equipment does not necessarily or appropriately address the real requirements.

Looking ahead, many expect that the network will become more important to their companies' overall success. To address this, networking investments related to support of server and storage virtualization are currently at the top of the list for consideration, followed by overall enhancement and optimization of the networking environment. To support virtualization of the entire IT infrastructure and to continue to optimize the network, IT organizations need to make architectural decisions in the context of the existing infrastructure, IT strategy, and overall business goals.

Developing a plan for the network and associated functional design is critical. Without a strong plan and a solid functional design, networking transitions can be risky, leading to reduced control of IT services delivered over the network, the potential for high costs with insufficient results, and unexpected performance or availability issues for critical business processes.

With a plan and a solid functional design, the probability of success is raised : a more responsive network with optimized delivery, lower costs, increased ability to meet application service level commitments, and a network that supports and fully contributes to a responsive IT environment.

3.2 Challenges réseau

3.3 Différents usages

3.4 Agilité

3.5 Sécurité

3.6 Changement des réseaux data centre

In order to enable a dynamic infrastructure capable of handling the new requirements that have been presented in the previous section, a radical shift in how the data center network is designed is required. The figures here show the comparison between the traditional thinking and the new thinking that enables this change of paradigm. Figure 1-3 illustrates a traditional, multitier, physical server-oriented, data center infrastructure.

Chapitre 4

Applications SDN

Ce chapitre redéfinira SDN et présentera ses réponses aux problématiques réseau rencontrées en général dans les data centre, discutées dans le chapitre précédent. Comment SDN approche la problématique ? Qu'est-ce que SDN ?

4.1 Redéfinition de SDN

4.2 Solutions

Chapitre 5

Apports de SDN aux data centres

Ce chapitre démontre les apports de SDN au sein des data centre par rapports aux problématiques présentées précédemment.

5.1 Différents usages

5.2 Agilité

5.3 Sécurité

Conclusion

Même avec le succès incontestable de l'architecture d'internet, l'état de l'industrie réseau et l'essence de son infrastructure se trouvent en phase critique. Il est généralement admis que les réseaux courants sont excessivement chers, compliqués à gérer, sujets aux blocages des fournisseurs et difficiles à faire évoluer.

On constate donc un réel besoin de faire évoluer cette architecture mais des résistances s'opposent à cette évolution en raison de la complexité et la possible saturation du système. En réponse, les réseaux programmables ont été un objet intensif de recherche par la communauté. Les travaux dans ce domaine s'orientent vers l'offre SDN, un nouveau paradigme transformant cette architecture.

L'approche SDN sépare le plan de contrôle et le plan de données, offrant un contrôle et une vision centralisés du réseau. Cela peut apporter certains bénéfices comme le contrôle directement programmable, la simplification du hardware réseau et la simplification de l'ingénierie du trafic. En revanche, des défis d'implémentation sont à surmonter tels que la concentration des risques dans un contrôle physiquement centralisé, l'équilibre entre flexibilité et performance et les conditions d'interopérabilité.

La flexibilité apportée par SDN est telle que de nombreuses possibilités d'applications sont à imaginer. Essentiellement pour l'administration de data centers, le contrôle d'accès et de la mobilité pour les réseaux campus ainsi que l'ingénierie du trafic pour les réseaux WAN.

Le marché suit de près les nouveautés dans le domaine et investit sur les technologies implémentant SDN. Les stratégies ne sont pas encore assez matures et les consommateurs potentiels attendent des offres plus consolidées. Cependant, des solutions innovantes commencent à surgir et certaines sociétés assument le rôle de tête dans le marché.

On s'aperçoit que l'ampleur des possibilités SDN, même si elle présente un avantage en théorie, freine son adoption. En raison de la grande variété de concepts et produits,

les consommateurs hésitent toujours à prendre une décision. En même temps, les grands fournisseurs cherchent à la fois à exploiter le nouveau marché et à protéger leurs solutions consolidées. Ces obstacles même s'ils sont confirmés, ne semblent pas être assez forts pour empêcher les échanges à long terme.

Au vu de cette étude, il semblerait que dans un futur proche, les clients les plus informés et les plus disposés à innover vont commencer à déployer SDN. Leurs expériences et les résultats obtenus vont fortement impacter le choix des prochains consommateurs. Il est possible que ceux qui dessineront le futur de la technologie des réseaux informatiques pour les prochaines années seront ceux qui auront osé se lancer les premiers. Cette démarche peut éventuellement représenter un risque, mais aussi l'opportunité de tirer des bénéfices plus durables et de prendre de plus larges parts du marché.

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
- DHCP** Dynamic Host Control Protocol, Protocole pour la configuration automatique d'hôte
- DNS** Domain Name System, Système de noms de domaine
- HTTP** HyperText Transfer Protocol, 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
- ONE** Open Network Environment, Environnement Réseau Ouvert
- ONF** Open Networking Foundation

QoS Quality of Service, Qualité de service

SDN Software-Defined Networking : Réseau Informatique Défini par Logiciel

SI Système d'Information

TI Technologie de l'Information

VLAN Virtual Local Area Network, Virtual LAN

VM Virtual Machine, Machine Virtuelle

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 où les couches correspondent à des niveaux de détails appliqués. [9]

Big Data Big Data est un terme appliqué aux ensembles de données dont la taille (ou le format) est au-delà de 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. [10]

Cloud Computing Cloud Computing, ou informatique dans les nuages, est une évolution de la fourniture de services **TI** qui 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**, fournissant un niveau plus haut d'automatisation. Diverses entreprises ont adopté le cloud computing et réalisent des avantages significatifs en agilité, réduction de coûts et soutien de la croissance du business. [11]

Cluster En réseaux informatiques, un cluster désigne un groupe des machines reliées entre elles à l'aide d'un réseau de communication. Cette configuration est souvent utilisée pour réaliser des calculs à haute performance. [12]

Data Center 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 center inclut alimentation et connexions des données redondantes, contrôles d'environnements comme la climatisation ainsi que divers dispositifs de sécurité. [13]

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 rassemblerait 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 transmettent des données l'un à l'autre à travers les switches d'interconnexion. [14] [15]

Middlebox Boîtier intermédiaire. Un middlebox est un serveur conservant des 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 concernant les sessions de communications. [16]

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. [17]

Open Source Logiciel avec code source ouvert, qui peut donc être utilisé librement, modifié et partagé par quelqu'un. Un logiciel open source est développé par plusieurs personnes et distribué sous des licences qui se conforment à la définition d'open source. [18]

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. [19] [20]

Paradigme Un paradigme consiste en une collection de règles, 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. [21] D'après Dosi [22], 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 sorte 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. [23]

Plan de Données Le plan de données traite les data-grammes entrants dans le média à travers une série d'opérations au niveau des liens qui collectent ces data-grammes et réalisent divers tests de cohérence basiques. Ensuite les data-grammes sont transférés en accord avec des tableaux pré-remplis par le **plan de contrôle**. [23]

Scalabilité Terme provenant de l'anglicisme *scalability* qui exprime la capacité d'être mis à échelle. En informatique cela désigne la capacité d'un système, d'un réseau ou un processus de gérer l'augmentation ou la réduction de la charge de manière à pouvoir la gérer. [24]. 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 ». [25]

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 aidé à réduire les coûts et a également augmenté la flexibilité et la performance. [11]

Bibliographie

- [1] 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.
- [2] 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.
- [3] *Examining the Impact of Security Management on the Business*. Executive Summary. An AlgoSec Survey. 2013.
- [4] *Understanding Data Centers and Cloud Computing, Section What Is a Data Center?* White Paper. Global Knowledge Training LLC. 2010.
- [5] Luiz André BARROSO et Urs HÖLZLE. « 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.
- [6] 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.
- [7] Sandeep RAGHURAMAN. *The Journey Toward the Software-Defined Data Center*. White Paper. Cognizant (NASDAQ: CTSI). Sept. 2013.
- [8] 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.
- [9] Robert M. KELLER. *Computer Science: Abstraction to Implementation, Section 1.1 The Purpose of Abstraction*. Article. Harvey Mudd College. Sept. 2001.
- [10] *Information Management and Big Data A Reference Architecture*. An Oracle White Paper. Fév. 2013.

- [11] *Intel's Vision of Open Cloud Computing, section Speeding Agility, Reducing Costs, and Accelerating Innovation via Cloud*. White Paper. Intel IT Center. 2013.
- [12] Qingkui CHEN, Haifeng WANG et Wei WANG. « Continuance Parallel Computation Grid Composed of Multi-Clusters. » In : *Journal of Networks* 5.1 (2010).
- [13] William TSCHUDI et al. « High-performance data centers: A research roadmap ». In : (2004).
- [14] Margaret ROUSE. *network fabric*. Web Site. <http://searchsdn.techtarget.com/definition/network-fabric>. Mar. 2014.
- [15] 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>.
- [16] 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.
- [17] *Open Daylight Project*. Web site. <http://www.opendaylight.org/project>. Avr. 2014.
- [18] *Open Source Initiative*. Web site. <http://opensource.org/>. Avr. 2014.
- [19] 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.
- [20] 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>.
- [21] 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.
- [22] G. DOSI. *Technological paradigms and technological trajectories, Research Policy*. Anglais. 1982.

- [23] 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.
- [24] 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>.
- [25] René J CHEVANCE. « Serveurs multiprocesseurs et SGBD parallélisés ». In : *Techniques de l'ingénieur. Informatique* H2068 (2001), H2068–1.

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. Comme par le passé pour le Big Data, conçu pour le traitement des énormes quantités de données ou le Cloud Computing pour le management de l'hébergement de ressources. Une évolution similaire est 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 est un nouveau paradigme créé pour adapter les infrastructures courantes aux enjeux de la communication actuelle : une plus haute bande passante et les exigences des applications modernes. SDN propose une nouvelle architecture plus dynamique, 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.

Mots clés : SDN, Réseaux Programmables, Plan de Contrôle, Plan de Données

Abstract

New technologies and concepts appear in response to new network and internet usage requirements. Such as Big Data for massive data processing and Cloud Computing for resources hosting management. Similar evolution is 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 is a new paradigm designed to adapt current infrastructures to the issues of the recent communication : increasingly need for high speed and the exigences of modern applications. SDN proposes a new architecture plus dynamic, ease to manage, profitable and flexible. 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.

Keywords : SDN, Programmable Networks, Control Plane, Data Plane
