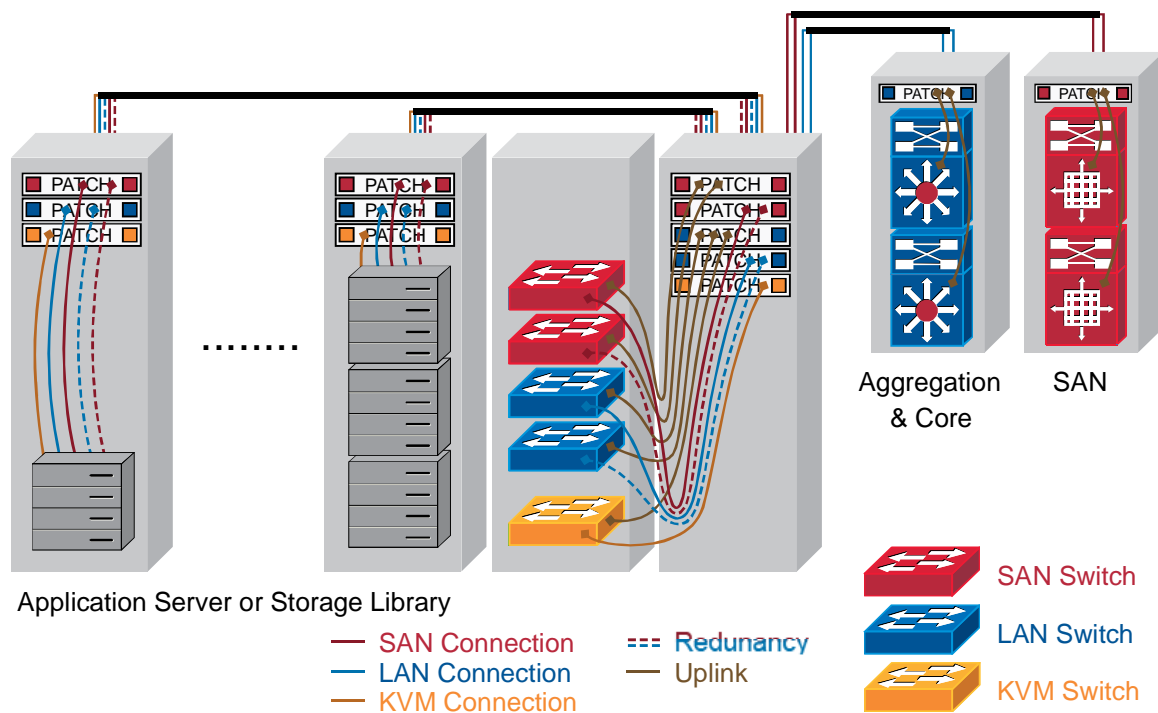


White Paper



Verkabelungsarchitekturen im Data Center

End of Row versus Top of Rack



Convincing cabling solutions

Verkabelungsarchitekturen im Data Center – End of Row versus Top of Rack

INHALTSVERZEICHNIS

1. Einführung.....	3
2. Netzwerkhierarchie und Verkabelungsarchitektur	4
2.1. Netzwerkhierarchie	4
2.2. Logische/physikalische Netzwerkstruktur	6
2.3. Verkabelungsarchitekturen	6
3. End of Row vs. Top of Rack.....	7
3.1. End of Row (EoR), Dual End of Row & Middle of Row (MoR).....	7
3.2. Top of Rack (ToR).....	8
3.3. Konsequenzen bei der Wahl der Verkabelungsarchitektur.....	9
4. Empfehlung	11
5. Zusammenfassung.....	12
6. Quellen und weitere Informationen	12

© Copyright 2011 Reichle & De-Massari AG (R&M). Alle Rechte vorbehalten.

Weitergabe und Vervielfältigung dieser Publikation oder von Teilen daraus sind, zu welchem Zweck und in welcher Form auch immer, ohne die ausdrückliche schriftliche Genehmigung durch Reichle & De Massari AG nicht gestattet. In dieser Publikation enthaltene Informationen können ohne vorherige Ankündigung geändert werden. Die Erstellung dieses Dokuments erfolgte mit grösstmöglicher Sorgfalt, es enthält den zum Zeitpunkt der Erstellung aktuellen technischen Stand.

Die Verkabelungsarchitektur im Data Center entscheidet über dessen Zukunftsfähigkeit

Das Web 2.0 hat Applikationen hervorgerufen, die grosse und kaum vorhersagbare Leistungsanforderungen an die I/O-Schnittstellen der Netzwerke stellen.

Im Access-Bereich der Data Center, wo die Switches die Netzwerkschnittstellen der Server für den Backbone-Anschluss bündeln, stehen die Verantwortlichen in Data Center mit konsequent virtualisierten Umgebungen zunehmend vor der schwierigen Aufgabe, die Voraussetzung für einen einwandfreien Betrieb der auf Multi-Core-Servern laufenden Anwendungen mit ihrem grossen, nur schwer planbaren Bedarf an Übertragungsraten und Speicher zu schaffen.

Umso wichtiger ist es, eine Verkabelungsarchitektur bereitzustellen, welche als Fundament die grösstmögliche Flexibilität, Leistungsfähigkeit und Skalierbarkeit (vertikal sowie horizontal) des Netzwerkes erlaubt.

Anwendung:	Data Center Cabling
Technologie:	Kupfer- und Glasfaserverkabelung, Verteilertechnik
Format:	White Paper
Themen:	Trends der Netzwerkhierarchien und deren Einfluss auf die Verkabelungsarchitekturen in Data Center
Ziel:	Verstehen der Vor- sowie Nachteile von EoR/MoR- & ToR-Verkabelungsarchitekturen, Erkennen, dass nur horizontal und vertikal durchgängige Strukturen die notwendige Flexibilität für zukünftige Systemarchitekturen bieten können
Zielgruppe:	Planer, Installateure, Netzwerk-Verantwortliche, Betreiber von Data Center
Autor:	Roland Diethelm
Erschienen:	November 2011

Dieses White Paper informiert über die Verkabelungsarchitekturen in Data Center und deren Einfluss auf die Migrationsfähigkeit des Netzwerkes.

1. Einführung

Mit dem Datenverkehr von multimedialen Inhalten, der I/O-Konsolidierung, der Konvergenz von Netzwerk- sowie Storage-bezogenen Daten, der Automatisierung von IT-Prozessen, der Integration immer intelligenterer Applikationen in die Netze, der Entwicklung weiterer Echtzeitanwendungen wie High Frequency Trading, High Performance Computing (HPC) und der steigenden Zahl an Unternehmen, die den Einsatz von Virtualisierungslösungen planen, stehen Data Center vor dem Problem, die notwendigen I/O-Leistungen nicht im Voraus planen zu können. Dies führt häufig zu Engpässen im Access-Bereich. Oft werden Netzwerke unter diesen Bedingungen an der I/O-Schnittstelle der Server überlastet, was zu Verlusten von Netzwerkpaketen und hohen Latenzzeiten für die Applikationen führt.

Diese Engpässe bei der Datenübertragung gilt es zu vermeiden, um den diskontinuierlichen Leistungsanforderungen zu begegnen. Die Hersteller von Switches bieten hierfür Lösungen zur Leistungsoptimierung an, wie grössere Bufferkapazitäten, höhere Backplane-Durchsatzraten, Unterstützung neuester Technologien wie dem Data Center Bridging (DCB), welche die verlustfreie Übertragung von verschiedenen Protokollen über Ethernet ermöglicht, sowie weitere Funktionalitäten.

Alle diese Lösungsansätze zielen auf eine möglichst schnelle, verlust- und verzögerungsfreie Weiterleitung der Daten durch die Switches ab. Kurze Kommunikationswege, im Sinne von wenigen Switches innerhalb einer Übertragungsstrecke, begünstigen eine effiziente Übertragung der Daten jedoch am effektivsten, da dadurch die Vermittlungsverzögerung in der Summe massiv minimiert wird.

Hierfür ist eine einheitliche, zukunftsorientierte Verkabelungsarchitektur erforderlich, die neben ihrer Flexibilität über grosszügige Leistungsreserven verfügt. Ebenso sollte eine ausreichende Anzahl von Verbindungen in der Horizontalverkabelung zur Verfügung stehen, um auch konzeptionelle Anpassungen vornehmen zu können, sei es zur Verflachung der Netzwerkhierarchie, zur horizontalen Skalierung (scale out) oder zur Leistungssteigerung mit zusätzlichen Aggregationen (Leitungsbündelungen) im Serverbereich.

2. Netzwerkhierarchie und Verkabelungsarchitektur

2.1. Netzwerkhierarchie

Klassische Netzwerkhierarchien in Data Center sind dreistufig aufgebaut, weshalb sie auch "Three Tier Network" genannt werden. Sie unterteilen sich in Core-, Aggregation- und Access-Bereich.

Hierarchische Netzdesigns haben viele Vorteile:

- Skalierbarkeit
- Leistungsfähigkeit
- Administrierbarkeit
- Redundanz
- Sicherheit
- Wartungsfreundlichkeit

Da hierarchische Netze ihrem Wesen nach modular und einfach zu skalieren sind, sind sie auch sehr wartungsfreundlich. Bei anderen Netzwerktopologien wird die Wartung umso komplizierter, je grösser das Netzwerk wird. Ausserdem gilt bei bestimmten Netzdesignmodellen ein festgelegter Grenzwert für das Wachstum des Netzes; wird dieser überschritten, so wird die Wartung zu kompliziert und zu teuer.

Beim hierarchischen Designmodell sind die Switch-Funktionen für jeden Bereich definiert, was die Auswahl des korrekten Switches vereinfacht. Das Hinzufügen von Switches in einem Bereich bedeutet jedoch nicht, dass es in einem anderen nicht zu einem Engpass oder einer anderen Einschränkung kommen kann.

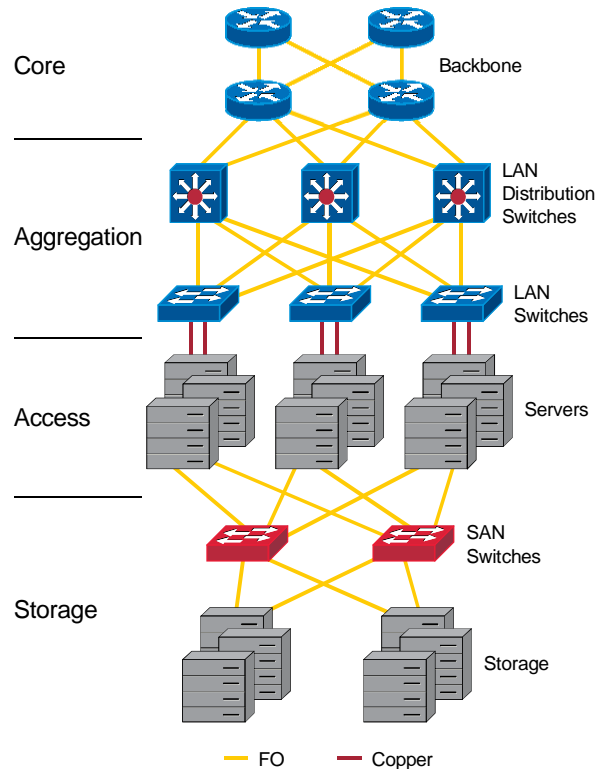


Abb. 1: Klassische Netzwerktopologie im Data Center

Beim Netzwerkdesign muss zuerst der Netzwerkdurchmesser berücksichtigt werden. Zwar wird ein Durchmesser traditionell als Längenwert angegeben, doch im Fall der Netzwerktechnik muss dieser Parameter die Grösse eines Netzes mittels der Anzahl der Geräte definieren. Als Netzwerkdurchmesser wird dabei die Anzahl der Geräte bezeichnet, die ein Datenpaket passieren muss, um seinen Empfänger zu erreichen. Kleine Netzwerke gewährleisten demzufolge eine niedrigere und vorsehbare Latenz zwischen den Geräten. Unter der Latenz versteht man diejenige Zeit, die ein Netzwerkgerät benötigt, um ein Paket oder ein Frame zu verarbeiten.

Denn jeder Switch im Netzwerk muss zuerst die Ziel-MAC-Adresse des Frames bestimmen, diese in seiner MAC-Adresstabelle nachschlagen und das Frame dann über den entsprechenden Port weiterleiten. Auch wenn der gesamte Prozess nur den Bruchteil einer Sekunde dauert, summieren sich die Latenzen, falls ein Ethernetframe zahlreiche Switches passieren muss.

Über „Data Center“-Switches werden meistens sehr grosse Datenmengen übertragen, da es sich hier um die Kommunikation sowohl Server/Server-Daten als auch Client/Server-Daten handelt. Aus diesem Grund sollten Switches, die für Data Center vorgesehen sind, eine höhere Leistung und niedrigere Latenzzeiten (Low Latency Switches) bieten als solche, die sich bei den Endgeräten befinden.

Servervirtualisierungen sorgen für zunehmende Komplexität. Auf einem physischen Rechner lassen sich beispielsweise zehn virtuelle Server betreiben, die zudem flexibel je nach Bedarf von Hardware zu Hardware verschiebbar sind. Wo also ein Netzwerk zuvor den Datenverkehr von 1'000 Servern zu verwalten hatte, sind es jetzt 10'000 virtuelle Maschinen, die sich zudem auch noch in Bewegung befinden.

Mit dieser Komplexität haben klassische dreistufig aufgebaute Netzwerke, welche das Data Center seit den späten 1990er-Jahren beherrschen, zunehmend Mühe. Der Ruf nach einer flachen Architektur wird laut, bei der das Netzwerk einem Gewebe (englisch: fabric) von gleichberechtigten Knoten gleicht. Statt reinen Punkt-zu-Punkt-Verbindungen sind auch Querverbindungen zwischen den Nodes möglich, was die Leistungsfähigkeit eines solchen Netzes erhöht.

Bei der Realisierung von Cloud Computing-Umgebungen und Storage-Lösungen ist die durch eine dreistufige Switching-Architektur hervorgerufene Verzögerung nicht mehr akzeptabel. Ausserdem ist die zur Verfügung stehende Bandbreite durch unzureichende Switch-Konzepte oftmals überbucht, was zu einem Flaschenhals beim Uplink führt. Die verschiedenen Anwendungen kämpfen dadurch permanent um eine begrenzte Bandbreite. Dies wirkt sich in einer Erhöhung der Reaktionszeit bzw. einer Verschlechterung der Services aus.

Echtzeit-Applikationen, wie beispielsweise Voice over IP (VoIP), Unified Communication (UC), Video Conferencing, Video on Demand (VoD) und insbesondere High Frequency Trading (elektronischer Finanzhandel in extrem schneller Abfolge) erfordern eine berechenbare Leistung, Verzögerung und einen definierten Quality of Service (Qos) für den Broadcast-, Multicast- und Unicast-Verkehr.

Cloud Computing und virtualisierte Data Center mit einer optimierten Switching-Architektur garantieren eine Ende-zu-Ende-Verzögerung von weniger als 10 Mikrosekunden. Um diese kurzen Übermittlungszeiten erreichen zu können, müssen so viele Schaltstufen im Netz wie möglich eliminiert werden. Bei der Server-Virtualisierung mittels Blade- oder Rack-Server, übernehmen die Server selbst die Rolle des Access Switches.

In einer zweistufigen Netzwerkhierarchie kann auf den Spanning-Tree-Algorithmus verzichtet werden. Bisher bestimmt dieser Mechanismus die aktiven Pfade zwischen dem Data Center und den Endgeräten. Falls ein Netzwerk mehrere Wege zwischen zwei Endpunkten bereitstellt, schaltet Spanning Tree alle Ports ab, die zum selben Ziel führen könnten. Dieses Verhalten beeinflusst die Skalierbarkeit von Netzen und führt zu einer wachsenden Ineffizienz, da ein wesentlicher Teil der theoretisch verfügbaren Bandbreite nicht zur Verfügung steht. Die Anforderungen an ein modernes Netzdesign, welches eine schnelle und verlustfreie Anbindung der virtualisierten Data Center bei geringsten Verzögerungen garantieren soll, sind mit der Nutzung des Spanning-Trees nicht mehr vereinbar.

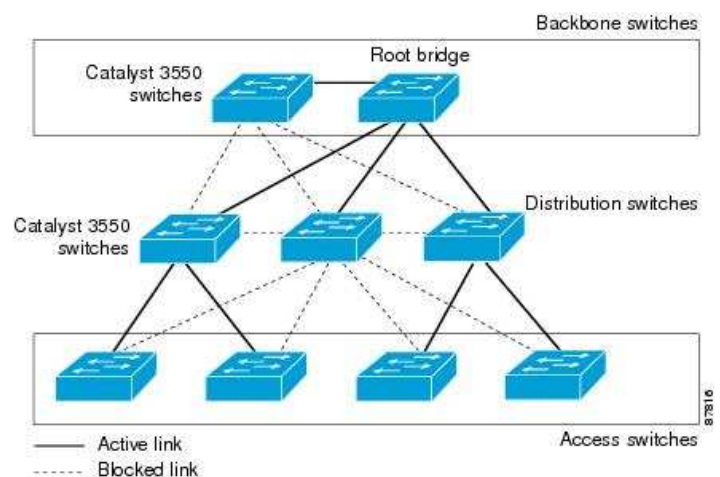


Abb. 8: Spanning-Tree im Einsatz (Quelle: Cisco)

Seit einigen Jahren arbeitet die IEEE (Institute of Electrical and Electronics Engineers) daran, Ethernet mit InfiniBand-ähnlichen Eigenschaften zu erweitern, um damit das neue Converged Enhanced Ethernet (CEE) zu bilden. Dieser Standard würde es möglich machen, Ethernet für horizontal skalierbare Data Center zu nutzen. Es gibt Hersteller, die bereits heute Ethernet-Lösungen anbieten, welche ohne Leistungseinbussen beliebig horizontal skalierbar sind. Mehrfachwege werden dabei parallel benutzt, so dass die verfügbare Bandbreite mit der Anzahl paralleler Verbindungen linear wächst. Selbst vermaschte Topologien sind möglich.

Marktforscher prognostizieren einen erheblichen Bedarf an 40- bzw. 100-Gigabit-Ethernet-Verbindungen in Data Center. In vielen Unternehmen werden bereits mit Hilfe der Link Aggregation-Funktion mehrere parallele 10 GBit/s-Verbindungen zu einem gemeinsamen Verbindungskanal mit höherer Geschwindigkeit gebündelt. Auch wenn die 40/100G-Ethernet-Produkte noch nicht von vielen Herstellern angeboten werden, sehen die Experten die Obergrenze des Durchsatzes noch lange nicht erreicht. Für das Jahr 2015 sagen sie deshalb bereits die Verfügbarkeit eines Terabit-Ethernet-Standards voraus.

2.2. Logische/physikalische Netzwerkstruktur

Den einzelnen Bereichen im Data Center sind bestimmte Funktionen zugeordnet (siehe auch Application Note: Data Center Standards). Das physikalische Layout basiert auf der logischen Netzwerkstruktur.

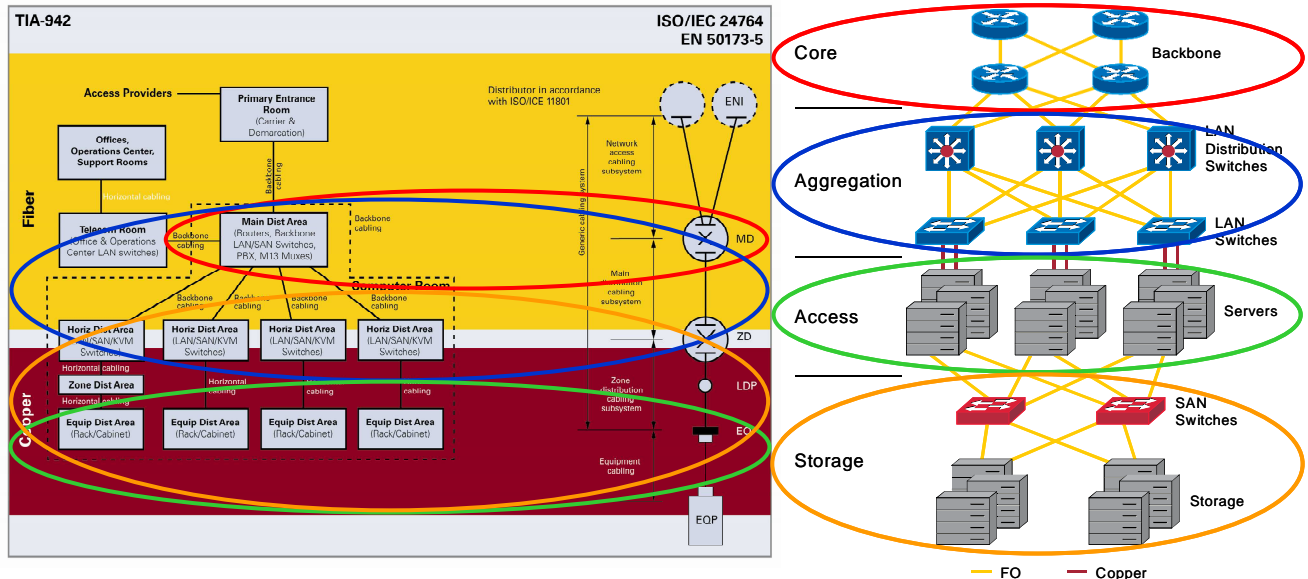


Abb. 2: Logische Abbildung der dreistufigen Netzwerkhierarchie im Data Center Layout

Ein zukunftsweisendes Data Center Layout ermöglicht es mit den Anforderungen zu wachsen. Es passt sich flexibel an neue Bedürfnisse an, gleichwohl welche Verkabelungsarchitektur gewählt wurde.

Die Festlegung der Verkabelungsarchitektur hat, neben der Leistungsfähigkeit und Skalierbarkeit des Netzwerkes, auch Einfluss auf dessen Verfügbarkeit (redundante Übertragungsstrecken) sowie auf die Anordnung der Racks, resp. der Rackreihen. Die Platzierung der Netzwerkkomponenten ist ein Gebiet mit hohem Optimierungspotenzial.

2.3. Verkabelungsarchitekturen

Verschiedene Verkabelungsarchitekturen werden seitens der Switch-Hersteller propagiert und kommen zur Anwendung. Jede dieser Varianten hat seine Vor- und Nachteile. Dabei gibt es zwei grundlegende Prinzipien, welche verfolgt werden und unterschieden werden können.

- End of Row (EoR), Dual End of Row & Middle of Row (MoR)
- Top of Rack (ToR)

Nicht alle Konzepte unterstützen eine Verflachung der Netzwerkhierarchie oder eine horizontale Skalierung.

3. End of Row vs. Top of Rack

3.1. End of Row (EoR), Dual End of Row & Middle of Row (MoR)

In den nachfolgend aufgezeigten Verkabelungsarchitekturen sind die Server über jeweils 3 Anschlüsse eingebunden. Bei geforderter Redundanz sind 5 Verbindungen je Server vorgesehen (+ 1 für LAN, +1 für SAN).

- **LAN:** Ethernet über Kupfer- oder Glasfaserkabel (1/10/40/100 Gigabit)
- **SAN:** Fibre Channel (FC) über Glasfaserkabel
- **KVM:** Keyboard/Video/Mouse-Signalübertragung über Kupferkabel

Bei einer **End of Row** Architektur wird eine Gruppe Serverracks von einem oder mehreren Switches aus versorgt. Von den Switches werden sternförmig die Datenkabel (Kupfer und/oder Glasfaser) zu den einzelnen Serverracks installiert. Pro Server sind dies 3 Kabel (5 bei redundanten LAN- und SAN-Verbindungen).

Bei 40 1HE-Rackservern werden somit 120 oder 200 Datenkabel pro Rack benötigt.

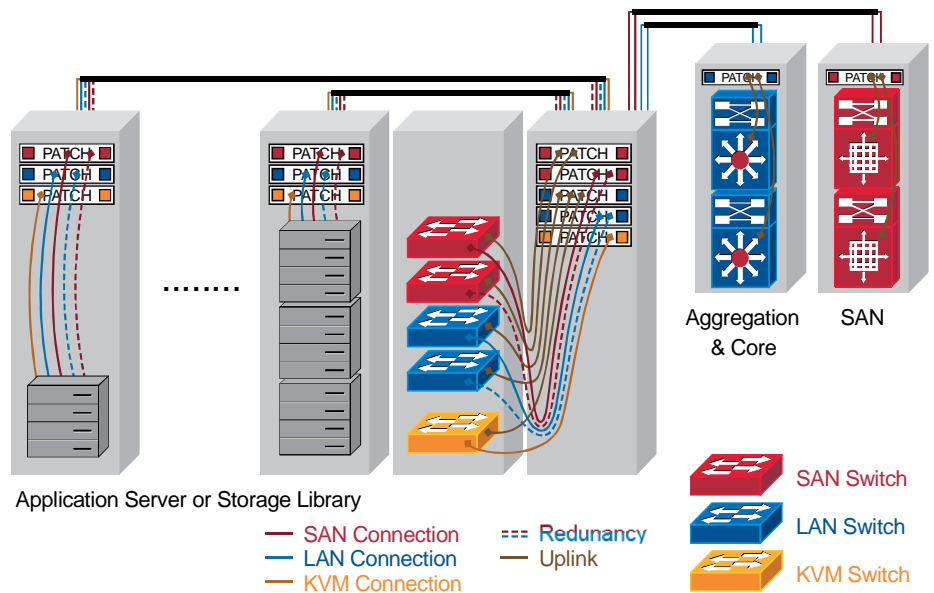


Abb. 3: End of Row

Die Variante **Dual End of Row** bietet sich aus Gründen der Redundanz an, da die Konzentration der Access Switches sich auf beide Enden einer Rackreihe aufteilt (verdoppelt bei Redundanz) und dadurch auch die entsprechenden Kabelwege separiert sind.

Die sternförmige Verkabelung erfolgt also beidseits der Rackreihen, wodurch jedes Rack mit beiden Endpunkten verbunden ist.

Ansonsten entspricht diese Lösung dem EoR-Prinzip.

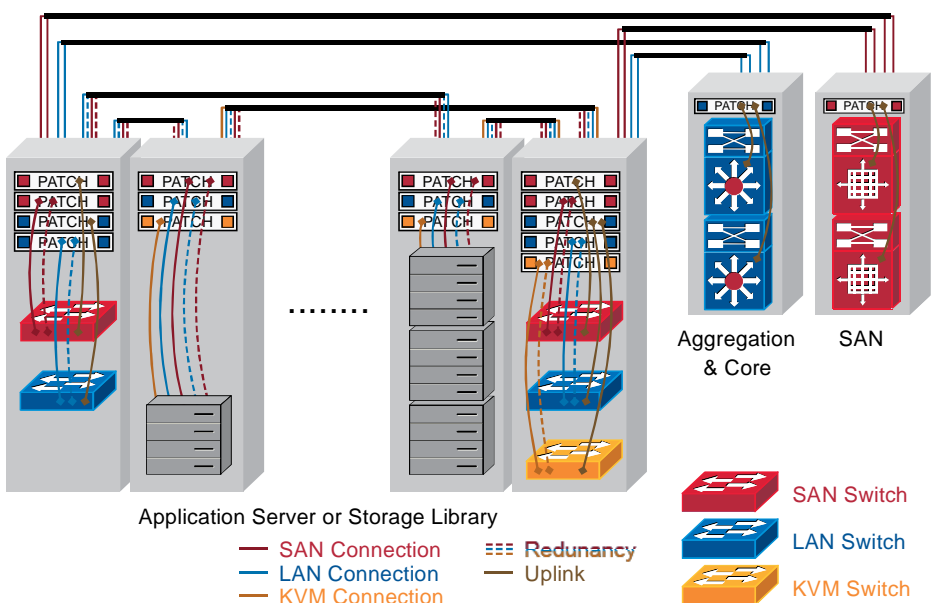


Abb. 4: Dual End of Row

Beim Konzept **Middle of Row** (MoR) liegen die Konzentrationspunkte, die die Access Switches beinhalten, nicht an den Enden sondern zentral innerhalb der Rackreihen.

Damit verkürzen sich die Übertragungsstrecken in der Horizontalverkabelung um die Hälfte.

Dies kann ein entscheidendes Kriterium sein, da sich die zulässigen Distanzen mit steigender Übertragungsrate verkürzen.

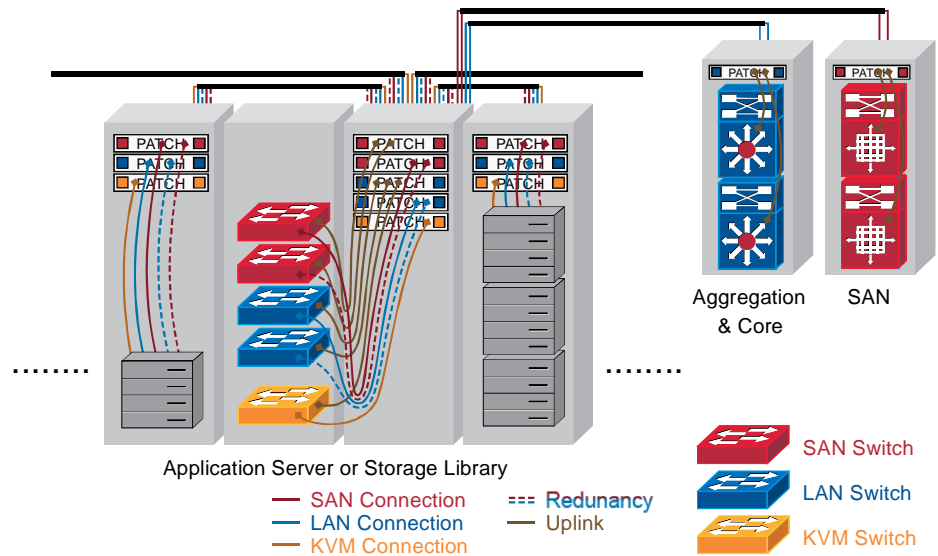


Abb. 5: Middle of Row

EoR/MoR-Verkabelungsarchitekturen weisen sich durch folgende Merkmale aus:

Vorteile	Nachteile
<ul style="list-style-type: none"> • Flexible, skalierbare Lösung (zukunftsicher) • Optimale Belegung der Switch-Ports, wodurch allenfalls Switches eingespart werden können • Konzentration der Access-Switches, was eine Vereinfachung für Moves/Adds/Changes (MAC) bedeutet • Platzoptimierung in den Serverracks für weiteren Ausbau 	<ul style="list-style-type: none"> • Grosses Kabelvolumen in der Horizontalverkabelung • Viele Kabelrangierungen in den Switch-Racks, da Server- und Uplink-Ports rangiert werden • eventuell längere Rackreihen (+1 oder 2 Racks)

3.2. Top of Rack (ToR)

Top of Rack (ToR) ist ein in Data Center eingesetztes Vernetzungskonzept bei welchem, wie aus der Bezeichnung hervorgeht, ein oder mehrere ToR- oder Rack-Switches (in der Funktion von Access Switches) oben in den Serverracks eingebaut sind.

Die Horizontalverkabelung wird vereinfacht, da die Server direkt mit Rangierkabeln an den Switches angeschlossen werden.

Aus den Serverracks führen lediglich die Uplink-Ports der ToR-Switches.

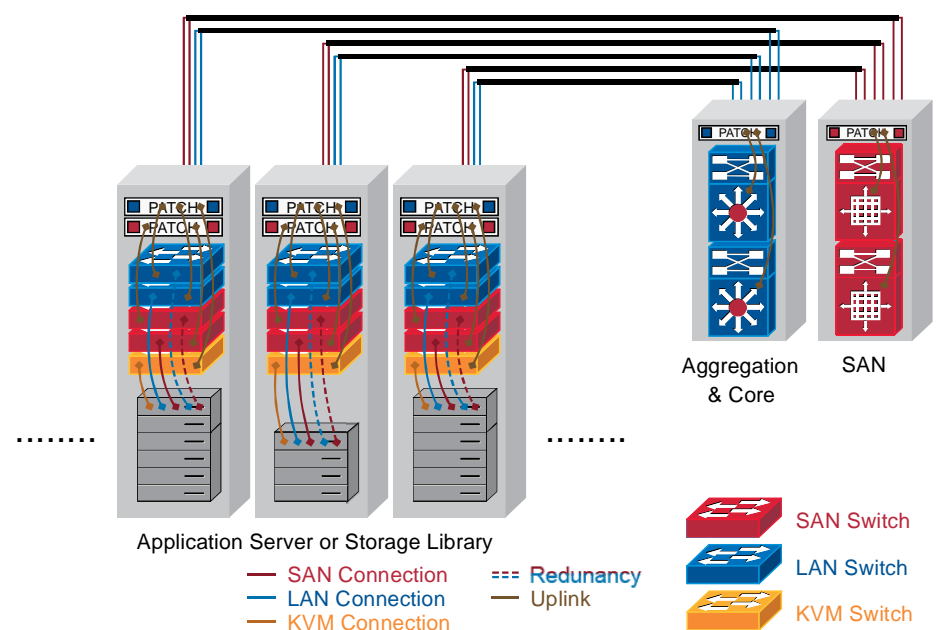


Abb. 6: Top of Rack

ToR-Switches verfügen typischerweise über 48 Ports (inklusive Uplinks) in der Bauform eines 1 HE-Rack-Switches.

ToR-Verkabelungsarchitekturen weisen sich durch folgende Merkmale aus:

Vorteile	Nachteile
<ul style="list-style-type: none">• Kleineres Kabelvolumen in der Horizontalverkabelung, Server-Rangierung repräsentiert Horizontalverkabelung	<ul style="list-style-type: none">• Keine optimale Belegung der Switch-Ports, wodurch allenfalls mehr Switches benötigt werden
<ul style="list-style-type: none">• Einfaches Hinzufügen weiterer Server (Rangierungen)	<ul style="list-style-type: none">• Starres Verhältnis zwischen Access und Aggregation, was die Skalierbarkeit des Netzwerkes einschränkt• Gefahr zur Überbuchung der Uplinks (bei zu vielen leistungsfähigen Servern pro Switch-Einheit/Uplink)

Ein häufiges ToR-Problem ist die Überbuchung der Uplinks. In Ethernet-Umgebungen gibt es traditionell eine hohe Überbuchungsrate (von 5:1 bis 10:1) zwischen Server und Uplink-Ports, insbesondere, wenn preisgünstige Access Switches zum Einsatz kommen. So ist beispielsweise ein ToR-Switch, an welchem 40 Server mit 1 GBit/s angeschlossen sind, mit einem Uplink von 10 GBit/s mit einer Rate von 4:1 überbucht. Sobald ein Switch überbucht wird, kommt es unweigerlich zu unvorsehbaren Vermittlungsverzögerungen.

Die zunehmende Virtualisierung in Verbindung mit der Mobilität virtueller Systeme generiert mehr „horizontalen“ Datenverkehr zwischen Serverracks und zwischen Servern und Storage-Subsystemen. In einer reinen ToR-Architektur (siehe Abb. 6) erfolgt der Kommunikationsweg in diesem Fall immer über die Aggregation- oder sogar Core-Schicht, wodurch der Netzwerk-Backbone unnötigerweise belastet wird, im Gegensatz zu einer EoR/MoR-Architektur.

3.3. Konsequenzen bei der Wahl der Verkabelungsarchitektur

ToR- haben gegenüber EoR/MoR-Konzepten Vor- und Nachteile. Für ToR spricht die effektive Verkabelung mit kurzen Wegen hin zum Server (Rangierung), Nachteil besteht in der frühen Aggregation und der potenziellen Überbuchung der Uplinks. Dies spricht wiederum eher für EoR/MoR-Konzepte, wobei dort dann Chassis mit hohen Portdichten zum Einsatz kommen und als zentrale Fabrics für alle Server dienen. Üblicherweise wird man einen Mix vorfinden, der im ToR-Bereich 1-GBit/s-Verbindungen zu 10 GBit/s aggregiert und diese dann zusammen mit direkten 10-GBit/s-Serververbindungen auf ein EoR/MoR-System geführt werden.

Ein einheitlich strukturiertes Verkabelungskonzept in Anlehnung an EoR/MoR erlaubt, dank der durchgehenden Sternstruktur, die Abbildung verschiedener Architekturen. Damit wird es möglich, eine ToR-Struktur zu betreiben und dennoch auf eine EoR/MoR-Architektur sukzessive wechseln zu können, falls dies erforderlich wird – und dies lediglich durch Umrangierungen. Betreffend Leistungsfähigkeit und Zukunftssicherheit ist eine solche Systemarchitektur kaum zu überbieten, da sie die grösstmögliche Flexibilität zur Verfügung stellt und eine Verflachung der Netzwerkhierarchie gestattet. Werden zudem Querverbindungen zwischen den Rackreihen (von EoR/MoR zu EoR/MoR) vorgesehen, steht auch einer horizontalen Skalierung nichts mehr im Wege. Bei der Auswahl der Verkabelungskomponenten sollten daher Übertragungstechnologien, welche heute eventuell erst im Backbone im Einsatz sind, auch bei der Horizontalverkabelung bereits berücksichtigt werden.

Wird die Horizontalverkabelung unter Berücksichtigung ausreichender Leistungsreserven und der Einhaltung der entsprechenden Längenrestriktionen bereits für die neuen Hochgeschwindigkeits-Ethernet-Technologien 40GBASE-SR4 und 100GBASE-SR10 nach dem EoR/MoR-Prinzip ausgelegt, können zukünftige Migrationschritte problemlos umgesetzt werden – ohne umständliche Nach- oder Umverkabelung in der Horizontalverkabelung.

Eine Server/Server-Kommunikation innerhalb derselben Rackreihe kann damit direkt über den EoR/MoR-Konzentrationspunkt erfolgen (rot markierter Bereich in Abb. 7). Denn im Gegensatz zur reinen ToR-Architektur, wo der Kommunikationsweg in diesem Fall über mehrere Switches und Schichten führt (ToR – Aggregation – Core – Aggregation – ToR), erfolgt diese Datenübertragung über lediglich eine Switch-Einheit (im EoR/MoR), was zu minimalsten Latenzzeiten führt und den Backbone nicht unnötig belastet. Direktverbindungen vom Access- zum Core-Bereich sind mittels Überführungen im EoR/MoR ebenso möglich (Hierarchieverflachung).

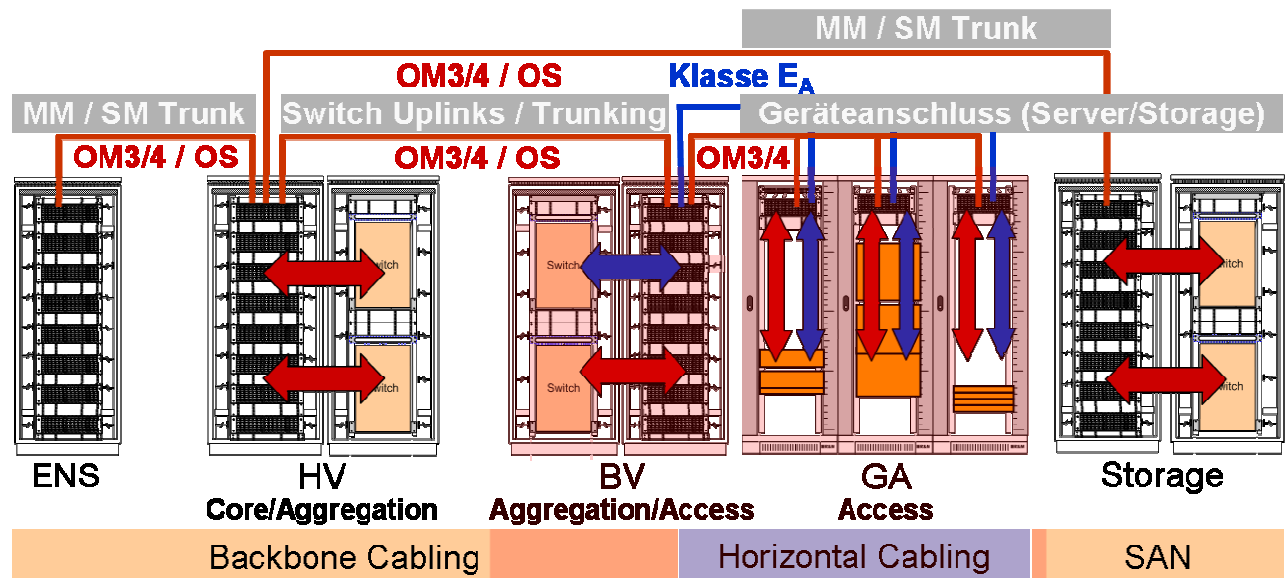


Abb. 7: Empfohlene Übertragungsmedien in den einzelnen Bereichen des Data Center (rot: Glasfaser, blau: Kupfer)

Hierbei gilt es zu beachten, dass die Leistungsreserven der Verkabelungskomponenten eine mehrfache Überführung erst erlauben. So hat sich R&M für die Güteklassen der Glasfaser-Steckverbinder eigene strenge Vorgaben auferlegt, mit welchen sich bei mehreren Stecküberführungen die Mindestdistanzen berechnen lassen. Denn die Länge eines optischen Kanals wird hauptsächlich von der Faser- und Steckerqualität beeinflusst (siehe Whitepaper: Design eines optischen Netzwerks – Planung des Verlust-Budgets mittels Statistik).

Wieviele dieser direkten Serververbindungen möglich sind, wird über die Anzahl der Übertragungsstrecken in der Horizontalverkabelung bestimmt. Wird eine EoR/MoR-Struktur aufgebaut, wo jeder Server direkt im EoR/MoR-Rack angebunden ist, steht für jede Server/Server-Kommunikation eine solche „latenzarme“ Verbindung zur Verfügung. Ein Wechsel vom „ToR-Betrieb“ zum „EoR/MoR-Betrieb“ ist demzufolge jederzeit und sukzessive möglich.

4. Empfehlung

Eine einheitlich durchgehende Sternstruktur vom Access- bis zum Corebereich mit Querverbindungen und ausreichend Leistungs- sowie Leitungskapazitäten in jeder Schicht der Verkabelungsarchitektur, kristallisiert sich nach all diesen Betrachtungen als beste Lösung heraus, um mit den zukünftigen Trends standzuhalten. Deshalb empfiehlt R&M eine Verkabelungsarchitektur in Anlehnung an das EoR/MoR-Konzept. MoR hat den zusätzlichen Vorteil der kürzeren Übertragungsstrecken in der Horizontalverkabelung und wäre daher zu bevorzugen, sofern das Data Center Layout diese Anordnung zulässt.

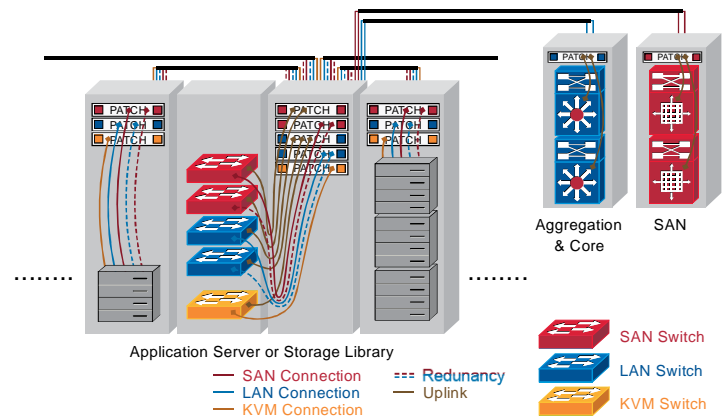


Abb. 9: Middle of Row – die flexibelste Verkabelungsarchitektur

Folgende Aspekte begründen diese Empfehlung:

- Die möglichen I/O-Leistungen eines virtualisierten Systems liegen heute bei etwa 20 – 30 GBit/s, weil der Hypervisor durch SR/IOV (Single-Root I/O Virtualization) und eine geeignete Hardware weitestgehend von den Vermittlungsaufgaben befreit wird. So gibt es Server-Blades, welche als einzelnes Blade mit 2 x 10 Gbit/s angeschlossen werden müssen, um ihr Potenzial zu erschliessen. Bei 8 Blades in einem System, wäre somit ein Systemanschluss mit 100 GBit/s erforderlich.
- Spricht man mit Herstellern von Blade-Systemen, arbeiten sie zurzeit an einer besseren inneren Aggregation, so dass der 100 GbE-Anschluss vermutlich in Kürze angeboten wird. Eine ToR-Struktur herkömmlicher Bauart wäre damit völlig überfordert.
- Heute sind am Markt keine günstigen ToR-Switches erhältlich, welche die notwendige Flexibilität haben und mit neuen hochwertigen Steuerungsfunktionen wie DCB ausgerüstet werden könnten, die End-to-End wirklich sinnvoll funktionieren.
- Dem ToR-Problem der Uplink-Überbuchungen kann mit Direktverbindungen zum EoR/MoR begegnet werden. Verfügbare Switch-Ports können jederzeit für die Aggregation bestehender Verbindungen ($n \times 1/10$ GBit/s) genutzt werden.
- Jede Ende-zu-Ende-Kommunikation kann mittels Hierarchieverflachung (Verkleinerung des Netzwerk-durchmessers) in den Latenzzeiten lediglich durch Umrangierungen optimiert werden.
- Bessere Ausnutzung/Auslastung aller verfügbarer Switchports durch die Hardware-Konzentrationen, d.h. dass die Switchanzahl allenfalls reduziert werden kann. Dies würde Kosteneinsparungen in der Anschaffung und im Betrieb (Maintenance, Energieverbrauch, etc.) bedeuten.
- Das Netzwerkmanagement und Trouble Shooting wird vereinfacht (Switch-Upgrades, Konfigurationen), da sich die Netzwerkhardware in den Rackreihen konzentriert (kein unnötiger Zugriff in Server racks).

Das Sparpotenzial bei den aktiven Netzwerkkomponenten ist im Gegensatz zu allfälligen Einsparungen bei den Verkabelungskomponenten nachhaltig, da der erforderliche Energiebedarf dauerhaft gesenkt werden kann (1 Watt Einsparung bei der IT = 3 Watt Einsparung des Gesamtenergiebedarfs!).

5. Zusammenfassung

Hierarchisch aufgebaute Netzwerke in Data Center bestehen typischerweise aus drei Schichten (Three Tier Network): der Access-, der Aggregation- und der Core-Schicht. Das Datenaufkommen der unteren Schicht muss möglichst verzögerungsfrei der nächsthöheren weitervermittelt werden. Somit muss jede übergeordnete Schicht deutlich leistungsfähiger sein, um eine Überbuchung der Uplinks und damit Verzögerungen zu vermeiden. Ein höherer Datendurchsatz wird mit einer schnelleren Übertragungstechnologie oder durch die Bündelung mehrerer Verbindungen in der Aggregation-Schicht erreicht.

Eine „End of Row“-Verkabelungsarchitektur bietet durch ihre Kabel- und Switch-Konzentration an den Enden oder zentral innerhalb der Rackreihen (Middle of Row) die leistungsfähigste Basis für eine effiziente Kommunikation zwischen den Schichten – heute wie auch morgen. Bei „Top of Rack“-Architekturen können zwar die Server direkt mit den ToR-Switches (Edge/Access Switches) in den Serverracks verbunden und somit das installierte Kabelvolumen reduziert werden. Dies geht jedoch auf Kosten der Skalierbarkeit und demzufolge auch auf Kosten der Zukunftssicherheit der Investitionen.

Steigt die Serveranzahl oder werden die Serverleistungen auf 10 GBit/s erhöht und in Blade-Systemen aggregiert, müssen die Daten mit 40 resp. 100 GBit/s in Richtung Core übertragen werden können. Spätestens dann führt dies bei „ToR“-Architekturen zu massiven Überbuchungen der Uplinks und unakzeptablen Latenzzeiten im Netzwerk.

Aus diesen Gründen empfiehlt R&M eine vermaschte „End of Row“- oder „Middle of Row“-Topologie: eine einheitliche, vertikal sowie horizontal skalierbare und demzufolge zukunftsorientierte Verkabelungsarchitektur.

6. Quellen und weitere Informationen

Weitere Informationen zu Produkten und Lösungen von R&M finden Sie auch im Internet unter:

www.rdm.com

www.datacenter.rdm.com