

Les défis du câblage dans les data centers haute performance

WHITEPAPER



Les data centers en tant qu'infrastructure critique

En Bavière, le premier jour des vacances forcées dues au coronavirus, le portail éducatif bavarois MEBIS s'est littéralement effondré. L'annonce officielle¹ faisait état d'une attaque DDoS². Pourtant, chez les enseignants et les élèves, tout le monde savait que MEBIS était loin d'être parfait. MEBIS n'était tout simplement pas conçu pour accueillir environ un demi-million d'élèves qui, du jour au lendemain, allaient se connecter au serveur à 8 heures précises³. On n'ose à peine imaginer ce qui se serait passé si, dans cette situation, des groupes actifs de pirates informatiques s'étaient trouvés sur le réseau.

La rapidité avec laquelle l'État fédéral de Bavière est parvenu à maîtriser le problème est tout simplement étonnante. Le dépannage a été mené parallèlement sur plusieurs couches OSI⁴. Pour commencer, une réorganisation au niveau des applications a été réalisée afin de répartir la charge sur plusieurs plateformes. Différentes offres ont été utilisées en fonction de la tranche d'âge. Ainsi, en dehors des offres de service web locales et bavaroises, des offres Cloud telles que Microsoft Teams ont également été utilisées. En outre, les capacités des data centers ont été réattribuées et l'expansion physique qui venait d'être amorcée a été poursuivie. Les capacités des data centers ont alors été multipliées par dix en l'espace de quelques jours seulement⁵. Certaines améliorations non détaillées ont été mises en œuvre afin de réduire les temps de latence et de chargement⁶.

Ainsi, « l'enfant mal aimé » MEBIS s'est transformé en histoire à succès et est parvenu à concilier la numérisation à la politique de l'éducation d'une manière exemplaire. Il nourrit désormais notre espoir de pouvoir, à l'avenir, profiter de davantage de flexibilité et d'offres de services plus rapides de la part des entreprises et des autorités.

Maintenant, on peut naturellement discuter sur la question de savoir si l'exemple cité est vraiment plus important qu'une numérisation globale du système de santé ou du plan d'urgence. Toutefois, cette nouvelle orientation en matière de numérisation met en avant un aspect crucial de notre réalité : pouvoir s'adapter rapidement aux bouleversements économiques et sociaux est essentiel. Ainsi, si la conduite autonome semble aujourd'hui secondaire, ce n'est pas le cas de l'utilisation de l'intelligence artificielle à des fins de prévision de propagation des virus ou de contrôle intelligent de la prévention des infections. Sans oublier qu'il est toujours question de réagir rapidement. Ainsi, les temps de latence représentent un aspect tout aussi critique que la disponibilité et le contrôle des ressources.

Il est incontestable que les data centers font partie des infrastructures critiques. Et au printemps 2020, chacun de nous a pris personnellement conscience de la nécessité d'un traitement de données rapide et fiable.

Les data centers : des fournisseurs de communication et processeurs d'information extrêmement flexibles

La virtualisation et le SDDC⁷ permettent aux data centers de réagir rapidement et de manière flexible face aux changements, d'assumer de nouvelles tâches ou de fournir de nouvelles interfaces. Toutefois, disposer du logiciel de virtualisation approprié, du système d'exploitation du

réseau NOS⁸ et des systèmes de gestion ne suffit pas. L'infrastructure sous-jacente des data centers doit également être hautement performante, flexible, fiable et extensible afin que les SDS, SDN et SDC⁹ puissent vraiment faire valoir leurs points forts.

¹ Entre autres https://twitter.com/mebis_bayern/status/

² Distributed Denial of Service (Attaque par déni de service) – Attaque informatique délibérée via l'inondation du serveur par un grand nombre d'ordinateurs

³ Nombre d'élèves en Bavière selon l'Office fédéral des statistiques : 1 252 674.

⁴ <https://de.wikipedia.org/wiki/OSI-Modell>

⁵ <https://www.mebis.bayern.de/infoportal/empfehlung/mebis-in-zeiten-von-corona/>

⁶ Vraisemblablement couche transport + session dans le modèle OSI

⁷ Software Defined Datacenter

⁸ Network Operation System

⁹ Software Defined Storage, Software Defined Network, Software Defined Computing

L'infrastructure des data centers – le matériel

Généralement, l'infrastructure informatique d'un data center comprend tous les processus et services non directement liés au traitement des données. Il s'agit de l'approvisionnement et de la distribution énergétique, de la climatisation, de la sécurité, de la gestion des bâtiments, des services, mais également du câblage relatif aux télécommunications¹⁰. Comment une infrastructure statique peut-elle alors être modularisée et dynamisée pour répondre aux exigences d'un data center dynamique ?

Malheureusement, dynamiser l'espace du data center relève quasiment de l'impossible. C'est pourquoi les opérateurs des data centers travaillent sur des concepts de plus en plus sophistiqués

afin de concevoir des espaces permettant une extension modulaire, laquelle va toujours de pair avec une densification des équipements. Concernant l'approvisionnement électrique, les systèmes d'alimentation modulaires sans interruption et les systèmes de distribution électrique modulaires et gérables avec suffisamment de redondance ont fait leurs preuves. Les systèmes de refroidissement requièrent une quantité de réserve appropriée afin de compenser facilement les pannes partielles et les pics de charge mais doivent également fonctionner régulièrement en mode économique. Cela signifie qu'il est préférable d'opter pour un contrôle intelligent et adapté aux besoins du volume d'air de refroidissement plutôt que de la vitesse de l'air.

Dynamisation du câblage ?

Le câblage semble présenter un cas particulier concernant les approches de dynamisation. Comment un câble peut-il devenir dynamique ? Le câblage d'un data center répond essentiellement aux exigences fixées par les normes

EN 50173 et EN 50600-2-4. Une architecture hiérarchique, telle qu'on l'a connue dans les data centers pendant de nombreuses années, présente certains inconvénients auxquels il convient de remédier. Par exemple :

- Les longs chemins de transmission impliquant la présence de nombreux commutateurs entre les serveurs entraînent des latences et des différences de latence.
- Les « goulots d'étranglement », c'est-à-dire la limitation des taux de transmission, quasiment déterminés par la conception même d'une architecture hiérarchique, conduisent à des pertes en matière de performance, même en cas de flux de données priorisés.
- Une extension du nombre de serveurs implique un câblage supplémentaire de grande ampleur.
- Malgré la redondance, la défaillance des switches de cœur¹¹ peut entraîner des pertes de performances considérables.

Le goulot d'étranglement, tout particulièrement, est à l'origine des problèmes liés à l'augmentation du trafic dans le data center. Depuis que les applications fonctionnent sur des machines virtuelles¹² entièrement détachées des serveurs dédiés, le trafic dit Est-Ouest, dont la conception

doit permettre d'éviter au maximum les blocages, augmente constamment. Des approches visant à créer des réseaux sans blocage ont été développées parallèlement à la mise en place des réseaux téléphoniques des années 1900 à 1950¹³.

¹⁰ voir EN-50600

¹¹ Switch de cœur : commutateur principal prenant en charge la majorité des transferts dans un réseau organisé hiérarchiquement.

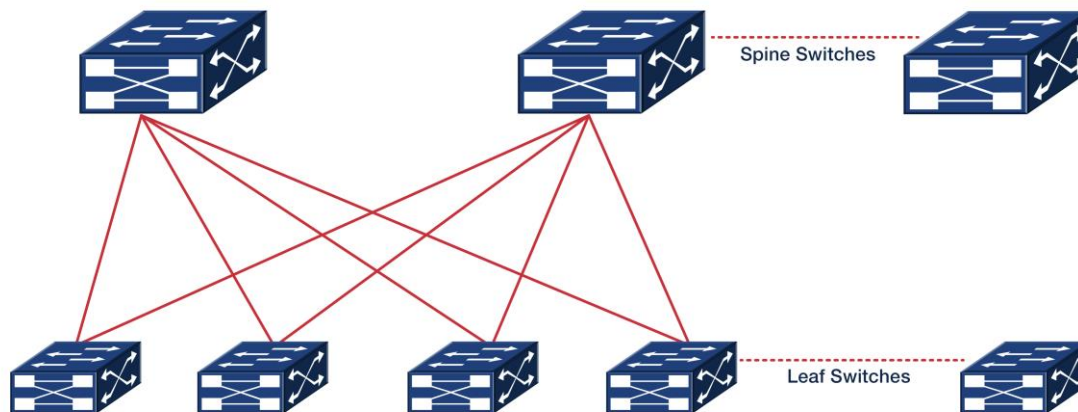
¹² MV

¹³ https://en.wikipedia.org/wiki/Clos_network

Des réseaux sans blocage : l'architecture Spine-Leaf

L'architecture Spine-Leaf, déjà largement utilisée, s'appuie sur l'idée du réseau sans blocage. Ici, chaque commutateur leaf est connecté à chaque commutateur spine. Dans la littérature, un tel

réseau entièrement maillé entre les commutateurs spine et leaf est souvent appelé mesh ou encore fabric.



Sur les graphiques, on reconnaît rapidement que le trafic d'un commutateur leaf à un autre ne

nécessite par principe qu'un seul bond (« hop »). Cette topologie présente les avantages suivants :

+

- Dans le concept maillé, le chemin des données est sélectionné au hasard.
- Le trafic des données est réparti uniformément dans la fabric.
- Si un lien existant entre un commutateur leaf et un commutateur spine menace d'être surchargé, un autre chemin vers un autre commutateur spine est utilisé de manière quasiment automatique, p. ex. en recourant à la LACP¹⁴.
- L'extensibilité à l'aide d'autres commutateurs leaf et, si nécessaire, d'autres commutateurs spine permet d'entretenir un réseau de fabric sans blocage adapté aux besoins.
- Le trafic de données entre les commutateurs leaf nécessite toujours un bond au maximum (spine). Les latences sont donc rapides et prévisibles, une condition indispensable pour les applications en temps réel, lesquelles deviennent de plus en plus importantes.

Les inconvénients d'une telle structure, par exemple les interruptions dues à des procédures

spanning tree¹⁵ involontaires, sont évitées grâce aux protocoles de routage ou à d'autres mesures.

¹⁴ Link Aggregation Control Protocol (couche OSI 2)

¹⁵ Fonction automatique Ethernet nécessitant beaucoup de temps et permettant de supprimer les boucles.

Réseaux superposés

Les avantages mentionnés ci-dessus constituent la base de différentes applications appelées d'une manière générale « réseaux superposés ».

L'exemple le plus simple et le plus connu de réseau superposé est vraisemblablement celui des VPN souvent utilisés dans le cadre des connexions de télétravail. Mais dans les data centers également, des réseaux séparés sont de plus en plus souvent superposés sur des réseaux existants de manière quasiment virtuelle. Les mots clés sont ici VXLAN¹⁶, NVGRE¹⁷, LISP¹⁸, etc. Cela montre clairement que, pour les architectures modernes Spine-Leaf, le soutien des couches supérieures est indispensable. Il est

recommandé d'utiliser au minimum les fonctionnalités de la couche 3 (c'est-à-dire les fonctions de routage dans les commutateurs).

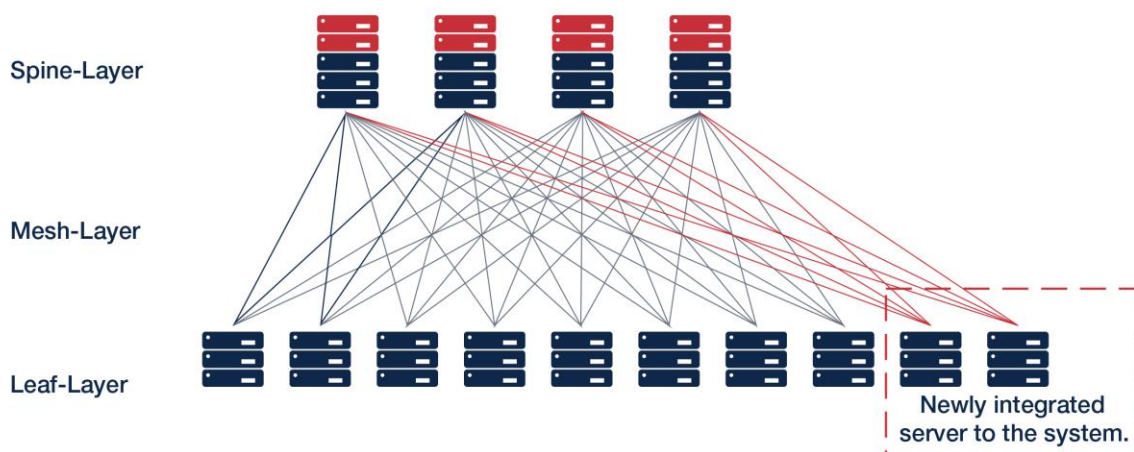
Une architecture Spine-Leaf permet de générer de tels réseaux superposés d'une manière particulièrement efficace et dynamique, lesquels font alors partie intégrante des SDN¹⁹. En dehors des réseaux superposés, les possibilités de contrôle flexible de la vitesse du port et de l'attribution des canaux dans les émetteurs-récepteurs²⁰ font également partie des fonctions utilisées par les SDN. Cela crée les bases permettant d'adapter les vitesses du port définies selon la norme IEEE²¹ 802.3²² aux besoins actuels.

Le câblage en fibre optique de la Spine-Leaf

Le câblage en fibre optique dans les data centers doit tenir compte des points mentionnés ci-dessus et fournir le plus grand nombre possible de connexions optiques parallèles pour des débits de données élevés. Ainsi, une architecture Spine-Leaf requiert la pose d'un câble en fibre optique entre chaque commutateur leaf et chaque commutateur spine. Il en résulte un câblage incroyablement compliqué qui, dans le pire des cas, rappelle les débuts des data centers et les montages de câbles entremêlés sur les double-planchers s'élevant sur plusieurs mètres. Examinons donc de plus près une installation

typique : dans le data center, les commutateurs switch sont généralement conçus en tant que commutateurs ToR²³ ou EoR²⁴, alors que les commutateurs spine occupent souvent une place centrale.

Des connexions en fibre optique reliant les serveurs à ces commutateurs soutiennent les débits binaires maximums des émetteurs-récepteurs des serveurs ainsi que des commutateurs. De la même manière, tous les ports des commutateurs leaf sont reliés à tous les ports des commutateurs spine.



© FiberCon GmbH

¹⁶ Virtual Extensible LAN (couche 2 transmission des données dans UDP de couche 4)

¹⁷ Virtualisation du réseau avec Generic Routing Encapsulation

¹⁸ Locator/Identifier Separation Protocol

¹⁹ Software Defined Networks

²⁰ Unité d'émission et de réception (ici pour la transmission optique de données)

²¹ Institute of Electrical and Electronics Engineers.

²² Groupe de travail IEEE pour Ethernet

²³ Top of Rack

²⁴ End of Row

On constate déjà que l'utilisation de câbles individuels pour une connexion point à point avec des « cordons directement raccordés »²⁵ est particulièrement compliquée, non seulement au niveau de l'installation, mais également au niveau de la maintenance. Un système de câblage consolidé avec des répartiteurs peut permettre de simplifier

ce travail. En outre, conformément à la norme EN 5060-2-4, cette approche correspond à la classe de disponibilité la plus basse.

Pour un système de câblage selon l'architecture Spine-Leaf, il existe donc plusieurs alternatives :

1. Utilisation de panneaux de brassage et de trunks classiques reliés à un point de commutation²⁶ central entre les commutateurs spine et leaf.

La répartition des connexions optiques parallèles des commutateurs spine vers les raccords duplex

sur les commutateurs leaf peut être réalisée via des modules MTP®/MPO.

+

- La couche de mailles peut être brassée librement.
- Les expansions ont lieu au niveau du point de commutation/point de brassage.

-

- En raison du grand nombre de raccords, le point de brassage requiert une extension massive.
- Le point de brassage devient alors une source d'erreur supplémentaire.
- En raison du grand nombre de raccords, l'acheminement des câbles devient particulièrement complexe.
- La documentation relative à la répartition doit être conçue de manière encore plus élaborée.

2. Utilisation de trunks spéciaux pour le maillage Spine-Leaf :

+

- L'utilisation de modules MTP®/MPO n'est plus nécessaire, les câbles s'adaptant et se mélangeant directement, par exemple de MTP®/MPO²⁷ à LC duplex²⁸.

-

- De tels câbles représentent des points de défaillance uniques très coûteux.
- Le concept est particulièrement complexe et le maillage peut difficilement être documenté directement sur le câble.
- Une modification ultérieure de la topologie quelle qu'elle soit est impossible.
- Les expansions sont difficiles à mettre en œuvre.

²⁵ DIN EN 50600-2-4 paragraphe 7.2

²⁶ DC/DZ DIN EN 50600-2-4 paragraphe 7.2.3

²⁷ Multiple-Fiber Push-On/Pull-Off (IEC61754-7:2014) / Multifiber Termination Push-On (USCONEC)

²⁸ LC duplex : connecteur à 2 fibres très couramment utilisé

3. Utilisation de cassettes spécialement conçues pour la mise en œuvre de l'architecture Spine-Leaf. Cette approche implique l'utilisation de « fiber shuffles » (mélanges fibreux) :

+

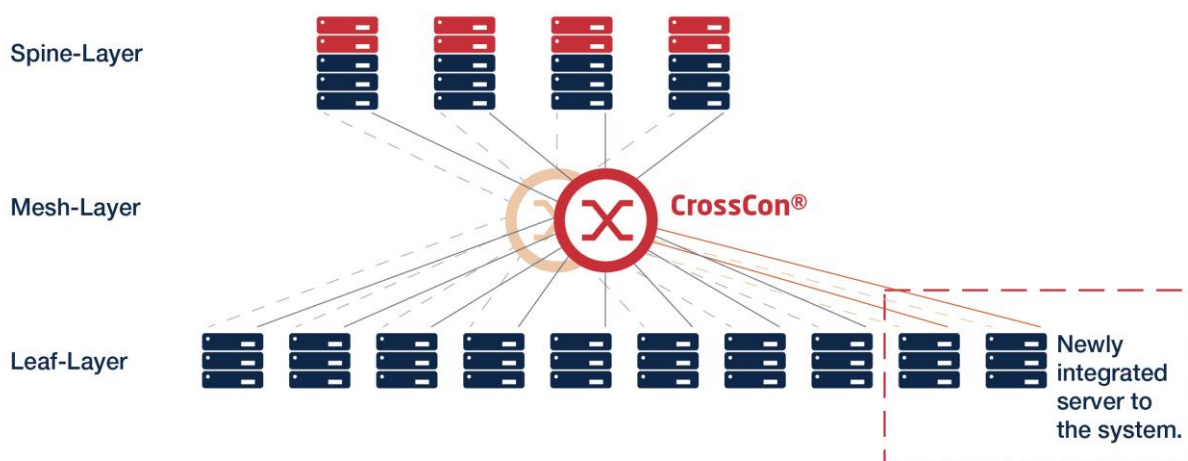
- Câblage conforme à la norme EN 50600 et possible dans toutes les classes de disponibilité.
- La fabric est un module facilement échangeable si nécessaire.
- Possibilités d'expansions simples.
- Ouvre des voies pour d'autres modules Spine-Leaf en fonction de la technologie de transmission choisie pour la couche 1.
- Manipulation simple.
- Facilement documentable.

-

- Atténuation légèrement plus élevée que pour les connexions point à point.



Cassette CrossCon® MTP®/MPO 4xSR4



© FiberCon GmbH

Résumé

Pour le contrôle intelligent de leur infrastructure, les Software Defined Data centers nécessitent une affectation dynamique des ressources matérielles aux machines virtualisées respectives (ou conteneurs). Il s'agit du seul moyen de garantir l'efficacité et la disponibilité maximales des processus.

Les fondements d'une infrastructure de data center moderne doivent permettre de répondre aux exigences élevées découlant des applications à débit binaire maximal tout en garantissant des temps de latences minimaux. Cela inclut un concept de câblage Spine-Leaf adaptable de manière modulaire.

À propos de Rosenberger OSI :

Depuis 1991, Rosenberger Optical Solutions & Infrastructure (Rosenberger OSI) est un expert dans les systèmes de câblage innovants en fibre optique et cuivre. Ces solutions se retrouvent dans un grand nombre de secteurs où de grands volumes de données doivent être échangés rapidement et de manière sécurisée : Data Centers, Réseaux Locaux, Telecom et Industrie.

Rosenberger OSI emploie un peu plus de 740 personnes et fait partie du groupe mondial Rosenberger depuis 1998. Rosenberger est une société allemande de 11000 personnes et est présente dans 32 pays. Ses solutions hautes fréquences, haute tension, liaisons fibre optique et cuivre sont reconnues pour leur fiabilité et leurs performances.

Pour plus d'informations, visitez : www.rosenberger.com/osi

Rosenberger

Rosenberger-OSI GmbH & Co. OHG

Optical Solutions & Infrastructure | Endorferstr. 6 | 86167 Augsburg | GERMANY | Phone: +49 821 24924-0 | info-osi@rosenberger.com | www.rosenberger.com/osi

Rosenberger® is a registered trademark of Rosenberger Hochfrequenztechnik GmbH & Co. KG. All rights reserved. © Rosenberger 2020

Transfer to third party only by authority of Rosenberger-OSI GmbH & Co. OHG- All rights reserved

Created: 16.06.2020
Valid since: 16.06.2020