

Transparente Spine-Leaf-Architekturen

Kreuzverbindungsmodule erleichtern den Umstieg auf East-to-West-Routing

Die Virtualisierung treibt RZ-Planer zu neuen Routing-Strategien. An die Stelle traditioneller dreischichtiger Architekturen treten zunehmend zweischichtige Spine-Leaf-Topologien. Die Umsetzung der hochkomplexen Verbindungen im Mesh-Layer lässt sich durch Kreuzverbindungsmodule deutlich vereinfachen.

Im Rechenzentrumsbereich ist schon seit geraumer Zeit eine steigende Anzahl virtualisierter Applikationen zu verzeichnen. Vor allem Cloud Computing, SaaS und IP Storage treiben diese Entwicklung voran. Damit aber wächst in den Datacentern auch der Bedarf an hochperformanten Gigabit-Ethernet-Verbindungen. Lange sind extern und intern geroutete Verbindungen im Verhältnis 4:1 gestanden – nun kehrt sich dieses Verhältnis um: Bedingt durch die massive Nutzung von Plattformvirtualisierungen – zum Beispiel auf Basis von Docker – kommen auf eine externe mittlerweile vier interne Verbindungen. Ein Großteil des Datenverkehrs findet also zukünftig innerhalb der RZ-eigenen Infrastruktur zwischen virtuellen Servern statt. Das erfordert eine Anpassung der Routing-Strategien. Ziel der RZ-Planer muss künftig sein, die hochgradig parallelisierte Infrastruktur mit laufzeitoptimierten Datenverbindungen im Bereich von 10 bis 40GbE auszustatten.

East-to-west als Lösungsansatz

Gegenwärtig werden die Bottlenecks solcher Datenverbindungen unter anderem durch eine traditionelle dreischichtige Router-Architektur gebildet. Dieses vertikale Routing-Schema wurde ursprünglich für das externe Routing konzipiert und besteht aus den Ebenen der Core-, Aggregations- und Access-Router. Eine solche klassische Client-Server-Topologie wird auch als North-to-South-Routing bezeichnet.

Da angesichts der wachsenden Anzahl virtualisierter Server jedoch zunehmend horizontale Routing-Verbindungen notwendig werden, erfordert die Optimierung der Datenlaufzeit eine andere Router-Topologie. An die Stelle der klassischen dreischichtigen tritt deshalb eine zweischichtige Spine-Leaf-Architektur: Die Router werden direkt über eine voll vernetzte Kreuzverbindung angebunden, um Zwischenebenen in Form des klassischen Aggregation Layers zu vermeiden. Für diese Topologie hat sich der Name East-to-West-Routing etabliert.

Beim East-to-West-Routing sind die Router direkt über faseroptische GbE-Module in Form von SFP- und QSFP-Transceivern verbunden. Auf der Leaf-Seite sind derzeit Transceiver für Datenraten von 0,1 bis 25 GBit/s, auf der Spine-Seite Transceiver für Raten von 10 bis 100 GBit/s gängig. Aufgrund dieser speziellen Architektur sind die Router jeweils nur einen Hop voneinander entfernt – was die Datenlaufzeit deutlich verkürzt –, und es entsteht innerhalb des Routing-Schemas eine hohe An-

zahl paralleler Pfade. Neuere Router-Protokolle wie TRILL (Transparent Interconnection of Lots of Links) oder SPB (Shortest Path Bridging) sollen diese Parallelpfade nutzbar machen und ältere Konzepte wie STP (Spanning Tree Protocol) ablösen. Gerade im Bereich IP Storage erwartet man sich hiervon signifikante Bandbreiten- und Latenzgewinne.

Module für Kreuzverbindungen

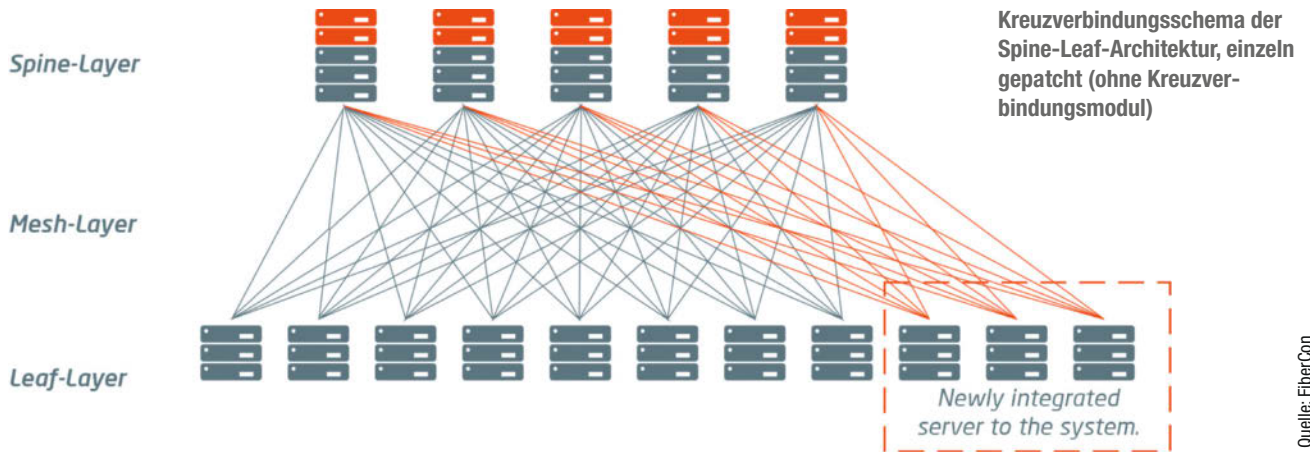
Bei Spine-Leaf-Architekturen besteht die besondere Herausforderung nun darin, diese hohe Anzahl paralleler optischer Verbindungen zwischen den einzelnen Routern – den sogenannten Mesh-Layer – lückenlos und zuverlässig zu realisieren. Eine East-to-West-Topologie setzt für n beteiligte Router bzw. Server genau n^2 faseroptische Verbindungen voraus. Das führt bei wachsender Anzahl von Routern zu immer komplexeren Leitungsverbindungen. Eine volle Kreuzverbindung von acht Routern auf acht Server erfordert genau $8^2 = 64$ Verbindungen – einschließlich der optischen Aktivkomponenten in Gestalt von SFP-Transceivern. Diese Kreuzverbindungen einzeln zu verlegen, ist äußerst aufwendig und kann insbesondere dann zu Schwierigkeiten führen, wenn eine rasche Skalierung der Datacenter-Infrastruktur erforderlich ist.

Einige Anbieter sind deshalb dazu übergegangen, die faseroptischen Kreuzverbindungen im Mesh-Layer in Form integrierter Kreuzverbindungsmodule bereitzustellen. Solche Module erleichtern die Erweiterung der Infrastruktur deutlich, da man nicht mehr jede Kreuzverbindung einzeln umsetzen muss, sondern zusätzliche Router bzw. Server lediglich per Trunkkabel an das Kreuzverbindungsmodul anschließt. Bei den von FiberCon entwickelten und von der euromicron-Tochter LWL-Sachsenkabel gefertigten CrossCon-Modulen etwa können beliebige Kreuzverbindungsweiten erzeugt werden. Möglich wird das durch ein neuartiges dreidimensionales Steckschema der faseroptischen Verbindungen (mit

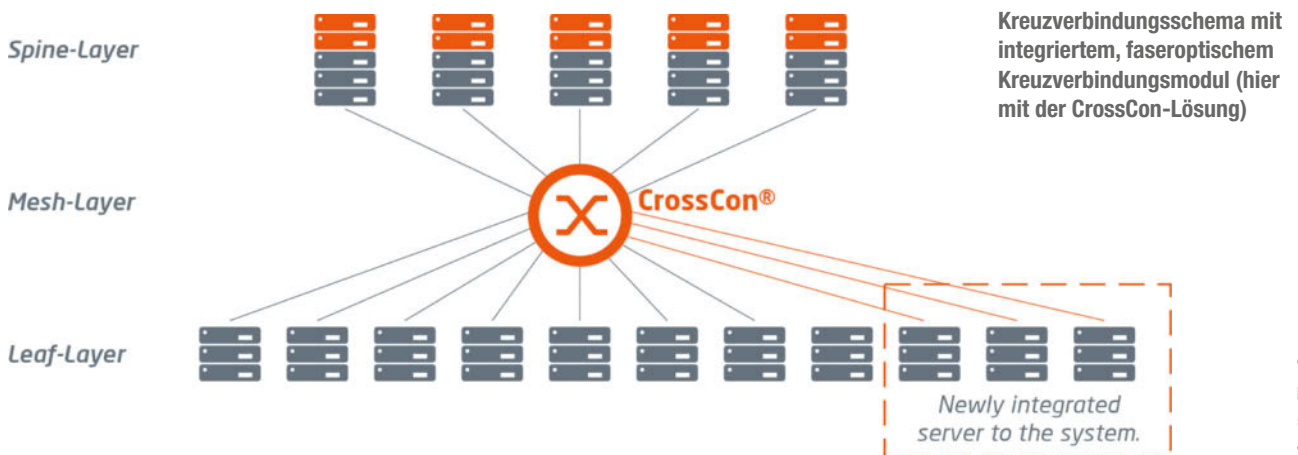
Das Kreuzverbindungsmodul CrossCon mit 8×8 faseroptischen URM-Verbindungen (MMF).



Quelle: FiberCon



Quelle: FiberCon



Quelle: FiberCon

integrierten Spalten und Zeilentauch innerhalb des Kreuzverbindungsmoduls). Die zuvor zweidimensionale Kreuzverbindung wird durch die Erweiterung auf drei Dimensionen „entflochten“ und ermöglicht eine strukturierte Umsetzung des Spine-Leaf-Routing-Schemas. Das heißt konkret, dass ein sogenanntes Shuffling faseroptischer Datenleitungen stattfindet – etwa in Form von ein- und ausgehenden Faserbündchen (Ribbon Fibers). Auf diese Weise sind aktuell gekreuzte Kanalbreiten von bis zu 32×32 Fasern möglich.

Bye bye, Patchkabel!

Durch solche Module werden Kreuzverbindungen jedoch nicht nur entflochten und klar strukturiert, sondern es wird, wie ein Blick auf standardisierte Verkabelungsinfrastrukturen in Rechenzentren zeigt, auch die Verbindung der einzelnen Faserstrecken vereinfacht. Diese wurde bei Kreuzverbindungen von Spine-Leaf-Servern bisher über Patchfelder im Bereichs- oder Hauptverteiler realisiert. Spine- und Leaf-Geräte wurden dabei gleichermaßen am Patchfeld abgebildet und anschließend mit Patchkabeln verbunden.

Bei Modulsystemen erübrigt sich hingegen der Einsatz von Patchfeldern und Patchkabeln, da die Kreuzverbindungsmodule diese Aufgabe selbst übernehmen. Die Verbindung zum Spine- oder Leaf-Gerät

wird dann einfach per Trunkkabel hergestellt. Das spart Zeit bei der Erstinstallation und erleichtert zudem – ein weiterer großer Vorteil – die Skalierung des Systems. So lassen sich beispielsweise Mesh-Layer-Bausteine definieren, die als Standard in die Topologie integrierbar sind und schnelle, transparente Erweiterungen ermöglichen.

Bei der Umsetzung von Spine-Leaf-Architekturen sind Kreuzverbindungsmodule eine enorm praktische Vereinfachung. Wo Verbindungen zwischen virtualisierten Servern hergestellt werden müssen, sind sie ein wichtiger Baustein bei der Realisierung einer strukturierten RZ-Verkabelung. Wer sie einsetzen will, muss sich allerdings von der Vorstellung verabschieden, dass Patchkabel universal notwendig seien. Das widerspricht zwar zunächst allen etablierten und standardisierten Verfahrensweisen, ermöglicht dafür aber den Aufbau einer enorm vereinfachten und dadurch leicht zu überblickenden Infrastruktur. Hinzu kommt, dass durch die klar definierten Verbindungswege eine klassische Fehlerquelle entfällt, was nicht zuletzt den Dokumentationsaufwand reduziert – ein Vorteil, dessen effizienzsteigernde Wirkung nicht unterschätzt werden sollte.

*Philipp Nölle,
FiberCon GmbH
Kai Wirkus,
LWL-Sachsenkabel GmbH*