

RECHENZENTREN UND INFRASTRUKTUR

KOMPONENTEN, KABEL, NETZWERKE

Netzwerkprotokolle
im Vergleich

**Konvergenz:
40-/100-Gigabit-
Ethernet bilden das
Fundament**

Seite 4

**RZ-Planung:
Überlieferte Werte
gehören auf den
Prüfstand**

Seite 18

**Kalt- oder Warmgang:
Einhausung spart
im Betrieb massiv
Kosten**

Seite 22

**Anwendungsbeispiel:
Freiluftkühlung im
Rechenzentrum
garantiert Effizienz**

Seite 24

**High-Speed-Netze:
Ethernet, Infiniband
und Fibre Channel**

Seite 12

**Arbeitsspeicher:
DRAM im Server
frisst Energie**

Seite 20

**Brandvermeidung:
Dem Feuer die Luft
abdrehen**

Seite 23



Less energy.
More speed.



The new 30 nano class Green DDR3

Samsung's 30 nano class 4G bit DDR3 server memory chip is the most advanced, best-performing chip we've ever created. It saves 86% more energy, processes two times faster and is far more reliable than its predecessor.* In fact, its energy usage is so small, operating and maintenance costs of your server farm are significantly reduced. Welcome the eco-innovation that doesn't compromise performance – just one more reason the leader in green memory technology is Samsung.

www.samsung.com/greenmemory



Visit us at CeBIT
Hall 17, booth D38



Wie viele Rechenzentren frisst das Cloud Computing?



IT-Ressourcen auf Abruf zukaufen, Bereitstellung selbst von komplexen Software-Architekturen binnen Stunden, Entfernen von nicht mehr benötigten Einheiten und die Abrechnung nach tatsächlichem Verbrauch: Ein derartiges Angebot für die Bereitstellung von IT erscheint unschlagbar. Wer diese Vorteile des Cloud Computing in die Waagschale wirft, wenn es darum geht, das unternehmenseigene Rechenzentrum selbst zu erweitern oder gar ein neues zu konzipieren, der hat schlechte Karten.

Sicher – heute gilt noch eine wichtige Trumpfkarte: Der Aspekt Sicherheit schlägt viele andere Argumente aus dem Feld. Der vielfach zitierte mittelständische Maschinenbauer ist nur schwer zu überzeugen, dass er seine Kronjuwelen wie die Konstruktionszeichnungen „irgendwo im Internet“ abspeichert. Was für private Urlaubsfotos funktioniert, das hat im Profibereich, etwa bei Finanzdienstleistern wie Versicherungen und Banken, nichts zu suchen – diese Argumentation hat allerdings bald ausgedient!

Denn die Anbieter von Cloud-Diensten erkennen recht schnell, dass sie hier nachbessern müssen. Mittlerweile sind mit den „Hybrid Clouds“ Architekturen im Gespräch, bei denen einige wichtige Informationen noch im eigenen Rechenzentrum residieren und das Gros der Daten – die weniger schützenswert sind – bereits bei einem Cloud-Anbieter liegen. Das ist nichts anderes als eine Einstiegsdroge.

Wer mittel- und langfristig mit seinem IT-Konzept gegen die „Cloud“ bestehen will, darf sich nicht darauf verlassen, dass es mit dem Zukauf von einigen energieeffizienteren Geräten getan ist. Die große Kunst liegt darin, eigene IT-Dienste zu konzipieren und die Bereitstellung dieser Services permanent zu verbessern. Dazu gehört dann auch das Beachten der einschlägigen Compliance-Vorgaben und das Qualitätsmanagement. Hier

müssen die Mitarbeiter im IT-Umfeld über die nötigen Zertifizierungen verfügen.

Wer hier aber sein Heil nur im eigenen Unternehmen sucht und die Aktivitäten von anderen, vergleichbaren Unternehmen außen vor lässt, der muss viel – vielleicht sogar zu viel – Engagement in diese Aufgabe stecken. Deswegen bietet sich die Mitarbeit in Arbeitskreisen an, wie sie zum Beispiel der ECO-Verband im Rahmen der ECO Datacenter Expert Group am 8. Februar 2011 startet (www.dcaudit.de/lang/de/2011/01/03/08-02-kick-off-meeting-datacenter-expert-group/).

Teil der Initiative sind zwei neue Arbeitskreise, die an diesem Tag zur konstituierenden Sitzung laden: Der Arbeitskreis „Nachhaltiges IT-Management“ wird sich um Themen wie Energieeffizienz und Nachhaltigkeit in der IT kümmern. Der Arbeitskreis „IT-Prozesse“ setzt seinen Schwerpunkt auf Prozess- und Standardisierungsthemen. Ebenfalls an diesem Tag einberufen wird der dritte Arbeitskreis in diesem Umfeld: Die vormals als „Datacenter“ bezeichnete Initiative bekommt nun den Titel „Arbeitskreis Infrastruktur“.

Wer die Ergebnisse dieser Arbeitskreise verwendet und sie mit den Anforderungen aus den einzelnen Fachabteilungen im Unternehmen elegant verknüpft, der hat gute Chancen. Denn letztendlich muss die IT das Geschäft unterstützen und den Weg für Innovationen frei machen. Und das kann ein eigenes IT-Team mit einem guten Verständnis für die Anforderungen sicher besser als ein Cloud-Anbieter, der auf möglichst standardisierte Dienste setzen muss. Damit kann das eigene Rechenzentrum gegenüber „Infrastructure/Platform/Storage/Software as a Service“ erfolgreich bestehen.

Rainer Huttenloher

Konvergenz im Netzwerk treibt das schnelle Ethernet

40- und 100-Gigabit-Ethernet legen das Fundament für die künftige Entwicklung

Mit der Spezifikation von 40- und 100-Gigabit-Ethernet (GbE) nach der Norm IEEE 802.3ba-2010 steht die Basis für die nächste Generation der Hochgeschwindigkeitsübertragung zur Verfügung. Die Konvergenz im Netzwerkbereich legt den Schluss nahe, dass sich Ethernet als Übertragungstechnologie für nahezu alle Bereiche durchsetzt. Dabei sind mit IEEE 802.3ba die Grundpfeiler gesetzt, auf die auch künftige Generationen aufbauen werden.

In bestimmten Bereichen steigt der Bandbreitenbedarf für die IT-Infrastruktur massiv an. Daher reicht oftmals der Durchsatz nicht mehr aus, wie ihn 10-Gigabit-Ethernet (10 GbE) bereits zur Verfügung stellt. Der primäre Grund liegt in den Datenmengen, die mittlerweile zu transportieren sind. Doch es sind nicht alle Unternehmen gleichermaßen von dieser Herausforderung betroffen. Typische Vertreter mit einem hohen Bandbreitenbedarf sind Unternehmen wie Google, aber auch die Betreiber von Internet-Austauschknoten gehören zu dieser Kategorie.

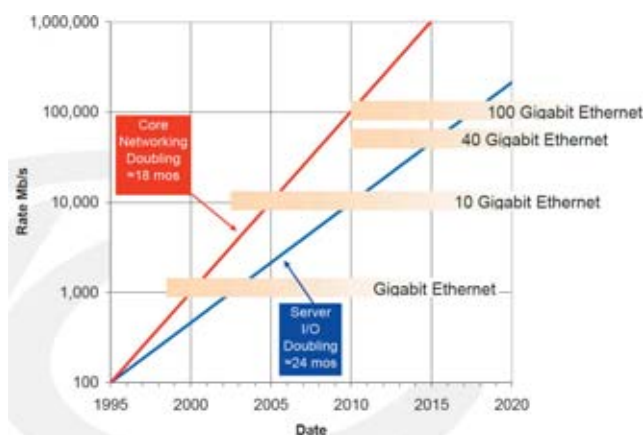
Doch die Forderung nach hoher Bandbreite muss man gewissermaßen genauer spezifizieren: Alle Vorreiter in Sachen Bandbreite benötigen sehr leistungsfähige Core-Netzwerke – so liegt etwa beim Amsterdamer Internet Exchange der Bandbreitenbedarf im Core-Netzwerk schon bei einem Terabit pro Sekunde (Tbit/s). Der weitere Anstieg in Bezug auf das Transfervolumen lässt sich allerdings nur sehr schwer abschätzen, denn die Zeiträume für die Nutzung der Infrastruktur sind recht lang. Dabei liegt die Phase für die Standardisierung der nächsten Generation im Ethernet-Bereich bei über fünf

Jahren – so die Erfahrungen aus dem Arbeiten der Taskforce zu IEEE 802.3ba.

Für die Arbeit in den IEEE-Standardisierungsgremien waren im Dezember 2005 erste Aussagen zur Nachfolge-Generation von 10 GbE relevant. In der Diskussion zeigte sich schnell, dass mehr als 10 GbE nötig werden und so kam es dann im Juli 2006 zu einem „Call for Interest“ (CFI) durch das IEEE 802.3. In dem Normungsgremium sitzen Vertreter von Hardware-Chip- und Kabelherstellern. Diese Unternehmen rechnen alle in erster Linie in der Größenordnung „verkaufte Ports“. Anwender sehen sich dagegen mit der Aufgabenstellung konfrontiert, wie sie mehr „Nutzen in die Netzwerke“ bringen können. Diese Zielsetzungen sind allerdings nicht immer deckungsgleich. Generell besteht die Erwartung, dass der Umstieg auf die Generation mit 40 und 100 GbE in den Mittelpunkt rückt, selbst wenn derzeit die Anzahl der verkauften 10-GbE-Ports noch nicht bei einer Million liegt. Erst wenn diese Größenordnung übertroffen wird, spricht man vom Durchbruch der Technologie.

Doch mittlerweile hat sich die Situation geändert. Das komplette Umfeld für Hochgeschwindigkeitsübertragung sieht anders aus, viele Unternehmen behelfen sich mit der „Netzwerk-Aggregation“, sprich sie kombinieren mehrere 10-GbE-Verbindungen zu einem Link. Denn das Anwenderverhalten hat sich massiv geändert. Aufgrund des Erfolgs von „Mitmach-Plattformen“, konsumieren die Benutzer nicht mehr nur die Inhalte, sondern sie tragen auch dazu bei. Ein gutes Beispiel dafür ist YouTube: Extrem viele Videos werden eingestellt, die dann auch angeschaut werden. Daher betrifft der Bedarf von schnelleren Netzwerken in erster Linie die Anbieter von Inhalten und die Backbone-Betreiber.

Doch auch die Betreiber von Rechenzentren schwenken auf die schnelleren Netzwerke um. Viele haben zwar noch Industrie-Standardserver mit Netzwerkschnittstellen mit 1 GbE im Einsatz, doch es werden sehr viele dieser Systeme „parallel“ eingesetzt. Derzeit halten die Schnittstellen mit 10 GbE Einzug in die Standardserver, denn diese Technologie wird zum einen deutlich günstiger. Zum anderen zieht der massive Einsatz der Server-Virtualisierung mit Lösungen von VMware, Microsoft oder Citrix einen steigenden Bedarf am Ein-/Ausgabedurchsatz für die physischen Server mit sich. Wenn die Industrie-Standardserver weitgehend auf 10 GbE-Interfaces umgestellt sind, lässt sich nur mit einem Einsatz von 100 GbE das Übertragungspotenzial der Systeme vernünftig ausnutzen.



Die zwei Linien im Diagramm zeigen zum einen die Entwicklung im Bereich der Übertragungsraten des Core Network (dabei verdoppelt sich die Übertragungsrates alle 18 Monate) und zum anderen die Entwicklung der Ein-/Ausgabeleistung von Servern – sie verdoppelt sich laut Moores Law alle 24 Monate (Abb. 1).

Quelle: IEEE802.org/Force10 Networks



datacenter.de



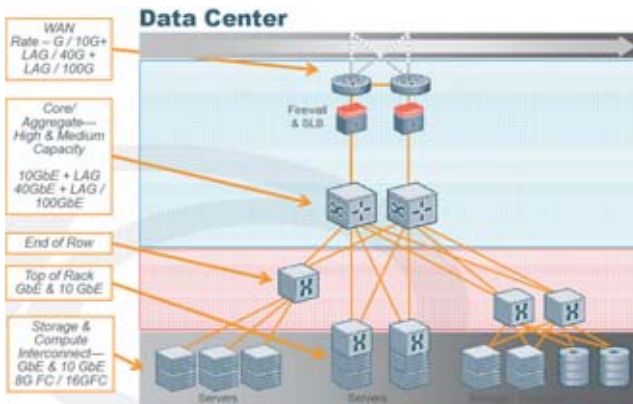
5-Sterne Premium-Rechenzentrums-Flächen

für Ihre Sicherheit und hohen Ansprüche.
Fordern Sie uns!

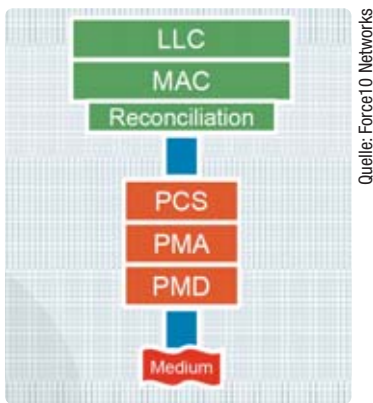


www.datacenter.de

datacenter.de ist eine Marke der noris network AG.



Verkabelungskonzept im Rechenzentrum (Abb. 2)



Übersicht über die Architektur von 40 und 100 GbE (Abb. 3)

In der Diskussion zur nächsten Ethernet-Generation nach dem ersten CFI stellte sich heraus, dass eine Normierung von zwei Stufen sinnvoll ist: 40 und 100 Gbit/s. Abbildung 1 zeigt ein Diagramm, das zwei Linien enthält: Einmal die Entwicklung im Bereich der Übertragungsraten des Core Network (dabei verdoppelt sich die Übertragungsrate alle 18 Monate) und zum anderen die Entwicklung der Ein-/Ausgabeleistung von Servern (sie verdoppelt sich laut Moores Law alle 24 Monate).

Diese Grafik zeigt auch, dass 1 GbE für viele Anwendungen angepasst hat, als es auf den Markt gekommen ist. Es hat genau die Erwartungen und Bedürfnisse getroffen, die zu der Zeit bestanden – die Zeitspanne zwischen den beiden Linien ist nicht allzu lang. Dagegen teilt 10 GbE immer noch die Einschätzung, dass sich der Durchbruch in Bezug auf die verkauften Ports auf dem Markt nicht ergeben hat: Denn die Volumenmärkte – sprich die Industrie-Standardserver sind noch nicht so weit. Das hat erst begonnen und setzt sich im Verlauf von 2011 entsprechend fort.

Core-Netzwerke brauchen Speed

Im Bereich der Netzwerke ist 10 GbE dagegen schon weit verbreitet. Die Frage zu den nächsten Geschwindigkeitsstufen lässt sich aufgrund der beiden Linien in der Grafik belegen: Eine neue Geschwindigkeitsstufe allein reicht nicht, denn der Abstand zwischen den beiden Linien – Netzwerkbandbreite und Server-I/O – wächst an. Daher passt die Variante mit 40 und 100 GbE sehr gut in das Bild. Denn die Server brauchen 100 GbE erst 2017, daher die Idee mit dem Zwischenschritt auf 40 GbE. Das wird schon 2015 benötigt.

Die Vorreiter unter den Anwendern, die höhere Bandbreiten in ihren Netzwerken brauchen, werden dabei wohl gleich auf 100 GbE umsteigen. Dazu gehören Unternehmen wie Google oder auch der neue Börsenstar Facebook. Im Jahr 2007 – nachdem die Standardisierungsgruppe zur nächsten Ethernet-Generation ihre Arbeit gestartet hatte – gab es weltweit 20 Millionen Facebook-Anwender. Im Juli 2010 stieg diese Zahl auf 500 Millionen. Zusammen mit dem datenintensiveren Nutzerverhalten – der Schwenk zu mehr Videoinhalten – verschärft für diese Unternehmen die Situation.

Anders sieht die Herausforderung dagegen in einem typischen Rechenzentrum aus (Abbildung 2). Das Gros der Server arbeitet heute noch mit Schnittstellen von 1 GbE, die Systeme mit 10 GbE kommen erst allmählich zum Einsatz. Die Switches – sei es in der Konfiguration „Top of Row“ (ToR) oder „End of Rack“ (EoR) arbeiten mit 1 GbE oder 10 GbE. Im Core des Netzwerks sind dann 10 GbE vorhanden – meist mit Link-Aggregation im Einsatz. Viele Anwendungen brauchen derzeit nicht mehr als 10 GbE.

Doch das Problem ist größer als die reine Nachfrage in Sachen Bandbreite. Denn es geht beim Interconnect auch um Themen wie Port-Typ, Port-Speed und „Media Reach“ – also wie weit man über die verschiedenen Übertragungsstrecken mit der spezifizierten Transfer-rate die Signale treiben kann. Mit der Konvergenz hin zu Ethernet, wie sie derzeit industrieweit zu sehen ist, kommen noch weitere Anforderungen ins Spiel:

- OTN (Optical Transport Network),
- FCoE (Fibre Channel over Ethernet),
- MPLS (Multiprotocol Label Switching),
- TRILL (Transparent Interconnect of Lots of Links),
- Virtualisierung und Automation

lauten dabei die Stichworte.

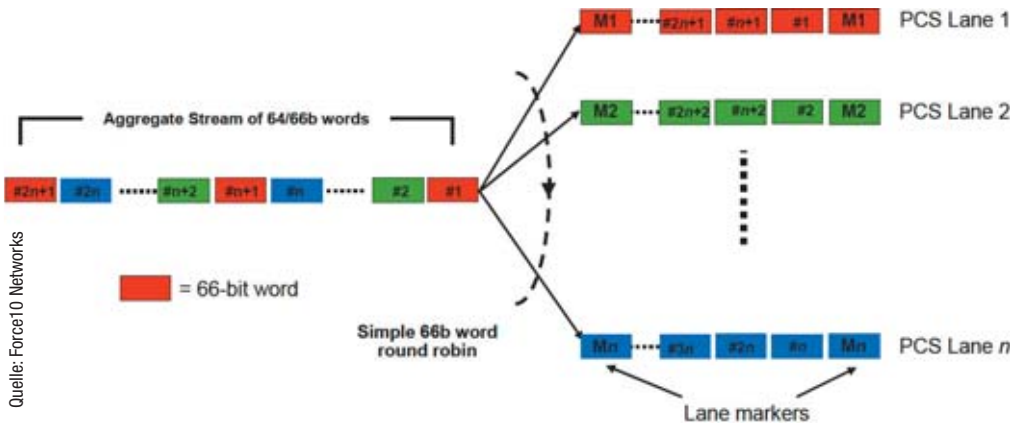
Beim Aufbau eines Netzwerks in einem Unternehmen ist es wichtig, eine zukunftsfähige Konfiguration aufzubauen. Dabei sollte die Verkabelung möglichst für die aktuelle und zudem noch für die kommende Transfertechnologie passen. Hier gilt es mit wenig Änderungen und Investitionen auch kommende Transfertechniken abzudecken. Speziell bei 40 und 100 GbE einigte sich die Standardisierungsgruppe, dass die Ebenen LLC (Logical Link Control) und MAC (Media Access Control) gleichbleiben.

Den Überblick über die Architektur von 40 und 100 GbE gibt die Abbildung 3. Für die MAC- und die übergeordnete LLC-Ebene ergeben sich dabei keine Änderungen. Die Modifikationen betreffen die drei Sublayer:

- PCS (Physical Coding Sublayer), damit wird die physikalische Codierung abgehandelt,
- PMA (Physical Media Attachment), mit diesem Sublayer wird die Information auf das Übertragungsmedium transferiert und
- PMD (Physical Media Dependency), dieser Sublayer ist dafür verantwortlich, wie die Signale auf den verschiedenen Übertragungsmedien getrieben werden.

Vom Prinzip gibt es bei 40 GbE mehrere Möglichkeiten, die 40 Gbit/s zu senden. Man könnte viermal 10 Gbit/s pro Lane übermitteln. Dazu haben viele Komponentenhersteller bereits Chips entworfen, die vier Kanäle zu je 10 Gbit/s abdecken. Will man dagegen auch noch 40 GbE auf einer „Lane“ transferieren, dann sind bereits zwei Implementierungen für den 40-GbE-Standard zu berücksichtigen, die alle spezifiziert werden müssen.

Komplexer erweist sich dann die Angelegenheit bei 100 GbE: Beim Start des Projekts war es noch nicht abzusehen, wie die Verteilung auf verschiedene Lanes aussehen wird, sprich wie die Signale auf dem Medium zu senden sind. Es gibt theoretisch die Option, einmal 100 Gbit/s



Die Verteilung auf verschiedene „Lanes“ findet im PCS-Sublayer statt (Abb. 4).

zu treiben, oder zweimal 50 oder viermal 25 GBit/s oder fünfmal 20 GBit/s, aber auch die Variante mit zehnmal 10 GBit/s erscheint vernünftig. All diese Aufteilungen waren für die Standardisierung sinnvolle Optionen, daher musste die Architektur all diese Optionen auch unterstützen. Das bedeutete viel Aufwand für den gesamten Prozess.

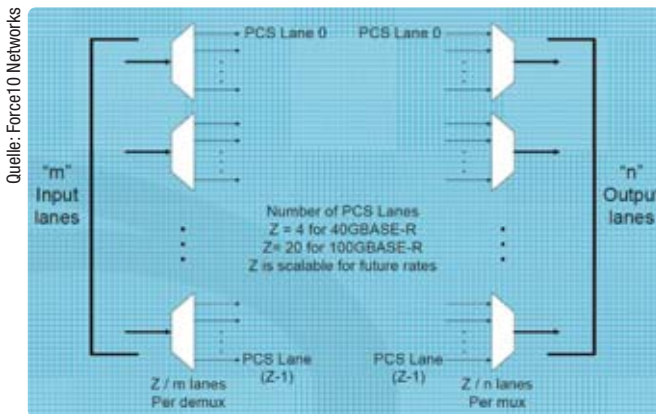
Im PCS ist der erste Teil für die Unterstützung der verschiedenen Skalierungsoptionen zu finden (siehe Abbildung 4). Hier erfolgt zum einen das Codieren der Datenbits und zum anderen das Zusammenfassen in Code-Gruppen, die letztendlich den seriellen Datenstrom bilden. Als Codierungs-Verfahren findet dabei das 64B/66B-Verfahren Verwendung, das bereits bei 10 GbE zum Einsatz kam. Damit die Informationen, also die Code-Blöcke, über mehrere Lanes verteilt, weiter geleitet werden können, wird in vordefinierten Lane-Anzahlen die das Senden nach einem Round-Robin-Verfahren abgewickelt. Dabei sind für 40 GbE bis zu vier Lanes vorgegeben. Dasselbe Prinzip gilt auch für 100 GbE, allerdings sind dabei bis zu 20 PCS-Lanes vorgesehen. Im Empfangskanal des PCS-Sublayer wird zudem noch sichergestellt, dass die Laufzeitunterschiede der einzelnen Code-Blöcke – auch als Skewing bezeichnet – kompensiert werden. Dazu werden bei der Übertragung in regelmäßigen Abständen noch Markierungsinfos pro Lane eingebettet.

Im PMA-Sublayer wird das Umstellen von der Anzahl der verwendeten Lanes und der Übertragungsrate pro Lane abgewickelt (Abbildung 5). Das ist notwendig, um mehrere Spezifikationen für die physikalischen Ebenen zu bieten. Damit wird eine flexible Abbildung von m Eingangs-Lanes zu n Ausgangs-Lanes erreicht, je nachdem wie viele

man für seine Übertragungsstrecke braucht. Für 100 GbE ist zehnmal 10 GBit/s heute Stand der Technik (etwa bei Multimode-Glasfasern, bei denen zehn parallele Pfade jeweils 10 GBit/s übertragen: 100GBASE-SR10). Die Variante mit viermal 25 GBit/s ist die zweite gängige Verteilung, etwa bei einer Singlemode-Glasfaser, bei der vier Wellenlängen mit jeweils 25 GBit/s zum Einsatz kommen (als 100GBASE-LR4 bezeichnet).

Für die Spezifikationen des physikalischen Layer gibt es bei 40 und 100 GbE verschiedene Port-Typen. Für den Einsatz im Bereich von Backplanes bei 40 GbE ist 40GBASE-KR4 vorgesehen. Es bewältigt eine Entfernung (auf Kupferleiterbahnen) von mindestens einem Meter. Mit 40GBASE-CR4 und 100GBASE-CR10 steht eine Spezifikation für den Transfer über Kupfer-Twinax-Kabel für beide Geschwindigkeiten zur Verfügung. Die Übertragungsstrecke darf dabei bis zu sieben Meter für 40 und 100 GbE betragen.

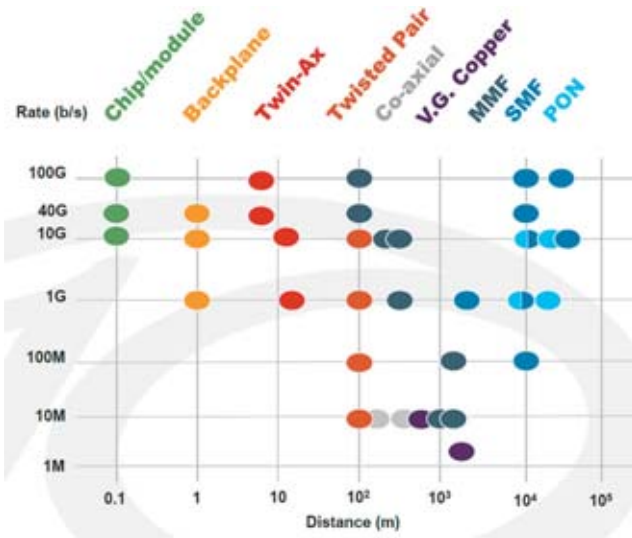
Im Bereich der Glasfaser-Übertragungsmedien sind größere Entfernungen mit 40 und 100 GbE machbar. Für Multimode-Fasern (MMF) gibt es die 40GBASE-SR4 und 100GBASE-SR10. SR steht hier für „Short Reach“, das sind für eine MMF der Güte OM-3 mindestens 100 Meter und für die hochwertigeren MMF, also OM-4, sogar 150 Meter. Dabei werden vier oder zehn Lanes mit jeweils 10 GBit/s getrieben. Sprich, es handelt sich um einen „Parallel Fibre Ansatz“. Als Lösungen für 10 Kilometer sind 40GBASE-LR4 und 100GBASE-LR4 (jeweils viermal 25 GBit/s) vorgesehen. Die Bezeichnung LR steht für „Long Reach“ – dabei sind für SMF (Singlemode-Fasern) Längen von mindestens



Die Abbildung von m Eingangs-Lanes zu n Ausgangs-Lanes erfolgt mit einem entsprechenden Multiplexen und Demultiplexen im PMA-Sublayer (Abb. 5).

		Planned for 1 st Generation		Not Planned for 1 st Generation	
Media	Reach	Speed	CFP	QSFP	CXP
Single-mode	10Km	100G	Green	Red	Red
		40G	Green	Future?	Red
Multimode (OM3)	100m	100G	Green	Red	Red
		40G	Green	Future?	Red
Multimode (OM4)	150m	100G	Green	Red	Red
		40G	Green	Future?	Red
Copper	3-7m	100G	Green	Red	Red
		40G	Green	Future?	Red

Definierte Formfaktoren für die Transceiver bei 40 und 100 GbE (Abb. 6)



Die Ethernet-Übertragungsgeschwindigkeiten in Abhängigkeit von der Übertragungsentfernung, viele Übertragungsmedien sind dazu spezifiziert (Abb. 7).

Quelle: David Law

10 Kilometern vorgesehen. Mit 100GBASE-ER4 sind über SMF sogar noch 40 Kilometer spezifiziert – mit der Aufteilung von viermal 25 GBit/s.

Für die neuen Standards sind sozusagen in der ersten Generation einige Formfaktoren für die Transceiver (sie werden auch als MSA, Multi Source Agreements, bezeichnet) vorgesehen (siehe Abbildung 6). Es gibt hier deutliche Größenunterschiede – das ist nötig, denn es sind zum Beispiel bei einer Konfiguration von zehnmal 10 GBit/s die elektrischen Interface-Bausteine anzusteuern. Das hat auch eine hohe Leistungsaufnahme zur Folge. Zudem ist die Signalintegrität ein bestimmender Faktor. Hier sind Parameter wie das Übersprechen zu berücksichtigen. Dabei gilt: je enger die Leitungsabstände, umso größer das Problem. Ein zweiter Grund ist der Aufwand für das Multiplexing.

Bei der LR-Optiken basiert 100 GbE auf einer Konfiguration von viermal 25 GBit/s und damit wird ein Multiplexer nötig, der die ursprünglichen zehn Lanes auf vier Lanes (in der optischen Übertragungsstrecke) zusammenführt. Dieser Vorgang benötigt erneut Power, die dann als Wärme abzuführen ist. Somit erweist es sich erneut als ein wichtiges Kriterium, wie die Wärme abgeführt werden kann. Als künftige Formfaktoren für die nächsten „Implementierungs-Generationen“ von 100 GbE sind bereits ein QSFP2 sowie ein CFP2 in der Diskussion.

Generell wäre es wünschenswert, wenn man bei diesen Formfaktoren eine Vereinheitlichung hinbekommen könnte. Schon bei 10 GbE gibt es sieben verschiedene Varianten – das vereinfacht die Handhabung auf keinen Fall. Eine gemeinsame Lösung wäre ideal, doch allerdings nur dann, wenn die Wärmeabfuhr bei den aufwendigeren Varianten funktioniert. Derzeit sind folgende Formfaktoren für die Transceiver von der Ethernet Alliance geplant:

- CFP (C Formfactor Pluggable, wobei das C für die römische Zahl 100 steht): Er ist für 40 oder 100 GbE geeignet. Er bietet 12 Send- und Empfangsleitungen mit jeweils 10 GBit/s. Damit lassen sich entweder ein 100-GbE-Port oder bis zu drei 40-GbE-Ports abdecken. Aufgrund seiner mechanischen Dimensionierung passt er auch für die Anforderungen von Singlemode- und Multimode-Glasfasern und außerdem für die von Kupferleitungen.
- CXP: Dieser Transceiver-Formfaktor bietet zwölf Lanes in jede Richtung (Senden und Empfangen), doch er ist deutlich kleiner als der CFP. Der CXP eignet sich für Multimode-Glasfasern und Kupferleitungen. Dabei passt dieser Typus auch für die Version SR10 bei 100 GbE.
- QSFP (Quad Small Formfactor Pluggable): Er ähnelt von der Größe her dem CXP und bietet je vier Lanes für das Senden und das Empfangen. Damit lassen sich 40 GbE über Multimode-Glasfasern und über Kupferleitungen übertragen. Es ist abzusehen, dass dieser Formfaktor auch für Singlemode-Glasfasern eingesetzt wird. Zudem kann die Unterstützung für 100 GbE kommen, wenn man pro Lanes 25 GBit/s übertragen kann.

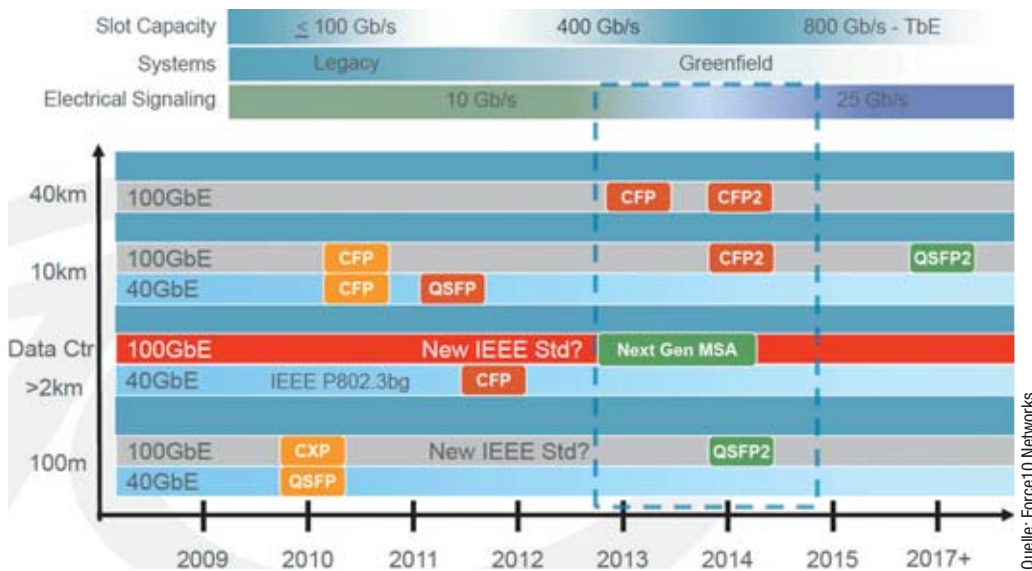
Herausforderungen für die Switching Engine

Aus dem Blickwinkel der Anschluss technik rücken weitere Anforderungen für die Hochgeschwindigkeitsübertragung in den Vordergrund. Dazu sind rechnerisch die Durchsatzraten bei den Switches zu bestimmen. Geht man bei den Line Cards von Netzwerkgeräten mit den Port-Anschlüssen auf der Frontseite aus, so lassen sich bereits 44 Ports (mit QSFP) realisieren. Wenn pro Port 40 GbE übertragen wird, addiert sich das zu 1,76 Tbit/s auf. Im Bereich von 100 GbE sind mit CXP-Anschlüssen bis zu 32 Ports auf einer Line Card machbar. Das ergibt dann eine Bandbreite von bis zu 3,2 Tbit/s pro Line Card. Übliche Größenordnungen

Quelle: Ethernet Alliance

GLOSSAR ZU 40- UND 100-GIGABIT-ETHERNET

AN	Auto Negotiation
CAUI	100 GBit/s Attachment Unit Interface („C“ steht für römisch 100)
CGMII	100 GBit/s Media Independent Interface („C“ steht für römisch 100)
CWDM	Coarse Wave Division Multiplexing
HSSG	Higher Speed Study Group
IEEE 802.3 Standard	Bezeichnung für den generellen Ethernet-Standard
IEEE P802.3ba	das Projekt, das den Ethernet-Standard für 40 und 100 GbE definiert hat
Ethernet IEEE Std 802.3ba-2010	der offiziell verabschiedete Zusatz zu Ethernet für die Standards 40 und 100 GbE
MAC	Media Access Control Layer
MDI	Medium Dependent Interface
MII	Media Independent Interface
MLD	Multilane Distribution (Verteilung der Signale über mehrere Lanes)
OTN	Optical Transport Network
PCS	Physical Coding Sublayer
PHY	Physical Layer Devices Sublayer
PMA	Physical Medium Attachment Sublayer
PMD	Physical Medium Dependent Sublayer
PPI	Parallel Physical Interface
RS	Reconciliation Sublayer
SoC	System on a Chip
WDM	Wave Division Multiplexing
XLAUI	40 GBit/s Attachment Unit Interface („XL“ steht für römisch 40)
XLGMII	40 GBit/s Media Independent Interface („XL“ steht für römisch 40)



Die Technologie-Roadmap für 40 und 100 GbE; es kommen mit QSFP2 und CFP2 weitere Transceiver-Typen ins Spiel (Abb. 8).

gen für Line Cards, die über eine Backplane verbunden sind, liegen bei acht oder 14 Line Cards pro System. Das hat zur Folge, dass in der Switching Fabric bis zu 25,6 Tbit/s (bei acht Line Cards) beziehungsweise 44,8 Tbit/s (bei 14 Line Cards) zu beherrschen sind.

Die Abdeckung durch die Ethernet-Technologie führt zu der Einschätzung, dass sich praktisch alle Anforderungen damit abdecken lassen. Die Abbildung 7 zeigt ein Diagramm, das die Übertragungsgeschwindigkeiten in Abhängigkeit von der Übertragungsentfernung – gekennzeichnet für die verschiedenen Übertragungsmedien – darstellt. Ganz links im Bild sind die kurzen Entfernungen zu sehen, wie sie etwa bei der Kommunikation zwischen Chips und einzelnen Modulen anfallen. Der Bereich um einen Meter bezieht sich auf die Backplane. Derzeit ist das bereits für bis hin zu 40 GbE spezifiziert, noch 2011 soll dies auch für 100 GbE stattfinden. Dazu wurde bereits ein „Call for Interest“ (CFI) abgesetzt und eine eigene Study Group ins Leben gerufen, die eine Standardisierung für diesen Aspekt vorantreiben soll.

Kupfer-Twinax ist nicht tot

Dagegen ist auf dem Kupfer-Twinax-Medium alles bereits von 1 bis zu 100 GbE normiert. Als neue Projekte stehen einige weitere Varianten zur Option: Bei Twinax-Verkabelungen (und im Bereich der Backplane) sind Konfigurationen für 100 GbE mit viermal 25 Gbit/s recht häufig gefragt. Anfang 2011 soll sich dazu bereits eine neue Study Group treffen. Die weitverbreitete Kupfer-Verkabelung nach dem Prinzip Twisted Pair geht bis zu 10 GbE, doch eine neue Gruppe wird sich um die Möglichkeiten kümmern, über Twisted Pair auch mit 40 und 100 GbE zu höheren Übertragungsgeschwindigkeiten zu gelangen.

Sie soll sich auf 40 und 100 GbE und auf Applikationen im Bereich der Rechenzentren fokussieren. Dabei wird die Entfernung sicher unter 100 Metern liegen. Mit Kupfer-Koaxialkabeln war dagegen bereits bei 10 Mbit/s Schluss.

Im Bereich der Glasfasern sind Multimode- (MMF) und Single-mode-Glasfasern (SMF) bereits entsprechend standardisiert. Hier sind künftige Entwicklungen zu erwarten, die bei MMF für noch weitere Entfernungen funktionieren sollen, oder die bei SMF sogar nur mit einer Wellenlänge arbeiten sollen. Auch hierzu wird ein CFI abgesetzt, der voraussichtlich im Juli 2011 definiert ist und sich von den Konfi-

gurationen zehnmals 10 Gbit/s beziehungsweise viermal 25 Gbit/s unterscheidet.

Zudem ist im Bereich der SMF-Übertragung an Anwendungen im Rechenzentrum gedacht. Sie stehen unter dem Aspekt der Optimierung in Bezug auf Energieeffizienz und Platzbedarf. Doch dazu ist wohl ein neues elektrisches und optisches Interface nötig. Das Thema energieeffizientes Ethernet ist bereits ratifiziert. Unter dem Begriff LPI (Low Power Idle) wird eine Reduzierung des Stromverbrauchs im Idle-Zustand angestrebt. Das ist derzeit nur für Kupferlösungen gedacht und für Übertragungsraten von 10 GbE und geringer. Doch die Hochgeschwindigkeits-Ethernet-Gruppen haben ihren Fokus auf die Energieeffizienz erweitert und umfassen dabei nicht nur die kupferbasierten Lösungen, sondern auch die optischen Übertragungsstrecken.

Die technologische Entwicklung für 40 und 100 GbE zeigt Abbildung 8. Dabei sind zusätzlich andere Varianten in der Diskussion, wie etwa 40 GbE über eine Entfernung von 40 km zu übertragen. Eine weitere Diskussion dreht sich bei 40 GbE um das Thema, 40 Gbit/s seriell zu übertragen. Dabei sind die Hersteller von optischen Übertragungskomponenten derzeit involviert.

Die Übertragungsraten der Signale sind liegt bei 10 Gbit/s (effektiv, wird die Codierung nach 64B/66B verwendet, bedeutet das eine Über-

Bit rate, Gb/s	Bit rate per lane, Gb/s	Number of lanes
100	10	10
	25	4
400	25	16
	40	10
	50	8
1000	25	40
	50	20

Zu viele Lanes sind bei den derzeitigen Übertragungsraten nötig, um schneller als 100 GbE zu übertragen (Abb. 9).

tragungsrate von 10,3125 GBit/s). Derzeit wird am Übergang zu 25 GBit/s (effektiv) gearbeitet. Dieser Prozess wird noch eine Zeit andauern. Zudem wollen die Anwender ihre Systeme im Feld möglichst lange behalten: Ideal ist für sie ein Upgrade, das die neuen Generationen unterstützt. Erst wenn das nicht mehr funktioniert, geht es an die Installation komplett neuer Systeme.

Vor allem bei den Einsatzbereichen im Rechenzentrum stehen einige Entscheidungen an. Es kommt voraussichtlich ein „Next Generation MSA“. Dabei dreht sich viel um die Frage bei der Konfiguration, ob zehnmal 10 besser ist als viermal 25. Unternehmen wie Google favorisieren die Variante zehnmal 10 GBit/s und wollen sogar 100 GbE im Masseneinsatz adaptieren. Somit steht auch die Forderung nach einem entsprechend niedrigen Preis für die nächste Generation im Vordergrund. Doch das ist ein schwieriger Trend für Ethernet – ganz anders als früher, als man 10 GbE standardisierte. Das stellte damals einen großen Schritt für die Technik dar, getrieben von den Herstellern.

Wer die Kosten für 100 GbE betrachtet, der wird Vorteile für die Konfiguration viermal 25 GBit/s erkennen: Das ist günstiger als zum Beispiel die zehnmal 10 GBit/s. Als Quintessenz erwarten Experten dadurch auch einen neuen MSA-Formfaktor. Bei einer Übertragungslänge von 10 Kilometern gilt CFP bereits als gesetzt, ebenso der QSFP. Eine neue Variante, der sogenannte QSFP2, soll etwa ab 2014 eine höhere Packungsdichte mit acht Ports zu je 100 GbE ermöglichen.

Die Nachfrage zu 100 GbE über 40 Kilometer ist speziell in Europa wichtig, doch sie findet in den USA keine große Unterstützung. Technologisch müsste man das über einen zusätzlichen Verstärker lösen, um die 100 GbE über diese Entfernung zu treiben. Doch dann steigt wieder die Latenzzeit auf der Übertragungsstrecke an. Und viele Anwendungen sind vor allem für diese Art von Verzögerungen anfällig. Wer den Ausblick bis ins Jahr 2015 wagt, der sieht schon erste Anfor-

derungen an 1 TBit/s aufziehen. So fordert man bei Google schon heute, dass bis 2013 die Generation mit 1 TBit/s, also TbE, nötig wird. Aufzeichnungen des Amsterdam Internet Exchange zeigen ebenfalls, dass deren Core Network bereits eine Bandbreite von 1 TBit/s und mehr benötigt. Das zeichnet sich auch bei vielen großen Internet-Austauschknoten ebenfalls so ab. 1 TBit/s reicht derartigen Anwendern bald nicht mehr.

Andere Projektionen von Unternehmen mit nicht so hohen Anforderungen sehen dagegen erst ab 2020 einen entsprechenden Bedarf. Die Standardisierer vertreten allerdings den verständlichen Standpunkt, dass „ein Standard speziell für einen Kunden“ nicht definiert wird. Daher steht die Standardisierungsgruppe vor der Frage, wann mit TbE zu starten ist. Dadurch erscheint die Forderung vieler Anwender nach einem „kostengünstigen 100 GbE“ mit höheren Packungsdichten als wichtiger, verglichen mit dem nächsten Schritt hin zu 1 TbE. Die Hersteller von Netzwerk-Equipment werden aller Voraussicht nach erst dann diesen Schritt angehen, wenn sich ein entsprechender Markt dafür abzeichnet.

Wer sich um die technische Machbarkeit kümmert, dem wird schnell klar, dass 25 GBit/s bei der elektrischen Signalisierung sich erst in der Entwicklung befindet und noch nicht einsatzbereit ist. Für TbE wäre sogar ein technologisches Umdenken nötig, denn die Signalisierung müsste anders erfolgen. Immer nur schneller Nullen und Einsen zu übertragen, lässt sich nicht so einfach umsetzen. Denn viele physische Parameter stoßen an die Grenze. Zudem steigt der Stromverbrauch massiv an, um die Signale wiederherzustellen.

Eine Lösung verspricht der Einsatz von Modulation – sprich es werden mehr Symbole pro Sekunde übertragen. Doch dabei steigt dann wieder die Latency an und der Stromverbrauch und der Aufwand für die Chipfläche steigen. Bei optischen Übertragungswegen könnte man mehrere Wellenlängen dazu nehmen. Ein anderer Ansatz ist das Parallelisieren – Fasern oder Leiter werden parallel dazu genommen. Die Abbildung 9 verdeutlicht das Dilemma: Je nachdem, wie schnell auf einer Lane übertragen wird (10 oder 25 beziehungsweise später mal mit 40 oder 50 GBit/s), wären bei 400 GbE 16, 10 oder 8 und bei 1TbE 40 oder 20 Lanes nötig. Das wirft enorme Probleme auf – zum Beispiel bei den Steckern.

Um eine gewissen Zukunftssicherheit zu bieten, hat sich die Standardisierungsgruppe zu 40 und 100 GbE in ihren Überlegungen von der Weiterverwendbarkeit einzelner Module leiten lassen. Alle Entscheidungen mussten sich auch den Fragestellungen unterziehen lassen:

- Was kann man weiter verwenden?
- Wie sehen die Building Blocks aus?
- Was bedeutet das für die Kosten beziehungsweise den Return on Investment?

Erst 100 GbE umsetzen und dann an TbE denken

Aus diesen Überlegungen leitet sich ab, dass die Debatte um die nächste Ethernet-Generation noch nicht ansteht. Es gilt zuerst, 40 und 100 GbE zum Durchbruch zu verhelfen. Dabei lässt sich aber auf einige Module zurückgreifen. Trotzdem zeichnet sich ab, dass TbE schwieriger wird als gedacht. Aber auch auf die Frage nach drahtlosen Alternativen in der Hochgeschwindigkeitsübertragung kommt von Experten die Aussage, dass sich zwar ursprünglich eine Gruppe dieser Thematik angenommen hat, doch das Resultat war eine Absage – auf absehbare Zeit eignet sich die Wireless-Technik nicht für derartig hohe Bandbreiten.

Rainer Huttenloher

40 UND 100 GbE BILDEN DIE BASIS FÜR SCHNELLE NETZE

Alles baut auf den einzelnen Modulen auf, die im Rahmen der Normierungsarbeiten für 40 und 100 GbE standardisiert wurden. John D'Ambrosia hat bei der IEEE P8023ba Task Force als Chairman die Normierung von 40 und 100 GbE mit gesteuert. Derzeit ist der Experte als „Acting Chair“ bei IEEE 100 GbE Electrical Backplane und der neu gegründeten „Cu Cable Study Group“ involviert.

Auf seiner Europa-Reise hat der Chief Ethernet Technologist von Force10 Networks einen Vortrag zum Thema „The state of 100G and Beyond“ gehalten. Dieser Beitrag basiert zum Teil auf seinen Ausführungen. Der in München im November 2010 gehaltene Vortrag wurde aufgezeichnet und in fünf Clips aufgeteilt. Ab dem zweiten Clip ist er auf YouTube zu finden (<http://www.youtube.com/watch?v=r6zCMDgQxDo>).



Über die Zukunft von 100 GbE und was danach kommt, referiert John D'Ambrosia von Force10 Networks.

Quelle: Rainer Huttenloher

Migration auf Microsoft Exchange 2010: Dell unterstützt bei der erfolgreichen Implementierung

Die Vorteile, die Microsoft Exchange 2010 bietet, sind weitreichend: von der Kostensenkung bis zur Erhöhung der Unternehmensproduktivität. Zentraler Nutzen der Lösung ist, dass Anwender produktiver arbeiten können, da sie die Flexibilität erhalten, von nahezu allen Plattformen, Browsern und Geräten über Standardprotokolle auf E-Mail, Voicemail oder Instant Messaging zuzugreifen. Bei der Einführung von Microsoft Exchange 2010 unterstützt Dell Unternehmen mit umfangreichem Consulting, mit Workshops, Design-, Implementierungs- und Deployment-Services.

Im Vergleich zur Vorgängerversion bietet Microsoft Exchange 2010 eine deutlich höhere Mobilität und Flexibilität. Mit dem neuen universellen Posteingang erhalten Anwender von überall aus Zugriff auf ihre gesamte geschäftliche Kommunikation. Das betrifft E-Mails, Voice-mails oder Textnachrichten. Weitere Optimierungen sind eine neue Benutzeroberfläche mit Touchscreen-Funktionalität, ein reduzierter Verwaltungsaufwand, integrierte Funktionen für die E-Mail-Archivierung und eine vereinfachte Notfallwiederherstellung. Exchange 2010 bietet zudem umfangreiche webbasierte Messaging-Funktionen bei den führenden Browsern Internet Explorer, Firefox und Safari. Empfehlenswert ist der Einsatz von Exchange Server 2010 auf einem Windows Server 2008 R2, denn die neue Exchange-Version ist für diese Betriebssystem-Umgebung optimiert.

Dell hat mit Microsoft Exchange 2010 mehrere Monate lang Performance-Tests auf den eigenen Server- und Storage-Plattformen durchgeführt. Microsoft Exchange Server 2010 wird heute von der gesamten Produktlinie der 64-Bit-PowerEdge-Server sowie von PowerVault-, EqualLogic- und Dell/EMC-Storage-Lösungen unterstützt.

Speziell für die Einführung von Microsoft Exchange 2010 hat Dell auch zwei Architektur-Modelle entwickelt, die den Einsatz der Lösung vereinfachen und die Bereitstellung beschleunigen. Das erste Modell heißt „Agile Consolidated Model“ und basiert auf einer Architektur, die sich für Kunden eignet, die anstatt spezieller Infrastrukturen für bestimmte Applikationen ein dynamisches Rechenzentrum mit Server- und Storage-Konsolidierung und -Standardisierung aufbauen wollen. In diesem Modell spielt die Server- und Storage-Virtualisierung eine zentrale Rolle. Die Grundlage für diese Variante bilden PowerEdge-Blade- oder -Rack-Server von Dell in Verbindung mit Speichersystemen der EqualLogic-PS-Serie oder Dell/EMC-SANs.

Basis des zweiten Modells mit dem Namen „Simple Distributed Model“ ist eine Server- und Storage-Infrastruktur, die darauf abgestimmt ist, Exchange 2010 möglichst einfach und kostengünstig bereitzustellen. Dieses Modell nutzt die Server- und Storage-Komponenten auf zweierlei Art: Es verwendet entweder einen PowerEdge-Server, der auf Grund seiner Speicherdichte keinen externen Speicher benötigt, oder einen PowerEdge-Standard-Rack-Server mit einem Direct-Attached-Storage-System.

Bei der Migration auf Exchange 2010 sollten sich gerade kleine und mittelständische Unternehmen für einen

breit aufgestellten Anbieter entscheiden, der ein One-Stop-Shopping ermöglicht und insbesondere über ein umfangreiches Serviceangebot wie Consulting, Installation und Integration verfügt. Dell startet jede Migration mit einem Planungs-Workshop, in dem gemeinsam mit dem Kunden die konkreten Anforderungen und Rahmenbedingungen ermittelt werden und eine entsprechende Migrations-Strategie erstellt wird. Anschließend werden in einer „Analyse & Design“-Phase die Details der Migration erarbeitet und alle notwendigen Maßnahmen festgelegt. Den Abschluss der Migration bildet das „Rollout & Deploy“-Projekt, in dem die Installation der Applikation durchgeführt wird.

Mit den neuen Referenz-Architekturen und den Services von Dell können Kunden die Komplexität und die Kosten des Deployments von Microsoft Exchange 2010 deutlich reduzieren. Die neuen Tools und Services erlauben eine rasche und effiziente Implementierung, so dass Anwender die Vorteile von Exchange 2010 schnell nutzen können, beispielsweise die optimierte Mobilität, eine größere Fehlertoleranz oder die bessere Voicemail- und E-Mail-Integration.



Meine Anhänger meistern
Straßenunebenheiten.
Auf meine Server kann
ich mich glücklicherweise
ebenfalls verlassen.

Andreas Gehlen, Gründer und
CEO von Zwei plus zwei



Weitere Infos unter: Dell.de/Exchange

Vielfältige Konfigurationen garantieren den Erfolg

Einsatz von 40- und 100-Gigabit-Ethernet

Im Juli 2006 vermeldete YouTube 100 Millionen „Hits“ pro Tag – sprich Seitenabrufe. Bis Mai 2010 stieg die Zahl auf zwei Milliarden täglich. Im Juni 2008 wurden bei iTunes täglich fünf Milliarden Songs heruntergeladen. Diese und ähnliche Anwendungen haben eine neue Art des Business geschaffen. Rückgrat dieser Applikationen sind Rechenzentren (RZ) mit einem steigenden Bedarf an Bandbreite und Speicherplatz. Das zeigt das Potenzial für die Hochgeschwindigkeitsübertragung auf Ethernet-Basis.

Vor fünf Jahren war kaum jemand in der Lage, die Flut an Informationen vorauszusagen, die heutzutage in Unternehmensnetzen gespeichert und übertragen wird. Virtualisierung spielt dabei eine immense Rolle. Unternehmen nutzen eine größere Anzahl an Soft- und Hardwareplattformen für ihre Geschäftsabläufe und integrieren in zunehmendem Maße Kollaboration, Video und andere fortgeschrittene Anwendungen. IT stellt nicht länger ein notwendiges Übel, sondern einen Wettbewerbsvorteil dar.

Die Lösung für diese schier unersättliche Gier der Informationstechnologie findet sich in der Erweiterung des bestehenden Ethernet-802.3-Protokolls zur Unterstützung von 40 und 100 Gigabit pro Sekunde (Gbit/s). Mit 40- und 100-Gigabit-Ethernet (40 GbE und 100 GbE) kommt eine neue Generation ins Spiel, die erheblich mehr Bandbreite bereitstellt und gleichzeitig die Kompatibilität der installierten Ethernet-Basis gewährleistet. Die 40-GbE-Spezifikation orientiert sich an High Performance Computing (HPC) und Speichersystemen und sollte sich besonders für die Rechenzentrums Umgebung eignen, da diese Server, Rechencluster, Blade Server aber auch Speichernetzwerke (Storage Area Networks, SANs) unterstützt.

Die 100-GbE-Spezifikation ist auf Core-Netzwerk-Applikationen ausgerichtet – Switching, Routing, Zusammenfassung von Rechenzentren, Internet-Knoten und Peering Points der Service-Provider gelten hier als die passenden Einsatzfelder. Insbesondere die Hochleistungs-Rechenumgebungen mit ihren extrem hohen Bandbreitenanforderungen wie Video-on-Demand (VoD) werden von den 100-GbE-Technologien profitieren.

Mit der Verabschiedung des 40/100-GbE-Standards durch die IEEE hat sich die Anzahl der möglichen Varianten der Hochgeschwindigkeitsübertragung erhöht. Der Standard IEEE 802.3ba spezifiziert folgende Parameter:

- 1 Meter über eine Backplane (40GBASE-KR4); derzeit nur für 40 GbE
- 10 Meter über Twinax-Kupferkabel (40GBASE-CR4/100GBASE-CR10); ursprünglich so geplant, jetzt verkürzt auf 7 Meter
- 100 Meter mit OM3 (40GBASE-SR4/100GBASE-SR10)
- 150 Meter mit OM4 (40GBASE-SR4/100GBASE-SR10)
- 10 Kilometer mit SMF (40GBASE-LR4/100GBASE-LR4)
- 40 Kilometer mit SMF (100GBASE-ER4)

Im Folgenden wird ein Überblick gegeben und ein Vergleich zwischen den verschiedenen Möglichkeiten der Hochgeschwindigkeits-Datenübertragung für moderne RZ angestellt.

Beim Übergang zu 40/100GbE besteht der wichtigste Unterschied für Backbone- und horizontale Multimode-Anwendungen in der Anzahl der verwendeten Fasern. 40GBASE-SR4 nutzt jeweils acht Fasern parallel über vier Übertragungspfade zum Senden und Empfangen. 100GBASE-SR10 nutzt dagegen zehn Übertragungspfade mit jeweils einer Sende- und einer Empfangsfaser, was insgesamt 20 Fasern ergibt.

Höhere Datenraten über Glasfaser

Bei SMF (Singlemode-Faser) werden derzeit nach wie vor zwei Fasern verwendet. Obwohl die Glasfaser an sich preiswerter ist, können die optischen und elektronischen Bauteile dafür 10-mal so teuer sei als das Equipment für die Kupferübertragung. Im RZ und im Backbone-Bereich ist es durchaus möglich, acht oder 20 Einzelfasern über verschiedene Glasfaserkabel bereitzustellen. Durch die Ungleichheit der Pfade werden allerdings erhebliche Laufzeitunterschiede auftreten (bekannt als Bitversatz), die zu Bitfehlern führen.

Aus diesem Grund treffen die 40/100-GbE-Standards Festlegungen zu Glasfaser-Trunk-Kabeln mit MPO (Multi-Fiber Push-on) beziehungsweise MTP (Mechanical Transfer Push-on) Mehrfaser-Steckverbindern. Bei diesen Kabelbaugruppen besitzen alle Fasern die gleiche Länge – auch als „parallele Optik“ bezeichnet – was Bitversatz/Laufzeitunterschiede minimiert.

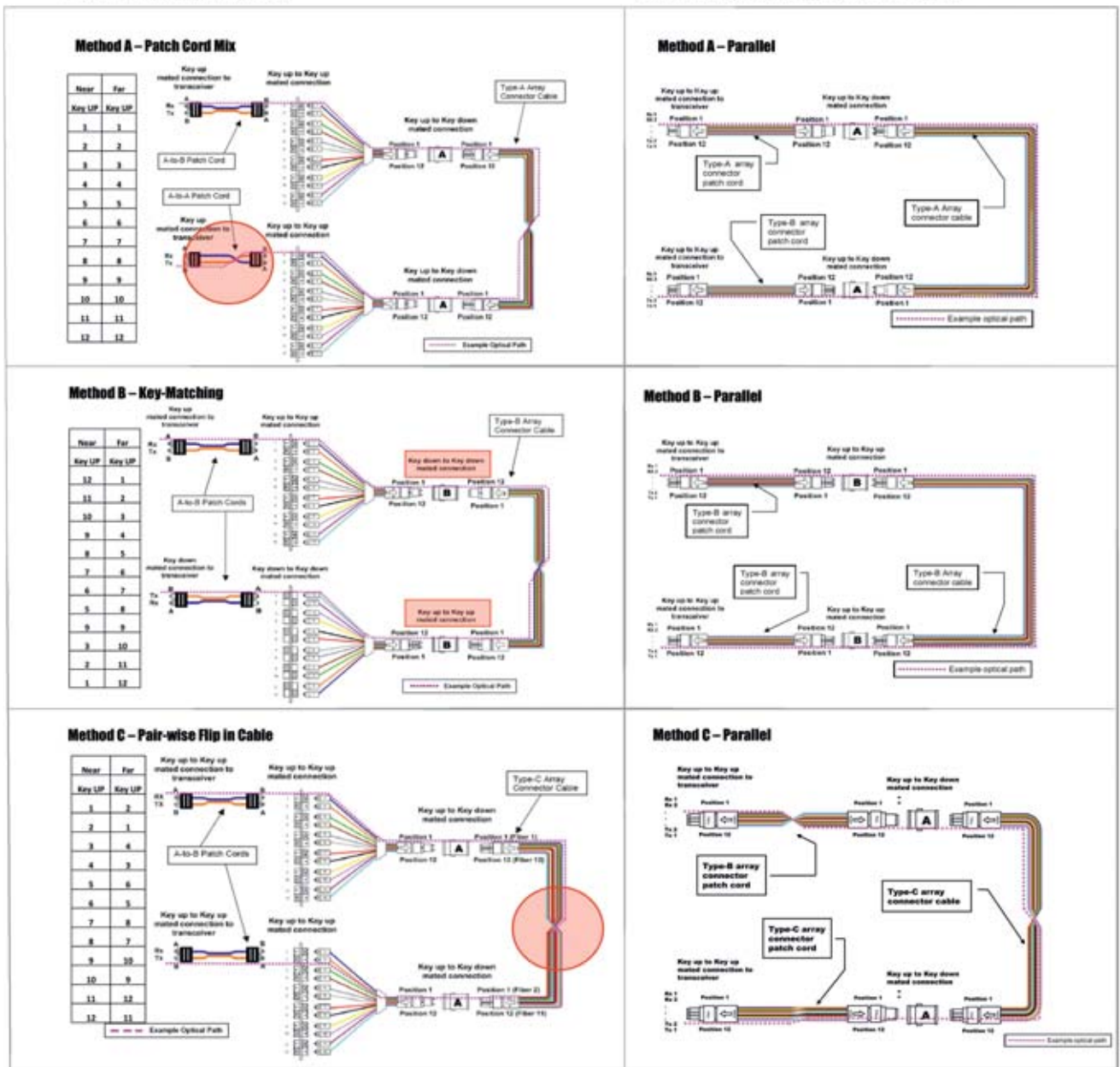
MPO- und MTP-Steckverbinder gibt es in der Anschlusskonfiguration mit 12 und 24 Fasern am Ende des Trunk-Kabels. MPO/MTP-Trunk-Kabel unterstützen ebenfalls 10GBASE-SR/SX Anwendungen, obwohl nur zwei Fasern genutzt werden. In diesem Fall werden die Trunk-Kabel an Kassetten beziehungsweise Hydra-Kabel angesteckt, welche die Mehrfachfasern in zweifaserige Verbindungen aufteilen (typischerweise mit LC- oder SC-Steckern).

Ein weiterer Unterschied bei den Highspeed-Glasfaserkonfigurationen ist die Polarität. Bei zweifaserigen Anwendungen, wie die 10-GbE-Übertragung, ist das Managen der Polarität einfach durch das Kreuzen der Fasern auf der Übertragungstrecke erledigt. Dabei sind drei verschiedene Methoden definiert (siehe Tabelle auf Seite 16).

40GBASE-SR nutzt acht Fasern eines zwölfaserigen MPO/MTP-Trunk-Kabels, (viermal Senden und viermal Empfangen). Die mittleren vier Fasern im MPO/MTP-Steckverbinder bleiben dabei unbeschaltet. Für die Schnittstelle am Equipment wird ein MPO/MTP-Mehrfaser-Steckverbinder statt eines herkömmlichen LC-Steckers benötigt.

Two Strand Polarity Illustration

Parallel Optics Polarity Illustration (12 strand)



Quelle: Siemon

Die Methoden A, B und C zeigen, wo auf der 40-/100-GbE-Übertragungsstrecke das Vertauschen der Polarität stattfinden kann (Abb. 1).

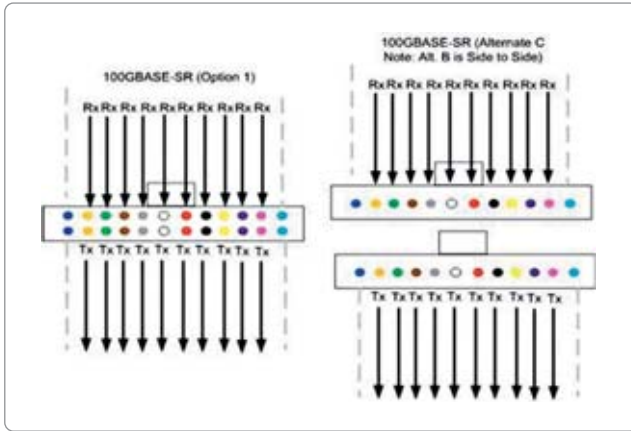
Für 100 GbE gibt es drei anerkannte Übertragungsverfahren, unter anderem mit einem 24-faserigen (links abgebildet) beziehungsweise zwei 12-faserigen Trunk-Kabeln in der Konfiguration als „over and under“ (doppelstöckig) oder „Side-by-Side“ (nebeneinander) (vergleiche Abbildung 2; dabei ist die Side-by-Side-Konfiguration nicht abgebildet). Bei der Übertragung werden zehn Fasern zum Senden und zehn zum Empfangen genutzt, wobei die äußeren ungenutzten Fasern unbeschaltet bleiben.

Higher Speed Infiniband

Bei Infiniband (IB) handelt es sich um eine Industriestandard-Spezifikation. Sie definiert ein Protokoll, das in High Performance Computing

Clustern zur Verbindung von Servern, Switchen, Speichersystemen und Embedded Systems verwendet wird. Infiniband ist eine latenzarme Fabric-Architektur mit hoher Quality of Service (QoS), die geschaltete Punkt-zu-Punkt-Verbindungen mit Datenraten von gegenwärtig maximal 32 Gbit/s über Glasfaserverbindungen realisiert beziehungsweise zulässt. Die direkt angeschlossene Architektur ermöglicht es einem „Host Bus Adapter“ (HBA), direkt mit einem „Target Channel Adapter“ (TCA) zu kommunizieren und erweitert damit im Wesentlichen den Bus eines Servers oder Speichergerätes.

IB nutzt ein ähnliches Verfahren für Multilane-Datenverkehr wie 40/100 GbE und wird über einen QSFP-Steckverbinder angeschlossen. Eine Reihe von Schnittstellen unterstützt derzeit Infiniband, wobei auch Übertragungspfade zum Einsatz kommen, bei denen ein Pfad



Die Signalbelegung beim Standard 100 GBASE-SR: zehn parallele Lanes sind zu beherrschen (Abb. 2)

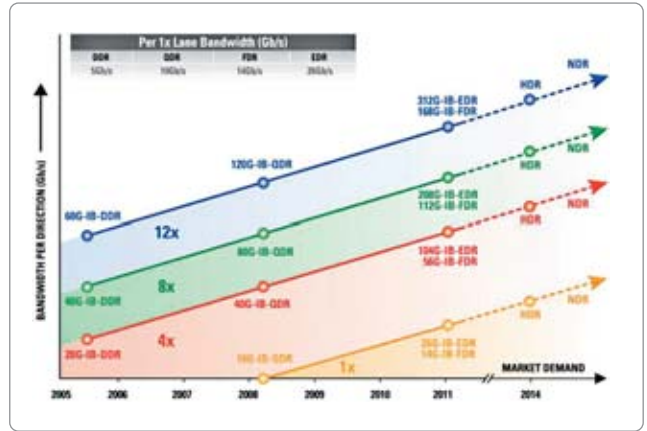
zwei Fasern entspricht (einmal für Senden und einmal für das Empfangen). MPO/MTP Trunk-Kabel können über originale Infiniband-Steckverbinder oder die Fasern entsprechend mit Hybridkabel angeschlossen werden.

Die Infiniband 4x-Spezifikationen (4x steht für 4 Lanes) definieren gegenwärtig drei Datenraten:

- eine 8 GBit/s Single Data Rate (SDR) mit 2,5 GBit/s pro Pfad,
- eine 16 GBit/s Double Data Rate (DDR) mit 5 GBit/s pro Pfad und
- eine „Quad Data Rate“ (QDR) mit 10 GBit/s pro Pfad.

Zudem sind „12x“ (12 Lanes) Infiniband-Technologien definiert worden, die maximal 96 GBit/s pro Channel ermöglichen. Für die nahe Zukunft werden zwei weitere Infiniband-Technologien erwartet: „FDR“ (Fourteen Data Rate) mit 14 GBit/s pro Pfad und „DER“ (Enhanced Data Rate) mit 26 GBit/s pro Pfad.

Ein weiteres Protokoll ist Fibre Channel (FC), das derzeit als Standard-Protokoll für Speichernetzwerke (SAN) in Unternehmens-RZ gilt und das sich ebenfalls in Richtung höhere Übertragungsraten entwickelt. Beim Fibre Channel kommen ebenfalls Kupferschnittstellen zum



Die Entwicklungsplanung für das Infiniband-Protokoll zeigt bei der Bandbreite nach oben (Abb. 3).

Einsatz. Der Standard umfasst Speicherung, Bearbeitung, Übertragung, Anzeige, Management-Organisation und Abrufen von Informationen.

Die Weiterentwicklung beim Fibre Channel

Traditionell verwenden Server gesonderte Adapter für Speicherverkehr und Netzwerkverkehr – Host Bus Adapter (HBA) für Speicherverkehr und Netzwerkkarten (NICs) für den LAN-Verkehr. Hybridadapter gibt es für FCoE (Fibre-Channel over Ethernet) mit einem Port, der als „Fibre Channel HBA“ fungiert (für den auf herkömmliche Weise angeschlossenen Fibre Channel) und einem zweiten Port für Ethernet. Dabei ist der Ethernet-Port der Schnittstelle der gleiche wie beim herkömmlichen kupfer- und glasfaserbasierten Ethernet.

Moderne „Converged Network Adapter“ (CNA) besitzen HBAs für den Fibre Channel und konventionelle Ethernet-Netzwerkkarten in einem oder mehreren Ethernet-Standard-Ports, über die sowohl die Speicherung als auch der Netzwerkverkehr erfolgen. Durch die Nutzung von Server-CNAs reduziert sich die benötigte Anzahl von Server-Adaptern,

Rittal – Das System.

Erleben Sie „Rittal – Das System.“ live

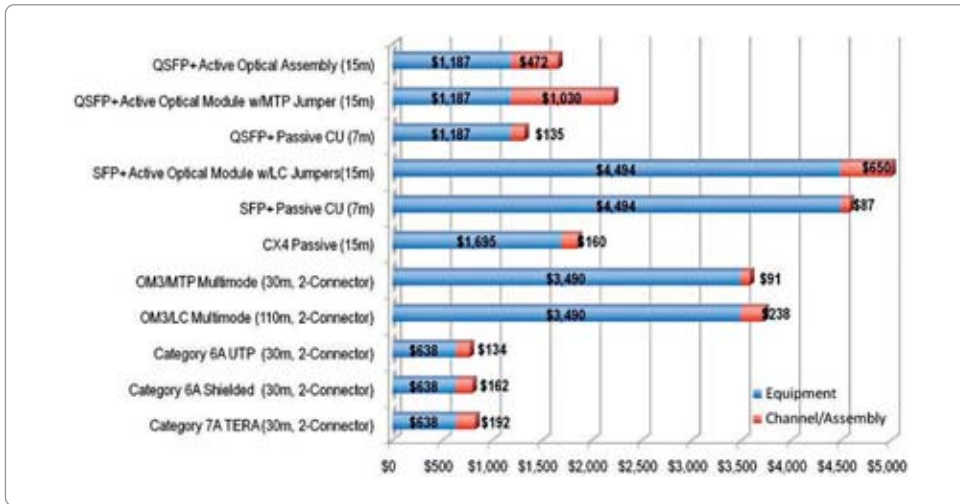


CeBIT Hannover
1. bis 5. März 2011
Halle 11, Stand E06

SCHALTSCHRÄNKE

STROMVERTEILUNG

KLIMATISIERUNG



Der Kostenvergleich für die verschiedenen Anschluss-Techniken; Kosten pro Port (Abb. 4)

Quelle: Siemon

wodurch wiederum Eingangs-/Ausgangs- (I/O)-Kabel eingespart werden und sich die Zahl der verwendeten Switchports verringert. Diese Konfiguration spart Hardware ein, vereinfacht die I/O-Konfigurationen der Server, senkt den Stromverbrauch und die Gesamtkosten.

High-Speed Interconnect Kupferkabel für alle Anwendungen

Fibre Channel unterstützt darüber hinaus steckbare Twinax-Kupferkabel, die über Transceiver (SFP+, Small Formfactor Pluggable Transceiver Plus) angebunden sind (SFP+ gibt es sowohl für optische als auch für kupferbasierte Übertragungslösungen). Damit kommt eine energie- und kostensparende Lösung für Anwendungen über kurze Entfernungen ins Spiel. Diese Highspeed-Verbindungskabel finden typischerweise in Top-of-Rack-Anwendungen Einsatz.

Die erste 10-GbE-fähige Kupferschnittstelle ist für die 10GBASE-CX4 (IEEE 802.3ak) Applikation entwickelt worden (bereits 2004 veröffentlicht). Diese Schnittstelle ist kompatibel zu X2, Xenpack und XPack

MSAs (als MSAs werden Multi-Source Agreements zwischen konkurrierenden Herstellern zu standardisierten Formfaktoren einer Schnittstelle bezeichnet). Die physikalischen Anforderungen an diesen geschirmten Vierpfad-Kupfer-Steckverbinder sind nach SFF-8470 genormt. Als eine passive Baugruppe besitzt das SFF-8470/CX4 Kabel (siehe Abbildung 3) eine Reichweite von 15 Metern und unterstützt 10 GbE, Infiniband, Fibre Channel und FcoE.

Andere serielle Übertragungen werden für I/O-Funktionen wie SATA (Serial Advanced Technology Attachment), SAS (Serial Attached SCSI) und RapidIO ebenfalls unterstützt. Bei diesen Kabeln handelt es sich um Twinax-Kabel, die aus zwei Innenleitern bestehen und von einem Folienschirm und einem Geflechtsschirm umgeben sind. Aufgrund der geringen Latenz-Zeiten über diese Verbindungen werden diese Kabel gern in Supercomputing Clustern, beim High Performance Computing und bei Storage-Anbindungen verwendet.

Kabelbaugruppen nach SFP+ für Geschwindigkeiten von 10 GBit/s sind relativ neu am Markt und werden als Kupfer- und Glasfaservariante angeboten. Die „SFP+ Interconnect-Kabel“ unterstützen die Protokolle

Schneller – besser – überall.



IT-INFRASTRUKTUR

SOFTWARE & SERVICE



www.rittal.de



Quelle: Siemon

Aktive optische Kabel für Hochgeschwindigkeitsverbindungen verbessern die Ausnutzung der Switch-Ports (Abb. 5).

nach Ethernet, Infiniband und Fibre Channel. Kupferbasierte Twinax-Kabel sind eine Option für „Top of Rack“-Anwendungsfälle (ToR) oder bei kurzen Distanzen für „Middle of Row“-Konfigurationen (MoR), während die Glasfasermodule mit SFP+ Anwendungen über mittlere bis lange Distanzen abdecken. Die weitgehende Standardisierung für SFP und SFP+ ist durch ein MSA zwischen konkurrierenden Herstellern geregelt.

Passive Kupferkabel für SFP+ im 10-GbE-Umfeld erfreuen sich bei Kurzstrecken als steckbare 10-Gbit/s-Anwendungen an zunehmender Beliebtheit und haben eine Reichweite bis 7 Metern bei Ethernet-Applikationen. Bei anderen Protokollen lassen sich damit auch größere Entfernungen überbrücken. Bei den aktiven Kabeln für SFP+ im 10-GbE-Umfeld verstärkt die in den Steckverbinder integrierte Elektronik das Signal. Daher können längere und auch dünnere Kabel verwendet werden als es bei den steckbaren passiven Kupferkabeln möglich ist.

Aufgrund der begrenzten Reichweite der Kabel für SFP+ eignen sich für ToR-Switch-Konfigurationen oder zum Switchen benachbarter Racks, im Bereich HPC und für Rechenclustern sowie für Server und Speicher. Mit dem Aufkommen der Steckverbinder für SFP+ werden CX4-Kabel typischerweise für Infiniband (über Kupfer) eingesetzt. Diese Steckverbinder wurden bevorzugt für das ToR-Switching genutzt, bis die effizientere Switching-Hardware nach 10GBASE-T auf den Markt gekommen ist.

Die Weiterentwicklung: Quad Small Form-factor Pluggable Plus

Eine Weiterentwicklung bei den Verbindungskabeln sind Kabelbaugruppen für QSFP+ (Quad Small Form-factor Pluggable Plus). Diese Vierpfad Hochgeschwindigkeitsverbindungen unterstützen ähnliche Anwendungen wie SFP+ und sind für hochdichte Anwendungen bis zehn Metern bei einer Datenrate von 10 Gbit/s pro Übertragungspfad

entwickelt worden. Eine Kabelbaugruppe kann bis zu vier standardmäßige Verbindungen nach SFP+ ersetzen, bietet eine höhere Dichte und reduzierte Systemkosten. Der IEEE-Standard 802.3ba für 40 GbE erkennt die Verwendung von derartigen Steckverbindern an (auch bekannt als Style 1, SFF-8436 Rev 3.4). Für Style 1 spezifiziert Multilane 40GBASE-CR4 die Verwendung von Vierpfad-QSFP+-Kabeln, um diese Geschwindigkeit für Entfernungen bis 7 Metern zu erreichen.

Aktive optische Kabel für Hochgeschwindigkeitsverbindungen

Aufgrund der Reichweitenbegrenzung der Punkt-zu-Punkt-Verbindungskabel kann die Port-Überbuchung sich bei ToR-Konfigurationen zu einem Problem auswachsen. Stromverbrauch und Kühlung gelten als die limitierenden Faktoren, die über die mögliche Anzahl der Aktivkomponenten in einem Schrank entscheiden. Um die Switchport-Nutzung innerhalb der Schrankreihe im RZ zu verbessern, ist ein aktives optisches Kabel (AOC) entwickelt worden. Es behält das den elektrischen Steckverbinder nach SFP+ bei und besitzt zusätzlich einen optischen Transceiver im Steckverbinder – zur optischen Übertragung bis zum Switch des entfernten Racks, wo das optische Signal im Steckverbinder wieder in ein elektrisches Signal umgewandelt wird.

Eine Bewertung der jeweiligen Technologie sollte anhand ihrer Interoperabilität erfolgen. Die Vorzüge der Applikation, die Skalierbarkeit und die Wartung sowie die Gesamtkosten der Übertragung pro Port gelten dabei als wichtige Kriterien. In diese Analyse sollten die Kabelbaugruppen, Kosten für Switch-Port und Netzwerk-Schnittstelle der Server oder Speichergeräte eingehen. Während diese Lösungen eine geringe Verzögerungszeit (Latenz) aufweisen und die Kosten der horizontalen Kupferverkabelung zu senken vermögen, werden sie dennoch meist nur für spezifische Anwendungsbereiche im RZ eingesetzt – begründet durch die höheren Kosten für Aktivtechnik und Wartung. Die Mehrkosten bei diesen Kabelbaugruppen und aktiven Ports veranlassen die meisten RZ weiter für 10GBASE-T zu verkabeln.

Beim Entwurf eines neuen Rechenzentrums sollte man MPO/MTP-Glasfaser-Trunk-Kabel für die Verbindungen im Backbone einplanen, damit der Übergang auf 40/100 GbE ohne die Verlegung neuer Übertragungskanäle gewährleistet ist. Die traditionellen Kupfer-Channel sind für das Monitoring, zentrales KVM (Keyboard, Video, Mouse), Management et cetera einsetzbar. Kupfer-Channel sollten für Übertragungen nach 10GBASE-T ebenfalls verlegt werden. Sowohl TIA 942-2 als auch ISO 24764 empfehlen im Minimum eine Kategorie-6A/Klasse-EA-Verkabelung. Bei einer „Any to All“-Verkabelungsvariante, die auch 10GBASE-T unterstützt, lassen sich Überbuchung und kostspielige Umkonfigurierungen (Moves, Adds und Changes, MAC) vermeiden.

*Carrie Higbie,
ist Global Director Datacenter Solutions
and Services bei Siemon.*

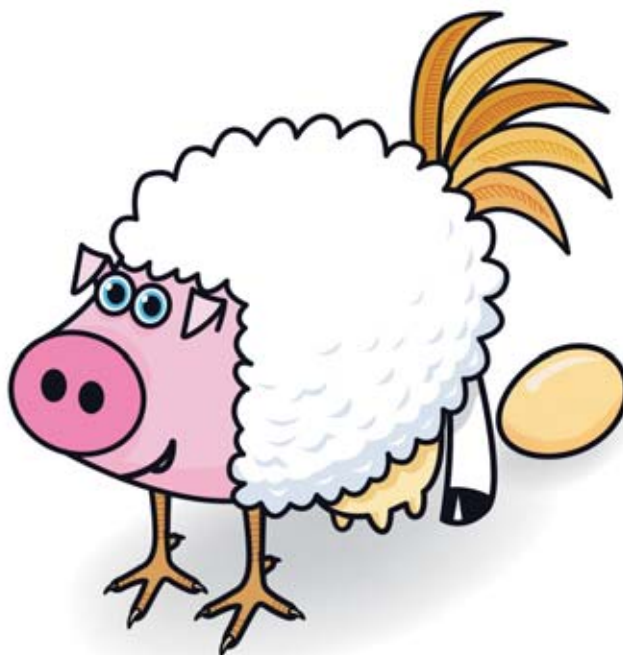
OPTIONEN FÜR DAS MANAGEN DER POLARITÄT

Method	Trunk	Cassette	Jumpers
A	same at both ends	same at both ends	polarity reversed at one end
B	same at both ends	Polarity is reversed in the cassette.	same at both ends
C	Polarity is reversed in the trunk.	same at both ends	same at both ends

Quelle: Siemon

Tabellarische Darstellung der Optionen für das Managen der Polarität; das Kreuzen der Fasern muss an einer beliebigen Stelle auf der Übertragungsstrecke erfolgen (Abb. 6).

Sind Sie als Unternehmen auch immer auf der Suche nach der Eierlegenden Wollmilchsau für Ihr Rechenzentrum?



Entdecken Sie, wie die Avocent Lösungen von Emerson Network Power Ihre Eierlegende Wollmilchsau werden können

Treffen Sie die RZ Sau auf
WWW.RZSAU.COM

Und mit etwas Glück gewinnen Sie sogar ein iPig.

Mit freundlicher Unterstützung unseres Distributionspartners:



Überlieferte Werte gehören auf den Prüfstand

35 Grad Celsius Raumtemperatur im Rechenzentrum sind machbar

Drei Faktoren ergeben einen energieeffizienten Rechenzentrumsbetrieb: das IT-Equipment, die Hilfsaggregate und die Energie-Bereitstellung. Nur wer alle drei Bereiche im Griff hat, wird die optimale Lösung finden.

Vor allem bei kleineren Rechenzentren und Serverräumen in Unternehmen lautet die übliche Vorgehensweise beim Ausbau der zentralen IT-Ressourcen: Es wird ein zusätzlicher Server plus Storage angeschafft. Doch im Kielwasser folgt die nächste Anforderung: „Nun benötigen wir noch eine neue Klimaanlage, um die zusätzliche Abwärme durch die neu hinzugekommenen IT-Systeme abführen zu können. Und interessanterweise wird nur ganz selten diese vermeintlich logische Konsequenz hinterfragt.“ Dieses Szenario hat Harald Rossol, Vorstand der Erecon AG, in seinem Vortrag zu den aktuellen Herausforderungen beim Betrieb von Rechenzentren (Vortrag auf dem Heise-iX-Event in Hamburg zu Verkabelung und Rechenzentrums-Infrastruktur) geschildert.

Fragt man nach, lautet die Antwort der Hersteller in den meisten Fällen, dass beim Überschreiten einer bestimmten Temperatur – meist werden 20 und 21 Grad Celsius genannt – die IT-Systeme kaputt gehen, und man daher für eine entsprechende Kühlung zu sorgen habe. Doch es gibt keinen technischen Beleg, warum die Temperatur im Serverraum 21 Grad sein muss. Vor dieser Herausforderung stand auch die Erecon. Bereits 2003 wurde ein Test gefahren: Es kam dabei zu keiner Neuinvestition in eine Klimaanlage, dafür wurde ein Testprojekt initiiert, das die Temperatur im Server und im Serverraum genau untersucht hat. Dazu wurden Messfühler im Server eingebaut. Das Ergebnis dieses Projekts war interessant: Die Raumtemperatur und die

Temperatur im Server selbst sind voneinander nicht linear abhängig. Der Aufwand für die Kühlung nach dem Motto „man steckt ein Kilowatt (kW) für die Rechenleistung rein und dazu noch ein kW, um den Server zu kühlen“ reicht oft nicht aus – das Verhältnis ist meistens schlechter. Im Vergleich dazu gibt es laut Harald Rossol keinen anderen technischen Bereich, in dem ein derartiges Verhalten wie in der IT nötig ist. Das Ziel bei Erecon war, das Verhältnis nicht mehr bei 1 : 1 zu belassen, sondern auf ein Verhältnis von 1 : 0,3 – oder besser – zu kommen.

In 2008 hat die IT weltweit den Kohlendioxid-Ausstoß des Luftverkehrs überholt. Die starke Steigerung des Energieeinsatzes zeigt die Dramatik: Im Jahr 2000 waren nur ein bis zwei Prozent der Gesamtausgaben im Bereich der IT für Energie nötig, 2008 bereits zehn Prozent und heute etwa zwischen 12 bis 15 Prozent. Wenn es so weiter geht, werden in zwei bis drei Jahren 18 bis 20 Prozent erreicht. Doch die IT-Verantwortlichen stehen damit vor einem Problem, denn in Deutschland wissen mehr als 90 Prozent von ihnen gar nicht, so Rossol, wie viel Energie für den Betrieb ihrer IT-Umgebung inklusive Kühlung notwendig ist, was der Stromverbrauch kostet und welchen Kohlendioxid-Ausstoß das entspricht. Der Grund dafür ist simpel: Die Energiekosten tauchen in den IT-Budgets gar nicht auf.

Dieses Problem wird sich noch verschlimmern. Vor allem das Wachstum der Speichersysteme nimmt den Hauptposten ein, denn sie sind die großen Energiefresser. Zudem kommen mehr Anforderungen auf die IT zu, sprich es gilt mehr Systeme anzuschaffen und mehr Applikationen zu betreiben. Zwar hilft die Virtualisierung, denn damit lassen sich mehrere Anwendungen auf einem physikalischen System betreiben – es sind dann im Zuge der Konsolidierung auch bis zu 20 Applikationen auf einem System möglich. Doch das spart nicht im selben Maß, denn zusätzliche Energie wird für den entsprechenden Systemausbau (mehr Arbeitsspeicher, mehr Storage) nötig.

Stellschrauben zur energieeffizienten Optimierung im Rechenzentrum

Die Steigerung der Energiekosten, wie sie die großen Anbieter in Deutschland derzeit durchsetzen, verschärft das Problem auf der Kostenseite. Doch es gibt Lösungsmöglichkeiten, allerdings, so Rossol, „müsse man die Thematik umfassend adressieren“. Die Maßzahl mit der Angabe in Terawattstunden für den Energieverbrauch der IT sei dazu wenig geeignet: Ende 2010 sind in Deutschland etwa fünf Millionen Server in RZ und Serverräumen im Einsatz. Für deren Energieversorgung sind vier Atomkraftwerke nötig – die Hälfte der Energie entspricht dabei dem Kühlungsaufwand.

CHECKLISTE MIT OPTIMIERUNGSMÖGLICHKEITEN

1. IT-Infrastruktur: Anordnung der Komponenten (wie Kaltgang-/Warmgang-Konzept und Verblechung der Leerräume im Rack), Effiziente Datenhaltung (nicht Benötigtes entfernen), Energiemanagement Betriebstemperaturen/Wirkungsgrade/Systemauswahl (Erecon betreibt das eigene Rechenzentrum mit 35 Grad Celsius Raumtemperatur), Konsolidierung/Virtualisierung (schnelle Erfolge: ein Server benötigt bei 10 Prozent Last fast genauso viel Energie wie bei 90 Prozent Auslastung)
2. Kühlung/Klima: Gebäude/Räume, kontrollierte Anhebung der Raumtemperatur, energieeffiziente Kühlung und Klimatisierung, intelligente Messung und Regelung Nutzung von Abwärme
3. Energie: Nutzung von Abwärme, intelligente USV-Systeme mit hohem Wirkungsgrad, Energielieferung (Energy Contracting)

Einzelne Faktoren bestimmen die Gesamteffizienz. Daher müssen in den Auswertediagrammen auch die einzelnen Faktoren – in Abhängigkeit vom Zielwert und den maximal erreichbaren Werten darstellen lassen (Abb. 2).



Quelle: Ereon

Bei den Temperaturvorgaben für den Betrieb von IT-Systemen haben die Hersteller sich einen großen Sicherheitsbereich geschaffen. Angaben wie die 21 Grad Celsius sind sonst nicht zu erklären. Die spezifizierte Temperatur für die Komponenten, die in einem Server verbaut sind, und die Raumtemperatur im Rechenzentrum oder Serverraum weichen massiv voneinander ab. Für die Prozessoren sind 70 bis 75 Grad Celsius angegeben, die DRAM-Bausteine verkräften 90 Grad Celsius, die Festplatten 50 Grad Celsius: Das „passt“ nicht zur Raumtemperatur von 21 Grad Celsius. Damit besteht aber auch viel Flexibilität.

Rossol sieht generell drei Bereiche, an denen man ansetzen kann:

- IT-Equipment (Server, Switches, Storage-Systeme),
- Hilfsaggregate (Kühlung und Klima) sowie
- die Energiebereitstellung selbst.

Zum letzten Punkt gehört das Einführen des Stroms über die Unterbrechungsfreien Stromversorgungen (USV). Sie kann einen Wirkungsgrad von mehr als 96 Prozent haben – doch den erreicht sie in der Realität leider fast nie. „Die meisten USVs, die wir in unserer Beratungstätigkeit gesehen haben, sind nicht optimal angeschlossen und überdimensioniert. Sie arbeiten in falschen Lastbereichen, und so weiter“, so Rossol. Damit haben die USVs oftmals nur einen Wirkungsgrad von 80 Prozent. Daraus ergibt sich dann: Über die Hauptverteilung und das USV-System – wenn es gut läuft – gehen zehn Prozent verloren, wenn es schlechter läuft, fallen sogar 20 bis 30 Prozent Verluste an.

Erst eine Potenzialanalyse könne zeigen, so Rossol, wie die einzelnen Wirkungsgrade sich zusammensetzen. Wer hier Aufwand spen-

dert, der kann dem Ansatz „Einfach alles weg kühlen“ eine Absage erteilen. Hier gilt es, bereits bei der Ausschreibung anzusetzen. Um die bestmöglichen Teilwirkungsgrade zu erzielen, ist bereits im Einkaufsvorgang bei den IT-Herstellern darauf zu achten. Das kann sich als recht komplex erweisen, denn das geht in Richtung Organisation.

Zum Beispiel lässt sich bei den Serverherstellern sehr wohl angeben, dass sie Server mit effizienten Netzteilen liefern sollen, die bereits einen Wirkungsgrad von 95 Prozent schon bei 50 Prozent der Nennlast aufweisen. Das können die Hersteller schon seit Langem – doch der Anwender muss es verlangen. In der Checkliste (auf Seite 18) finden sich dazu einige Beispiele. Dazu bildet sich dann auch ein Regelkreis heraus – Potenzial erkennen, Konzept aufstellen, Realisierung angehen und dann die weitere Betreuung.

Überwachung der Maßnahmen

Generell ist ein laufendes Monitoring nötig, um die Wirksamkeit der energetischen Maßnahmen zu prüfen und weiter zu verfeinern. Denn es sind immer die schwächsten Glieder in der Kette zu bestimmen. Diese Faktoren beeinflussen die Gesamteffizienz. Dazu gibt es verschiedene Metriken. Abbildung 2 zeigt ein Diagramm, in dem die Ist-Werte zu den Zielwerten und zum Nonplusultra – den Best Practices – für verschiedene Parameter zu sehen sind. Damit lässt sich sehr schnell zeigen, wie weit man in seinen Optimierungsbestrebungen gekommen ist.

Neben den Einsparungsmöglichkeiten bei den Energiekosten, stellen sich weitere positive Aspekte ein. Von den Erfahrungen aus Optimierungsprojekten der Ereon berichtet Harald Rossol, dass sich in den meisten Fällen auch eine bessere Ausnutzung der Ressourcen einstellt. Weitere Optimierungen betreffen den Abbau der Umgebungs-komplexität. Einige Systeme werden nicht mehr benötigt. Die Wartung und auch der Betrieb der IT-Umgebung werden einfacher. Ein Beispiel dazu ist der Ausfall der Klimaanlage für das Rechenzentrum. Dann setzt der Betrieb binnen kurzer Zeit aus. Wer dagegen ein Projekt aufsetzt – wenn etwa die Server 35 bis 40 Grad Raumtemperatur aushalten – und in dem gar keine Klimaanlage mehr benötigt wird, der hat derartige Szenarien nicht zu fürchten. Das lässt sich heute – unter Berücksichtigung einiger Randbedingungen durchaus erreichen.

Rainer Huttenloher



Ansatzpunkte für den energieeffizienten IT-Betrieb (Abb. 1)

Quelle: Ereon

Arbeitsspeicher im Server frisst Energie

DRAM-Technologie und Fertigungsprozess sind die bestimmenden Faktoren

Wer mit der Server-Virtualisierung die Ressourcen in seinem Rechenzentrum besser ausnutzen will, der muss mehr Arbeitsspeicher in seine Server stecken. Damit entwickelt sich der Arbeitsspeicher zur bestimmenden Größe für die Leistungsaufnahme des gesamten Servers. Mit der aktuellen Generation von Dynamic Random Access Memory (DRAM) lässt sich die Energieaufnahme allerdings reduzieren.

Die Datenmengen und die Aufgabenvielfalt im Rechenzentrum (RZ) wachsen. Um einen möglichst effizienten Betrieb zu gewährleisten, wird der Ausnutzungsgrad der Industrie-Standardserver massiv erhöht. Hier bietet sich in erster Linie die Virtualisierung an – wo früher ein Server mit einer Applikation und einem oftmals nur einstelligen Ausnutzungsgrad für die CPU-Rechenleistung vor sich hinarbeitete, sind nun 20 und mehr virtuelle Maschinen über die Hypervisor-Virtualisierung – sei es von VMware, Microsoft oder Citrix – auf einer physikalischen Serverhardware im Einsatz.

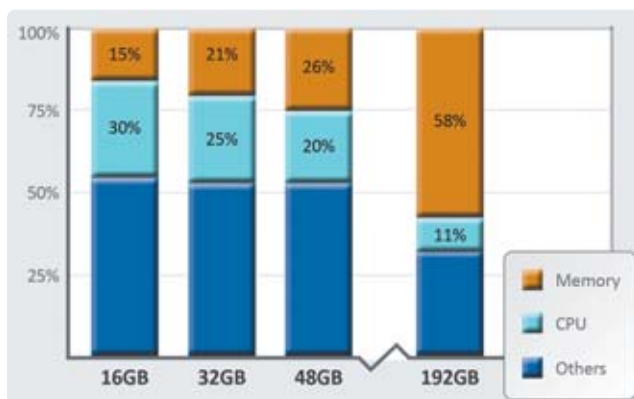
Virtualisierung verlangt mehr Arbeitsspeicher

Eine effiziente Kühlung der IT-Systeme im RZ ist wichtig, hier sind vielfältige Design-Kriterien zu berücksichtigen. Generell aber gilt, je weniger man sein IT-Equipment kühlen muss, umso besser sieht die Energiebilanz aus – ganz zu schweigen vom ökologischen Gewissen. Doch wenn die Virtualisierung von Servern zum Einsatz kommt, zieht das eine logische Konsequenz nach sich: Damit die bestehenden Ressourcen besser ausgenutzt werden, müssen die Industrie-Standardserver in bestimmten Bereichen gut ausgerüstet sein. Speziell die Be-

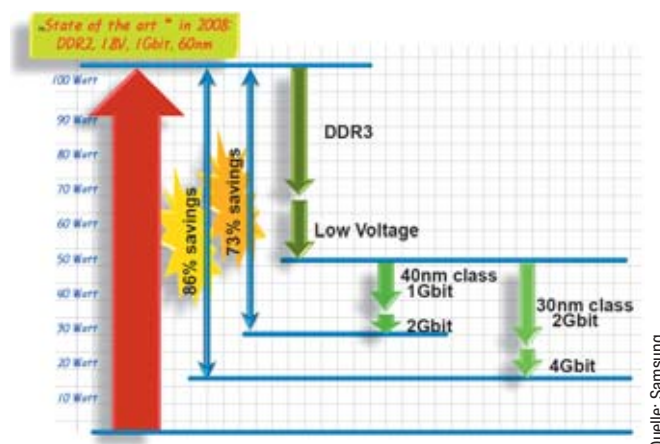
reiche Arbeitsspeicherausbau und Ein-/Ausgabe müssen für mehr Durchsatz ausgelegt sein. Nur wenn entsprechend viel DRAM-Kapazität im System zur Verfügung steht, lassen sich 20 und mehr virtuelle Maschinen auf einem Server betreiben.

Teilbereich Memory braucht mehr Strom

Vor dem Einsatz der Server-Virtualisierung waren Systeme mit durchschnittlich 16 GByte Arbeitsspeicher in weiten Einsatzbereichen gut versorgt. Zu diesen Zeiten entfielen weniger als 15 Prozent des Energieverbrauchs eines Servers auf den Arbeitsspeicher. Verfügt dagegen ein Virtualisierungs-Host über 192 GByte an Arbeitsspeicher, fallen fast 60 Prozent des Energieverbrauchs des Servers für diesen Bereich an. Bei den Technologien für die DRAM-Speicherchips kommt derzeit Double Date Rate (DDR) in den Industrie-Standardservern zum Einsatz. Dabei reduziert sich über die verschiedenen DDR-Generationen die nötige Versorgungsspannung für die Chips und damit auch der Energieverbrauch. Für die erste Generation waren noch 2,5 Volt (V) nötig, bei DDR-2 nur mehr 1,8 V, mittlerweile ist DDR-3 im Einsatz mit 1,5 V. Noch im Jahr 2011 wird die Spannung auf 1,35 V sinken. Die



Bei einem Arbeitsspeicherausbau in Höhe von 192 GByte in einem auf Intel-Prozessoren basierten Industrie-Standardserver fallen 58 Prozent der Leistungsaufnahme für den Arbeitsspeicher (DRAM) an (Abb. 1).



Bis zu 86 Prozent weniger Stromaufnahme sind bei einem Server zu verzeichnen, wenn die kommende DRAM-Technologie Verwendung findet – im Vergleich zu den Speicherbausteinen, die noch 2008 State of the Art waren (Abb. 2).

entsprechenden Normierungsarbeiten werden vom JEDEC-Gremium vorgegeben. An diese Standards halten sich dann auch alle Hersteller von Speicherchips.

Verschiedene Parameter helfen, um bei DRAMs Energie zu sparen

Als ein Schritt hin zu weniger Stromverbrauch gilt der Umstieg von DDR-2 auf DDR-3. Dabei hat man bei Samsung die Werte in Serversystemen mit 48 GByte Arbeitsspeicherausbau gemessen. Beide Systeme hatten DRAM-Bausteine auf der Basis der Fertigungstechnologie mit 60 Nanometern (nm) Strukturbreite und Module auf 1-GBit-Basis im Einsatz. Dabei reduziert sich der Energieverbrauch bereits um 35 Prozent, wenn man von DDR-2 auf DDR-3 mit den angegebenen Kenngrößen umsteigt. Daher sollte man bei Neuinvestitionen in seine Serverhardware darauf achten, dass möglichst schon DDR-3-Chips eingesetzt werden können.

Doch die Fortschritte in der Fertigungsprozessstechnologie führen ebenfalls zu einer Reduzierung der Energieaufnahme im Server (für die Berechnung wurden wieder 48 GByte angenommen): Für DDR-3 wurde bei 60 nm Strukturbreite (das war Stand der Technik vor zwei Jahren) eine Leistungsaufnahme in Höhe von 66 Watt gemessen. Sie reduziert sich bei 50 nm bereits um 25 Prozent. Erste Hersteller sind bereits in der Lage, DRAMs in 40-nm-Technologie zu fabrizieren. Der Umstieg auf 40 nm bringt eine weitere Reduzierung um 18 Prozent: Damit fallen nur noch etwa 41 Watt bei einem Speicherausbau von 48 GByte an. Und für den kommenden Umstieg in 2011 auf 30 nm erwartet man eine weitere Reduzierung auf 35 Watt.

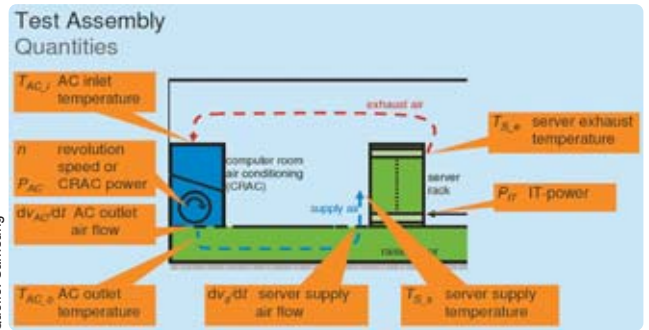
Ein weiterer Parameter, um die Energiekosten zu reduzieren, ist die Speicherdichte auf den Modulen. Der Umstieg von der 1-GBit-Basis auf Module mit 2-GBit-Basis reduziert die Leistungsaufnahme von 41 Watt auf 34 Watt. Der Grund: Es befinden sich weniger Bausteine auf einem Modul. Der Einsatz der nächsten Technik, die 4 GBit basierten Module, lässt dann die Leistungsaufnahme wohl auf 19 Watt zurückgehen (alles Werte bei 48 GByte in einem Server). Als vierte Stellenschraube gilt die Skalierung in Bezug auf die Betriebsspannung für den Arbeitsspeicher. Bei Komponenten auf Basis von 2 GBit, einem Ausbau von 48 GByte und dem Einsatz von DDR-3-Speicher im Server verbraucht das System mit 1,5 Volt Signalspannung 34 Watt. Das reduziert sich bei 1,35 Volt um 18 Prozent auf 28 Watt.

Prozesstechnologie ist entscheidend für den Energieverbrauch eines Servers

Wer Server aus dem Jahr 2008 betreibt, die sozusagen als „State of the Art“ mit DDR-2-Arbeitsspeicher, einer Signalspannung von 1,8 Volt, auf der Basis von 1-GBit-Komponenten und mit einer 60-nm-Fertigungstechnologie, der kann 86 Prozent dieser Leistungsaufnahme einsparen, wenn er im neuen Server Speicherchips mit 30 nm Strukturbreite, mit geringer Signalspannung, und mit Komponenten auf 4-GBit-Basis einsetzt (siehe Abbildung 2, alles gerechnet pro 48 GByte).

Fallstudie Datacenter: Gemessene Werte zum Energiebedarf

Um in einer echten Umgebung belegen zu können, dass die rechnerischen Werte auch zutreffen, wurde eine Testumgebung aufgestellt. Dabei haben zusammen mit Samsung auch andere Unternehmen mitgearbeitet: Microsoft mit seiner Virtualisierungs-Technologie, Fujitsu Technology Solutions mit den Servern; von Samsung stammen die



Quelle: Samsung

Messaufbau für das Test-RZ (Abb. 3)

Set-up:

- **Location and test equipment:**
 - TÜV Rheinland / Köln
- **System:**
 - 2 PRIMERGY TX300 platforms from Fujitsu
- **Processor:**
 - Xeon L5640 Westmere EP from Intel
- **OS:**
 - Server OS, virtual machine app., from Microsoft
- **Memory:**
 - 12 units of 8GB 1Gb based 56nm 1,5V from Samsung,
 - 12 units of 8GB 2Gb, 35nm based 1,35V from Samsung

Quelle: Samsung

Das „Server-Setup“ für die beiden Rechenzentren (Abb. 4)

Speichermodule, und von Intel das Know-how zu den Prozessoren. Die Messungen zum Energieverbrauch hat der TÜV Rheinland ausgeführt – den Messaufbau zeigt Abbildung 3 und in Abbildung 4 ist das „Server-Setup“ für die beiden Rechenzentren zu sehen.

In einem Rechenzentrum (das Data Center A) wurden 25600 Server gerechnet, mit 96 GByte pro Server Arbeitsspeicher und mit den aktuellen Speicherbausteinen (56 nm). Es kamen jeweils 20 virtuelle Maschinen pro Server zum Einsatz. Das andere Rechenzentrum, Datacenter B, mit derselben Anzahl Systeme verwendete die neuesten Arbeitsspeicher (35 nm Fertigungsprozess, 4 GBit Module, 1,35 Volt Spannung).

Das ergab im Bereich des Energieaufwands allein für den Arbeitsspeicher eine Reduzierung um 86 Prozent. Für das gesamte Rechenzentrum gerechnet, betrug die Reduzierung immerhin noch 43 Prozent – die sich natürlich im entsprechenden Verhältnis auf die Stromkosten auswirkt. Die Reduzierung der Stromkosten ergab 12 Millionen Dollar auf 6,8 Millionen Dollar – für diesen extremen Ausbau der Rechenzentren.

Und als positiver Nebeneffekt: Die Performance war beim Data Center B sogar noch besser: Fujitsu Technology Systems hatte in seinen Labors in Paderborn nachgemessen, dass die aktuelleren DRAM-Generation zu einer 40 Prozent höheren Effizienz führen (mit dem SPECpower-Benchmark gemessen). Der Rechendurchsatz lag um 22 Prozent besser.

Rainer Huttenloher

Der Beitrag basiert auf dem Vortrag „Green IT beginnt mit grünen Komponenten“, den Peyman Blumstengel, Head of DRAM Marketing EMEA, bei Samsung Semiconductor Europe gehalten hat.

Aus Heiß mach Kalt – oder besser umgekehrt?

Diskussion über die Vorteile der Warm- oder Kaltgang-Einhausung

Beim Aufbau von Rechenzentren oder Serverräumen scheiden sich oftmals die Geister. Denn für das Abführen der Kühllasten bieten sich prinzipiell zwei Ansätze an. Dabei gilt es für die jeweilige Konzeption zu klären, welches Prinzip besser passt: die Kaltgang- oder die Warmgang-Einhausung. Allerdings stellt eine Einhausung auch zusätzliche Forderungen an den Brandschutz, um eine umfassende Sicherheit für den Rechneraum garantieren zu können.

Wenn es um das Bereitstellen von hochverfügbaren Daten und den Einsatz von energieeffizienten Systemen geht, spielt die Kühlungslösung im Serverraum und Rechenzentrum eine entscheidende Rolle: Sie übt nicht nur maßgeblichen Einfluss auf die Verfügbarkeit geschäftskritischer Anwendungen aus, sondern auch auf den Wirkungsgrad. Die steigenden Leistungsdichten in den IT-Racks – wie sie etwa durch den Einsatz von Blade-Rechnern gegeben sind – treiben den Energieverbrauch in die Höhe und fordern daher zusätzliche Kühlungsleistung. Eine bewährte Methode zur Steigerung der Kühlungseffizienz ist die Trennung von Kalt- und Warmluftströmen zwischen den einzelnen Rack-Reihen. Hierbei kann zwischen zwei Funktionsweisen unterschieden werden.

Beim Warmgangmodell werden die IT-Schränke mit den Rückseiten gegenüber in zwei Reihen aufgestellt. Das IT-Equipment saugt von außen Kaltluft an, die Abwärme sammelt sich zwischen den beiden Schrankreihen und bildet den Warmgang (Hot Aisle). Eine Warmgang-Einhausung (Hot Aisle Containment System) schottet den Warmgang mittels Türen und Dachelementen zum umgebenden Raum hin luftdicht ab und konzentriert so die erwärmte Luft im Inneren. Leistungs-

fähige, in die Schrankreihen integrierte Kühlgeräte und Präzisionsklimaanlagen entziehen dann dem Warmgang die erhitzte Luft, kühlen sie über einen geschlossenen Wasserkreislauf oder Kältekreis ab und leiten sie wieder nach außen in den Raum, aus dem die IT-Geräte dann wieder die Kühlluft ansaugen.

Kaltganglösungen gehen den umgekehrten Weg: Die Vorderseiten der IT-Schränke stehen sich gegenüber und die Abwärme des eingebauten IT-Equipments wird frei in den Serverraum oder das Rechenzentrum abgegeben. Die erwärmte Luft steigt nach oben und sammelt sich als Warmluftpolster unter der Decke. Von dort aus wird sie mittels Raumklimaanlage abgesaugt, gekühlt und anschließend über einen Doppelboden wieder in den Kaltgang zwischen den IT-Schränken eingeblasen. Die Kaltgang-Einhausung vermeidet – analog zum Hot Aisle Containment – die Vermischung von Warm- und Kaltluft. Dadurch erreichen beide Systeme eine effektive Kühlung.

Beim Prinzip der Kaltgang-Einhausungen mit Raumkühlung ist allerdings ein Doppelboden erforderlich. Sonst kann die gekühlte Luft nicht mehr in den abgeschlossenen Kaltgang zwischen den Serverracks gelangen. Würde wie in vielen Fällen bei der Planung kein Doppelboden vorgesehen, muss dieser nachträglich eingebaut werden. Das Manko: Da der Doppelboden eine Höhe von mindestens 60 Zentimetern benötigt, vermindert sich die Deckenhöhe im Rechneraum. Unter Umständen kann das den Warmluftstrom zu den Klimaanlagen behindern.

Warmganglösungen mit Reihenkühlung sind hingegen auch ohne Doppelboden einsetzbar. Damit lassen sich beispielsweise auch Büroflächen für den IT-Betrieb nutzen, die ursprünglich für einen anderen Zweck konzipiert wurden. Beim Prinzip der Warmgang-Einhausung werden die Klimageräte in der Schrankreihe platziert. Speziell für diesen „In-Row“-Einsatzzweck entwickelte Kühlgeräte und Präzisionsklimaanlagen weisen die gleichen Höhen und Tiefen wie 19-Zoll-Schränke auf.

Als weitere Alternative für die Kühlung bei Warmgang-Einhausungen bietet zum Beispiel die InRow OA von APC by Schneider Electric den Vorteil, dass sie keinen weiteren Platz im Rechneraum einnimmt. Die Montage erfolgt an der Raumdecke oder auf den Dächern der sich gegenüberstehenden 19-Zoll-Schränke. Durch die In-Row-Bauweise wird eine sehr hohe Leistungsfähigkeit bis über 30 Kilowatt pro 19-Zoll-Schrank erreicht – vorausgesetzt, die gegenüberstehenden Schrankreihen werden symmetrisch aufgebaut.



Quelle: APC by
Schneider Electric

Eine Warmgang-Einhausung (Hot Aisle Containment System) schottet den Warmgang mittels Türen und Dachelementen zum umgebenden Raum hin luftdicht ab (Abb. 1).

Die Redundanzbildung bei den Kühlgeräten ist einfach zu erreichen und bietet damit große Vorteile gegenüber in sich geschlossenen, gekühlten 19-Zoll-Schränken. Idealerweise sollte zudem Wasser- beziehungsweise ein Wasser-Glykol-Gemisch den Wärmeabtransport übernehmen, da erst so die indirekte freie Kühlung möglich ist. Gegenüber einer offenen Aufstellung der Schrankreihen lässt sich so eine Energieeffizienzsteigerung von 20 bis 50 Prozent erzielen. Im Vergleich zu sehr einfach ausgestatteten und mangelhaften Installationen mit Kaltluftzufuhr durch den Doppelboden können sogar bis zu 80 Prozent der Klimakosten eingespart werden.

Modular aufgebaute Warmganglösungen bieten weiterhin den Vorteil der vergleichsweise einfachen Skalierbarkeit. So lassen sich die Reihen durch zusätzliche Schränke und Kühlgeräte bedarfsgerecht erweitern. USV und Stromverteilung sind beim Infrastruxure-Modell beispielsweise ebenfalls modular aufgebaut. Das erlaubt es dem Anwender, im unteren Leistungsbereich zu starten und sein System später zu erweitern und durch integrierte Reihenkühlung und ein „Hot Aisle Containment“-System zur High-Density-Lösung zu erweitern.

Packungsdichten und Energieeffizienz des Rechenzentrums lassen sich somit zu verhältnismäßig geringen Investitions- und Betriebskosten erhöhen. Des Weiteren kann man bestehende Geräte zur Raumklimatisierung dabei als Ergänzung oder Redundanz weiter verwenden.

Bei der Realisierung eines Warmgangkonzepts sollten IT-Manager auch auf das Brandschutzkonzept achten, denn eine Lösung mit einem Sprinklersystem, das Wasser von der Raumdecke her zuführt, scheidet wegen der Abdeckung durch die Einhausung üblicherweise aus. Um mögliche Brände in den IT-Geräten, den USV- oder Klimaanlage schon weit vor dem Auftreten von offenen Flammen zu vermeiden, empfiehlt es sich sowohl im Raum als auch in der Einhausung ein Rauchsaugsystem (RAS) mit Detektionsrohren zu installieren.

Da eine raumbasierende Luftführung verwendet wird, sollte das IT-Management auch auf ein raumbasierendes Löschsystem setzen. Möglichkeiten hierfür sind die Flutung von großen Räumen mit Inertgasen wie Argon oder Stickstoff und bei kleineren Räumen mit künstlichen Löschgasen wie FM200 oder NOVEC 1230. Als Vorteil bei diesen Konzepten ergibt sich ein wichtiger Punkt: Die IT-Systeme und Klimageräte müssen nicht abgeschaltet werden.

Häufig wird auch die Inertisierung für den Brandschutz diskutiert, also die Reduzierung des Sauerstoffgehalts der Luft im Rechneraum. Inertisierung ist auch beim Einsatz von Einhausungen möglich. Wichtig sind hier die Dichtigkeit des Raumes und eine geringe Zugangsfrequenz, um den Sauerstoffeintrag bei geöffneter Tür so gering wie möglich zu halten. Am besten eignen sich dafür Türen mit Schleusenfunktion (siehe Kasten „Dem Feuer die Luft abdrehen“).

Kaltgang-Einhausungen können eine Lösung bei geringer Leistungsdichte sein. Speziell in existierenden Rechnerräumen mit Doppelboden und vorhandenen Downflow-Klimaanlagen können diese Komponenten weiterverwendet werden. Der Aufbau einer Kaltgang-Einhausung in Verbindung mit offen aufgestellten, nicht eingehausten Rack-Reihen sollte allerdings vermieden werden.

Warmgang-Einhausungen mit In-Row-Kühlung bieten sich für alle Anwendungen im höheren Leistungsbereich an und lassen sich auch problemlos in große Rechenzentren mit offen aufgestellten Rack-Reihen als sogenannte Hybridlösungen integrieren. Kalt- und Warmgang-Einhausungen im selben Raum sind dagegen kontraproduktiv und sollten keinesfalls verwendet werden. Es empfiehlt sich daher, entweder die kalte oder die warme Luft einzuhausen.

*Michael Schumacher
ist Senior Systems Engineer bei APC by Schneider Electric.*

DEM FEUER DIE LUFT ABDREHEN



Quelle: APC by
Schneider Electric

Bei einer Warmgang-Einhausung ist das Thema Brandschutz wichtig, weil Sprinkleranlagen von oben normalerweise keine Löschung erlauben.

Eine Alternative zur Branderkennung und Brandbekämpfung sind die Brandvermeidungssysteme. Mithilfe einer genauen Steuerung wird in einem abgeschlossenen Bereich die Sauerstoffkonzentration soweit reduziert, dass keine Brände entstehen können. Dazu wird in die Luft zusätzlich Stickstoff eingebracht und so die Sauerstoffkonzentration reduziert. Bei einer Konzentration von weniger als 16 Volumenprozent Sauerstoff ist ein selbstständiges Brennen nicht mehr möglich. Die Regelung stellt sicher, dass dieses reduzierte Sauerstoffniveau auch beibehalten wird. Die Sauerstoff-Reduzierung liegt dabei in einem Bereich, sodass die Räumlichkeiten nach wie vor durch Menschen begehbar bleiben.

Für die Dimensionierung der Brandvermeidungssysteme sind verschiedene Einflussgrößen bestimmend. Als besonders wichtig gilt die Dichtheit der Raumhülle und das zu schützende Rauminvolumen – es muss ja die Sauerstoffkonzentration unter einem bestimmten Niveau gehalten werden. Vor diesem Hintergrund ist zum Beispiel die Zugangsfrequenz einzurechnen, also wie häufig Personen in den Raum gehen und dazu Türen öffnen müssen, durch die dann wieder normale Luft mit der höheren Sauerstoffkonzentration in den Schutzbereich gelangt. Auch der Faktor Klimatisierung geht in die Gleichung ein. Hier wird kühle Luft – ebenfalls mit höherer Sauerstoffkonzentration zugeführt.

Die Auswirkungen auf den Menschen, der in Schutzräumen mit geringerer Sauerstoffkonzentration arbeiten muss, ist vergleichbar mit dem Aufenthalt im Gebirge. Hier liegt die Sauerstoffkonzentration in einer Höhe von 2250 Metern bei etwa 16 Volumenprozent. Die gewerblichen Berufsgenossenschaften haben im Arbeitskreis Feuerschutz eigene Definitionen aufgestellt, um eine Gefährdung von Personen zu vermeiden. Demnach besteht bei einer Sauerstoffkonzentration von 20,9 bis 17 Volumenprozent eine freie Begehbarkeit für alle Personen ohne bekannte Herz-, Kreislauf-, Gefäß- und Atemwegs-Erkrankungen. Bei Sauerstoffkonzentrationen von 17 bis 13 Volumenprozent können Personen, die eine arbeitsmedizinische Vorsorgeuntersuchung genossen haben, je nach Restsauerstoffgehalt zwischen zwei und vier Stunden ohne Unterbrechung in diesem Bereich arbeiten.

Rainer Huttenloher

Freiluftkühlung garantiert effizienten Betrieb

Kundenspezifische Lösung verspricht effiziente Kühlung in Rekordzeit

Der IT-Dienstleister für die deutsche regionale Tageszeitung „Donaukurier“, die Aligia, hat am Standort Ingolstadt ein Rechenzentrum errichtet, um externen Kunden Serverhousing- und Storage-Kapazitäten anbieten zu können. Das Systemhaus Cancom realisierte innerhalb von zwei Monaten eine energieeffiziente Infrastruktur und eine passende Klimatisierungslösung. Dabei setzte Cancom Komponenten von Schäfer IT-Systems ein, die das Kühlkonzept unterstützen.

Kunden, die bei uns ihre Hardware einstellen, sollen sich mit ganzer Kraft um ihr Kerngeschäft kümmern können, ohne sich Sorgen um die Sicherheit ihrer Server und die Energiekosten im Rechenzentrum machen zu müssen“, erläutert Geschäftsführer Wolfgang Stiegler das Selbstverständnis der Aligia. „Neben einer qualitativ hochwertigen Infrastruktur und absoluter Termintreue spielte deshalb das Thema Energieeffizienz eine zentrale Rolle bei der Errichtung des neuen Rechenzentrums.“

Für maximale Energieeffizienz ist eine Infrastruktur von Alcatel-Lucent im Einsatz, die dank ihres geringen Leistungsverbrauchs bereits Preise für „Green IT“ erhielt. Geeichte Stromzähler sichern nicht nur die exakte Verbrauchskontrolle, sondern auch die größtmögliche Preistransparenz für die Aligia-Kunden. 22 000 Meter von Kategorie-7-Datenkabel sorgen für die nötigen Querverbindungen innerhalb des Rechenzentrums. Für die Umsetzung des Klimakonzepts verwendete Cancom vier Einhausungen des Typs Cold Section der Firma Schäfer IT-Systems, die auch 40 Serverschränke sowie fünf SP-Rack-Netzwerkschränke lieferte.

„In Sachen Kosten haben wir sehr vom Pricing der Schäfer-Produkte profitiert. Das konnten wir an die Aligia weitergeben“, erklärt Cancom-Geschäftsführer Christian Steininger. Und das hat sich gelohnt, wie Wolfgang Stiegler bestätigt: „Dass wir uns unter fünf Anbietern für Cancom entschieden haben, lag am überzeugenden Lösungskonzept und am guten Preis-Leistungs-Verhältnis“, erläutert der Aligia-Geschäftsführer.

Klimaprojekt in drei Monaten realisiert

Die Errichtung des gesamten Rechenzentrums dauerte von der Beauftragung bis zur Abnahme drei Monate. Die Konzeption und Umsetzung der Klimatisierungslösung gelang Cancom in nur fünf Wochen. Wie es dazu kam, erklärt Stiegler so: „Das Rechenzentrum der Aligia befindet sich im Keller des Donaukurier-Verlagsgebäudes. Deshalb bot es sich an, es über die Klimaanlage des Hauses zu kühlen.“ Entsprechend wurde Cancom im April 2010 zunächst nur mit der Infrastruktur des Rechenzentrums – ohne Kühlung – beauftragt.



Die Einhausungslösung Cold Section ist ein integraler Bestandteil des Klimatisierungskonzepts im Aligia-Rechenzentrum (Abb. 1).



Das modulare Doppelbodensystem – hier Rahmen mit je neun Lüftungskacheln – erzielt einen freien Querschnitt von bis zu 60 Prozent (Abb. 2).

Alle Bilder: Schäfer IT-Systems

Das Rechenzentrum auf 24 Grad Celsius zu kühlen, was eine Kältebelastung von immerhin 120 Kilowatt mit sich bringt, erwies sich im Rahmen der Gebäudeklimatisierung jedoch als unwirtschaftlich. Darum bat Aligia das Systemhaus zusätzlich um ein Klimatisierungskonzept. Innerhalb von zwei Arbeitstagen lag dieses samt Hydraulikschema auf dem Tisch. „So konnten wir Cancom Anfang Juni zusätzlich mit der Einrichtung der Klimasysteme beauftragen.“

Bereits Ende Juli erfolgte die Abnahme des fertigen Rechenzentrums inklusive Klimatisierung. Andere Anbieter hätten laut Stiegler frühestens Ende August liefern können, und die Montage hätte weitere Zeit gekostet.

Drei Kaltwassersätze bieten Redundanz

Die erforderliche aktive Kühlung ist durch ein redundantes System von drei Kaltwassersätzen auf dem Dach des Verlagsgebäudes mit insgesamt 180 Kilowatt Bruttoleistung sichergestellt. Entscheidend für die Energieeffizienz des gesamten Systems ist jedoch das Zusammenspiel mit der Freiluftkühlung, die bis zu 400 Kilowatt Kühlung erzielt. „Unsere SP Racks und Einhausungen unterstützen eine effiziente Freiluftkühlung, weil ihre Konstruktion für eine klare Trennung von Kalt- und Warmgängen sorgt“, erläutert Peter Wäsch, Leiter Verkauf IT DACH bei Schäfer IT-Systems.

Für das Klimatisierungskonzept erwies sich das Schäfer-Doppelbodensystem Swap Panel 9 als entscheidende Komponente. Das modulare System besteht aus 600×600 Millimeter großen Rahmengeräten, in die neun Kacheln – geschlossen, gelocht oder mit Kabeldurchlass versehen – beliebig eingebaut werden können. Beim Ausbau mit neun Lüftungskacheln erzielt das System einen freien Querschnitt von 60 Prozent.

„Wir hatten ursprünglich einen konventionellen Doppelboden vorgesehen“, erklärt Steinger. „Die angestrebte Kühlung von drei Kilowatt pro Rack hätten wir damit gerade eben erreicht. Durch Einsatz der Swap Panels konnten wir die Kühlleistung pro Rack um weitere 50 Prozent erhöhen, zumal Schäfer auf die Modifikation des Lieferumfangs schnell reagiert hat und die Komponenten sofort verfügbar waren.“

Der Clou des Systems: Die Spitzenkühlleistung stellt nicht nur die Kühlung der Server sicher. Über eine Trennstation wird die übrige Käl-



Das Zusammenspiel aus aktiver Kühlung (Bild), Freikühlung und Einhausung sorgt für eine hohe Energieeffizienz in einem Rechenzentrum (Abb. 3).

IT-ABTEILUNG WIRD ZUM EXTERNEN DIENSTLEISTER

Die Aligia GmbH in Ingolstadt bietet Serverhousing- und Storage-Kapazitäten sowie flankierende Beratungsleistungen speziell für den regionalen Markt an. Das Unternehmen ist der IT-Dienstleister der Donaukurier Verlagsgesellschaft mbH & Co. KG, einem der führenden Medienunternehmen im nördlichen Oberbayern. Die IT-Abteilung des „Donaukurier“, aus der die Aligia hervorgegangen ist, verfügt über langjährige Erfahrung in der Versorgung von derzeit rund 550 Mitarbeitern am Standort Ingolstadt. Diese Service- und Infrastrukturkompetenz steht durch die Aligia nun auch externen Kunden zur Verfügung.

leistung in die Klimaanlage des Verlagshauses eingespeist – und senkt dort den Stromverbrauch der Klimaanlage. So profitiert die Energiebilanz des gesamten Gebäudes von der Energieeffizienz des Rechenzentrums. „Bei Einsatz dieser Freiluftkühlung ist eine Einsparung von 40 000 Euro pro Jahr an Stromkosten – bei einem angenommenen Strompreis von 0,11 Euro – gegenüber einer herkömmlichen Klimaanlage zu erwarten“, hat Christian Steinger errechnet.

Perspektiven durch Leistungsreserven

Die in Ingolstadt ansässige deutsche Niederlassung des Softwarehauses Kaspersky Lab, einer der ersten Mieter, nutzt bereits ein Drittel der Housing-Kapazitäten von Aligia. Eine baldige Erweiterung des Rechenzentrums liegt damit durchaus im Bereich des Möglichen. Stiegler sieht sein Unternehmen für diesen Fall bestens gerüstet: „Die Leistungsreserven der von Cancom realisierten Klimatisierung und die modulare Erweiterbarkeit der Schäfer-Systeme geben uns enorme Sicherheit. In diesen Partnern sehen wir die Garantie, dass wir notwendige Erweiterungen unseres Rechenzentrums schnell und auf hohem technischem Niveau durchführen können.“

*Gerald Fiebig
ist in der Konzept PR als Redakteur tätig.*

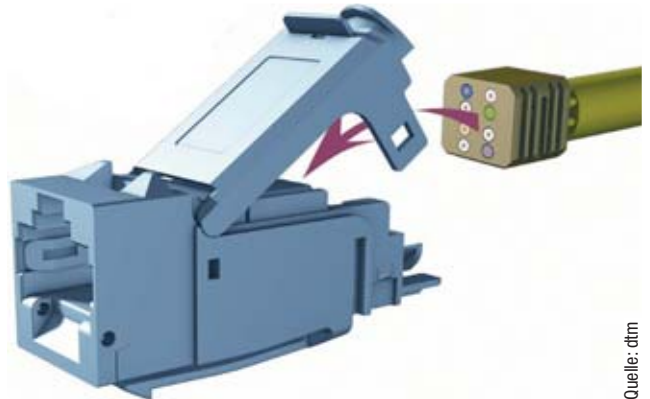


Hinter einer mit Fingerprint-Scanner gesicherten Tür befindet sich das Aligia-Rechenzentrum im Keller des Donaukurier-Verlagsgebäudes (Abb. 4).

Rechenzentrum und Infrastruktur – Komponenten, Kabel, Netzwerke

Die nächste Verlagsbeilage erscheint mit der *iX*-Ausgabe 5/2011 am 28. April 2011.
Dabei ist unter anderem das folgende Thema geplant:

LEDs im RJ45-Anschlussystem bringen den Überblick: Anwendungsneutrale Übertragungsstrecken errichten und dabei das Steckgesicht zum Anwender auch nachträglich ohne weiteren Installationsaufwand bestimmen – das sind Vorteile des Anschlussystems Quicklink. Es eignet sich für Kupferkabel der Kategorien 6 und 6A. Das System ist geeignet zur Übertragung von analogen und digitalen Sprach-, Bild- und Datensignalen. Die Leistungsmerkmale entsprechen der Klasse EA bis 500 MHz gemäß EN 50173-1 und 10GBase-T bis 500 MHz. Auch die Eignung für Power over Ethernet (PoE) gemäß IEEE 802.3at ist gegeben.



Schematische Darstellung
des Einführens eines Kupferkabels in das
Quicklink-Keystone-Modul

Quelle: dftm

Impressum Themenbeilage Rechenzentren & Infrastruktur

Redaktionsbüro Huttenloher

Telefon: 088 56/99 75, Fax: 088 56/99 76, E-Mail: rhu@heise.de

Verantwortlicher Redakteur:
Rainer Huttenloher (088 56/99 75)

Autoren dieser Ausgabe:
Gerald Fiebig, Carrie Higbie, Rainer Huttenloher, Michael Schumacher

DTP-Produktion:
Enrico Eisert, Wiebke Preuß, Matthias Timm, Hinstorff Verlag, Rostock

Korrektur:
Wiebke Preuß

Technische Beratung:
Duc-Thanh Bui

Verlag
Heise Zeitschriften Verlag GmbH & Co. KG, Postfach 61 04 07,
30604 Hannover; Helstorfer Straße 7, 30625 Hannover;
Telefon: 05 11/53 52-0, Telefax: 05 11/53 52-129

Geschäftsführer:
Ansgar Heise, Steven P. Steinkraus, Dr. Alfons Schröder

Mitglied der Geschäftsleitung:
Beate Gerold

Verlagsleiter:
Dr. Alfons Schröder

Anzeigenleitung (verantwortlich für den Anzeigenteil):

Michael Hanke (-167), E-Mail: michael.hanke@heise.de

Stellv. Anzeigenleiter und Ltg. International:

Oliver Kühn -395, E-Mail: oliver.kuehn@heise.de

Assistenz:

Stefanie Frank -205, E-Mail: stefanie.frank@heise.de

Anzeigendisposition und Betreuung Sonderprojekte:

Christine Richter -534, E-Mail: christine.richter@heise.de

Anzeigenverkauf:

PLZ-Gebiete 0–3, Ausland: Oliver Kühn -395, E-Mail: oliver.kuehn@heise.de,
PLZ-Gebiete 8–9: Ralf Räuber -218, E-Mail: ralf.raeuber@heise.de
Sonderprojekte: Isabelle Paeseler -205, E-Mail: isabelle.paeseler@heise.de

Anzeigen-Inlandsvertretung:

PLZ-Gebiete 4–7: Karl-Heinz Kremer GmbH, Sonnenstraße 2,
D-66957 Hilst, Telefon: 063 35/92 17-0, Fax: 063 35/92 17-22,
E-Mail: karlheinz.kremer@heise.de

Teamleitung Herstellung:

Bianca Nagel

Druck:

Dierichs Druck + Media GmbH & Co. KG, Kassel

Eine Haftung für die Richtigkeit der Veröffentlichungen kann trotz sorgfältiger Prüfung durch die Redaktion vom Herausgeber nicht übernommen werden. Kein Teil dieser Publikation darf ohne ausdrückliche schriftliche Genehmigung des Verlages verbreitet werden; das schließt ausdrücklich auch die Veröffentlichung auf Websites ein.

Printed in Germany

© Copyright 2011 by Heise Zeitschriften Verlag GmbH & Co. KG

Die Inserenten

Die hier abgedruckten Seitenzahlen sind nicht verbindlich.
Redaktionelle Gründe können Änderungen erforderlich machen.

Avocent www.avocent.de S. 17
Dell www.dell.de S. 11

IP Exchange www.ip-exchange.de S. 28
noris network www.datacenter.de S. 5
Rittal www.rittal.de S. 14/15
Samsung www.samsung.com/greenmemory S. 2
Thomas Krenn www.thomas-krenn.de S. 27

EXZELLENT SERVERHARDWARE

ZU UNSCHLAGBAREN PREISEN



1HE Intel Single-CPU SC813M Server

- Platzsparend durch 503 mm Einbautiefe
- Neueste Intel Nehalem - 2.93GHz und 4 Kerne pro CPU
- Bis zu 8 TB Festplattenspeicher und 16 GB Speichermodule

STATT 649,-
589,- EUR



2HE Intel Single-CPU SC825 Server

- Neueste Intel Nehalem - 2.93GHz und 4 Kerne pro CPU
- Bis zu 16 TB Festplattenspeicher und 16 GB Speichermodule
- Hohe Ausfallsicherheit durch redundantes Hot-Swap-Netzteil

STATT 1.059,-
899,- EUR



3HE INTEL DUAL-CPU SC836 SERVER

- Neueste Intel Westmere - 3.33GHz und 6 Kerne pro CPU
- Bis zu 32 TB Festplattenspeicher und 192 GB Speichermodule
- Optional erweiterbar auf 4 Onboard LAN-Ports
- Integriertes IPMI on Board

STATT 2.099,-
1.999,- EUR



01.03. - 05.03.2011
HALLE 2 STAND B46 HANNOVER
www.thomas-krenn.com/cebit_2011

Infos unter:
www.thomas-krenn.com/server_hardware

Thomas-Krenn.AG[®]
Die Server-Experten



Flächendeckendes Händler- und Servicenetz
in der Schweiz: www.thomas-krenn.com/ch

Business Class Rechenzentren

Outsourcing für Serverstrukturen, von individuellen Stellflächen bis full managed Hosting



Outsourcing

Housing

Hosting

Management

Netzbetrieb

Consulting

24/7 Service

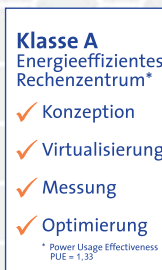
Datentransport

global CDN

Streaming

Archivierung

Virtualisierung



*Sprechen Sie mit uns über Ihre Anforderungen
für den **sicheren** Betrieb Ihrer IT Systeme in
erstklassigen deutschen Serverstandorten*

IP Exchange ist einer der führenden Anbieter sicherer Rechenzentrumsflächen für Outsourcing in Deutschland. Wir sind darauf spezialisiert den höchsten Standard physischer Sicherheit und betrieblicher Stabilität zu gewährleisten. Ein modulares Produktspektrum passt sich an die spezifischen Kundenwünsche an, bietet dauerhaft Flexibilität und wirtschaftlich effiziente Rahmenparameter. Von reiner IT Fläche bis hin zu full managed IT Ressourcen werden in den Standorten Nürnberg und München ausschließlich B2B Kunden auf einer Nutzfläche von über 10.000qm mit Dienstleistungen und 24/7 Service betreut. Des Weiteren berät, plant, baut und betreibt IP Exchange ganze dedizierte Rechenzentren und Serverräume im Kundenauftrag an strategischen Standorten in ganz Europa – oder direkt im jeweiligen Unternehmen. Gerne beraten wir Sie unverbindlich über Kosten- & Aufwandsminimierung bei gleichzeitiger Effizienzsteigerung Ihrer Serverstrukturen bei zukünftigen Migrationsvorhaben und Budgetgestaltungen.

