Eine Themenbeilage der
Heise Medien GmbH & Co. KG

IV
2016

RECHENZENTREN UND INFRASTRUKTUR

KOMPONENTEN, KABEL, NETZWERKE

Was Zertifikate von Uptime und TÜV IT wert sind

powered by:

Was TÜV- und Uptime-Zertifikate wert sind

Von Snowden haben deutsche Cloud-Anbieter nur profitiert. Ähnlich dürften sich die dienstbeflissenen Mail-Filter von Yahoo auswirken. Microsoft hat sein europäisches Cloud-Geschäft bereits nach Frankfurt und Magdeburg umgezogen, worauf Telekom-Geschäftsführer Dr. Ferri Abolhassan nicht wenig stolz ist. „Sicherheit made in Germany“, sagt er, „entwickelt sich zum Gütesiegel“ (Seite 4). Und Gütesiegel werden für das RZ-Geschäft demnächst kritisch; sowohl im IT-Sicherheitsgesetz als auch in der neuen europäischen Datenschutz-Grundverordnung sind Zertifizierungen ausdrücklich vorgesehen. Bislang kommen die Prüfer meist von TÜV IT, doch das könnte sich ändern, wenn The Uptime Institute auf den Europamarkt startet. Dabei spielt auch die Wahl von London als Sprungbrett eine entscheidende Rolle. Ariane Rüder hat bei den Anbietern nachgehakt und erklärt die Unterschiede der Prüfkriterien und -konzepte ab Seite 12: „Bitte eine Build-Prüfung!“

Die wichtigsten Neuerungen, die sich aus der EU-DSGV ergeben, hat sie außerdem ab Seite 22 zusammengestellt. Für Anbieter, die bereits dem Bundesdatenschutzgesetz unterliegen, wird sich zwar auch ab dem 25. Mai 2018 nicht viel ändern, aber einiges doch. Unklar ist vor allem die Situation mit Großbritannien, da der Brexit erst nach Wirksamkeit der EU-DSGV vollzogen wird. Auch sonst wird der internationale Handel mit Daten inklusive Analytics ein paar Schrauben kräftig anziehen müssen, denn die Verordnung verlangt künftig gewissermaßen einen lückenlosen Herkunftsnachweis inklusive Einwilligung der Betroffenen, ähnlich wie im Lebensmittelrecht. Wie das rechtskonform in der Praxis umzusetzen ist, wird sich vermutlich erst nach ein paar Gerichtsurteilen sagen lassen. Die Landesdatenschutzbeauftragten haben bereits bewiesen, dass sie auch gegen höhere Gewichtsklassen klagelustig sind. Wer in dieser Situation ein Information Security Management System erwägt, findet grundlegende Auskünfte zur Einführung ab Seite 7. Zuvor erklärt Patrick Schraut außerdem die Varianten eines Security Operation Centers (Seite 6).

Damit haben die Anwälte in diesem Heft genug Lektüre bekommen. Wenden wir uns lieber der Technik zu. Ein Schwerpunkt ist hier die Infrastruktur von Cloud-Rechenzentren. Für die Planer bedeutet das immer eine gute Strecke Blindflug. Flexibel skalierbare Ressourcen erfordern im RZ einen hohen Grad an Virtualisierung und starke Layer-2-Umgebungen. Neben TRILL und Fabrics kann man ein hochskalierbares Design auch mit MLAGs und Stacks aus mehreren ToR-Switches bauen. Dass so etwas gut funktioniert, zeigt Olaf Hagemann an einem Realbeispiel (Seite 16). Handfest argumentiert auch Gordon Haff: Hinter der Cloud, sagt er, „steckt immer reale Hardware, die an physikalische Gesetze gebunden ist.“ Sein Beitrag auf Seite 18 beschäftigt sich konkret mit Latenzen.

Die Performance der passiven Netzkomponenten hat unterdessen André Engel untersucht (Seite 20), der gleichfalls auf ein Praxisbeispiel verweisen kann: Der Max-Planck-Campus Tübingen hat sein Data Center mit einem modularen, extrem packungseffizienten Verkabelungssystem umgebaut, das obendrein die Integration von Glasfaser- und Kupferkabeln vereinfacht. Wenn das Ganze dann auch noch in sauber standardisierten Racks steckt (Seite 10), sparen sich Admins eine Menge Ärger bei Wartung und Erweiterung. Andererseits können sie das Infrastrukturmanagement auch gleich auslagern; was es mit Data Center as a Service auf sich hat, erklärt Frank Neubauer auf Seite 19.

Oder es wird alles ganz anders. In der Data-Center-2025-Studie rechnet die Mehrheit der Fachleute damit, dass Infrastruktur-Equipment und IT-Ausstattung deutlich an Effizienz zulegen. Dann müssen die Kühlkonzepte noch näher an die Komponenten rücken. Nicht nur die großen Hyperscale-Rechenzentren dürften von Anfang an eigene Anlagen zur Energieerzeugung mit einplanen. Und vermutlich bequeme Unterkünfte, damit die Reisegruppe der Zertifizierer dort ein paar Wochen Station machen kann.

Thomas Jannot

Die Security-Krauts

Sicherheit made in Germany entwickelt sich zum Gütesiegel

Obwohl die Cyberkriminalität stetig zunimmt, vernachlässigen viele Unternehmen immer noch ihre IT-Sicherheit. Dabei setzt Deutschland derzeit Maßstäbe beim Security-Know-how und bietet die besten Rahmenbedingungen für den Schutz sensibler Daten. Entsprechend gefragt sind deutsche Rechenzentren.

Vor vier Jahren musste ein Unternehmen aus dem Silicon Valley eingestehen, dass ihm 6,5 Millionen Nutzerpasswörter gestohlen worden waren. Nun stellt sich heraus, dass das Ausmaß des Datenlecks tatsächlich um ein Vielfaches größer ist: 117 Millionen Passwörter hat ein Unbekannter unter dem Pseudonym „Peace“ im Darknet angeboten – eine ungeheure Menge persönlicher Daten, die im Internet-Untergrund kursieren. Das gehackte Unternehmen hat die Echtheit der Daten bereits bestätigt.

Und die Auswirkungen des Datenlecks gehen weit über das eine Portal hinaus. Denn nicht selten verwenden Internetnutzer ein und dasselbe Passwort für die Accounts mehrerer Netzwerke – sei es Facebook, Google+ oder Twitter. Darüber hinaus nutzen Hacker echte Passwörter, um damit Cracker-Programme und deren Algorithmen zu füttern, die sich daraufhin noch besser auf das Knacken von Codes trainieren lassen.

Doch wie können solche wichtigen Daten – und noch dazu in einer derartigen Größenordnung – verloren gehen? IT-Fachleute werfen dem kalifornischen Unternehmen vor, dass die Sicherheitsvorkehrungen des Portals nicht einmal elementaren Standards genügten. Demnach waren die Kundendaten zwar verschlüsselt, jedoch nur mit einem veralteten Algorithmus der Variante SHA-1 – und ohne Salt, also eine zusätzlich angehängte Zeichenfolge nach dem Zufallsprinzip. Für Angreifer waren die Passwörter somit eine leichte Beute.

Das Cloud-Geschäft in Europa

IT-Sicherheit wird oft noch stiefmütterlich behandelt – obwohl die Gefahren allgegenwärtig sind. Zu diesem Schluss kommt auch der Cyber Security Report der Deutschen Telekom: Zwar gerät bereits jedes dritte Unternehmen mehrmals pro Woche ins Visier von Cyberkriminellen, doch spezielle Sicherheitsvorkehrungen hat erst die Hälfte der Firmen getroffen. Für den Report führten das Institut für Demoskopie Allensbach (IfD) und das Centrum für Strategie und Höhere Führung im Sommer 2015 Telefoninterviews mit insgesamt 645 Top-Entscheidern aus Politik und Wirtschaft, darunter 247 Führungskräften aus großen und 285 Führungskräften aus mittleren Unternehmen.

Die Lage ist umso merkwürdiger, als es gerade hierzulande riesiges Know-how und bestmögliche Rahmenbedingungen gibt, um Geschäftsdaten zuverlässig vor unbefugtem Zugriff zu schützen – zum Beispiel in einer deutschen Cloud-Umgebung. Global Player, die die Themen IT-Sicherheit und Datenschutz ernst nehmen, haben das längst erkannt. „Security made in Germany“ entwickelt sich seit dem NSA-Skandal und dem Safe-Harbor-Urteil mehr und mehr zum international anerkannten Gütesiegel. Besonders für US-IT-Firmen ist eine sichere Aufbewahrung der Kundendaten inzwischen ein unerlässlicher Erfolgsfaktor für das Cloud-Geschäft in Europa.

Die Nachfrage nach deutschen Cloud-Lösungen steigt kontinuierlich. Eine ganze Reihe namhafter US-Firmen setzt inzwischen auf Rechenzentren in Deutschland. Besonders konsequent geht der größte Softwarehersteller der Welt vor: Im November 2015 kündigte Microsoft ein neues Cloud-Angebot an – mit einem deutschen Datentreuhänder, welcher der deutschen Rechtsordnung unterliegt. Seit September 2016 bringt Microsoft die Dienste Azure, Office 365 und Dynamics CRM Online mit diesem Modell schrittweise auf den Markt.

Als Datentreuhänder agiert T-Systems, eine Tochtergesellschaft der Deutschen Telekom. Sie prüft Zugriffsanfragen auf ihre Rechtmäßigkeit und lehnt unrechtmäßige Zugriffsanfragen ab. Dies bedeutet unter anderem, dass T-Systems selbst Microsoft nur in vertraglich vorab festgelegten Fällen Zugriff auf die Kundendaten gewährt. Diesen Zugriff überwacht, protokolliert und beendet der Datentreuhänder.

Rechenzentren setzen Standards

Um größtmöglichen Schutz der Kundendaten zu gewährleisten, werden die Standorte der hochsicheren deutschen Rechenzentren sorgfältig ausgewählt: keine Ansiedlungen, Autobahnen oder Flugschneisen in unmittelbarer Nähe, die die Sicherheit der Anlagen und Verfügbarkeit der Daten beeinträchtigen könnten. Zudem sorgen physische Sicherheitsmaßnahmen dafür, dass nur hineinkommt, wer hineingehört. Dazu zählen meterhohe Erdwälle, Sicherheitszäune mit Stacheldraht, Videokameras und Bewegungsmelder, Multifaktor-Authentifizierungen sowie Handflächenscanner. Das T-Systems-Rechenzentrum in Biere erfüllt so die Sicherheits-, Service- und Qualitätsstandards als zertifiziertes Information Security Management System (ISMS) nach ISO 27001, CSA STAR Level 2 und TÜV Trusted Cloud Service.

Doppelt ausgelegte Stromversorgungen und Twin-Core-Technologien heben die Ausfallsicherheit auf ein Maximum. Dazu trägt auch ein umfassendes Qualitätsprogramm bei. Mit klaren Standards entlang der drei Dimensionen Personal, Prozesse und Plattformen erhöht ein solches Programm die Datenverfügbarkeit deutlich. Und wenn es doch einmal zu einer Störung kommt – 100 % Ausfallsicherheit gibt es nie –, sorgt idealerweise ein ein Manager on Duty für die schnelle Wiederherstellung der Services. T-Systems hat die Verfügbarkeit seiner Dienste so auf bis zu 99,999 % gesteigert und die Zahl der Major Incidents um 95 % reduziert. Als erstes Qualitätsprogramm dieser Art wurde Zero Outage daher 2015 vom TÜV Rheinland zertifiziert.

Die digitale Sicherheitstechnik der deutschen Data Center ist darauf ausgerichtet, Hacker-Angriffe und Cyberattacken zuverlässig abzuwehren. Daher fließen die Daten oft in einem geschlossenen System durch gesicherte IP-VPN-Tunnel, abgeschottet von den öffentlichen Netzen. Die Firewalls der Rechenzentren werden ergänzt durch Sicherheitsmechanismen wie Intrusion-Detection- und Intrusion-Prevention-Systeme,

Im Herbst 2016 erscheint bei Springer-Gabler das von Dr. Ferri Abolhassan herausgegebene Buch „Security Einfach Machen – IT-Sicherheit als Sprungbrett für die Digitalisierung“.

Quelle: T-Systems/Marc-Steffen Unger

die Datenströme auf verdächtigen Traffic und Schadcode filtern sowie bedenklichen Zugriffsmustern nachspüren. Moderne Datenverschlüsselungen nach SSL-/TLS-Protokollen sorgen dafür, dass die Daten nur von berechtigten Personen eingesehen werden können.

Provider mit D-Zentrale bevorzugt

Mit seinem Angebot für den deutschen Markt ist Microsoft im Cloud-Zeitalter nicht allein: Rund drei Viertel der Unternehmen legen größten Wert darauf, dass ihr Provider seinen Hauptsitz in Deutschland hat (72 %) und auch ausschließlich hier seine Rechenzentren betreibt (76 %). Zu diesem Ergebnis kommt der Bitkom in seinem Cloud-Monitor 2016. „Der Standort Deutschland genießt einen Vertrauensvorsprung“, heißt es darin.

Neben der hochsicheren Technik beruht das Vertrauen in deutsche Cloud-Anbieter vor allem darauf, dass diese dem Bundesdatenschutzgesetz unterliegen, einem der strengsten Gesetze seiner Art weltweit. An dem Vertrauen in deutsche Cloud-Anbieter wird auch die EU-Datenschutz-Grundverordnung (EU-DSGV) nichts ändern, die im Mai 2016 in Kraft getreten ist. Sie wird zwar nach einer Übergangsfrist bis spätestens zum Frühjahr 2018 das Bundesdatenschutzgesetz (BDSG) ablösen, im Wesentlichen werden dabei aber die Grundzüge des BDSG fortgeschrieben, sodass sich die EU-DSGV nicht negativ auf das Datenschutzniveau in deutschen Rechenzentren auswirkt.

Das heißt beispielsweise, dass personenbezogene Daten nur unter Beachtung strenger, gesetzlicher Erlaubnistatbestände des Betroffenen verarbeitet werden dürfen. Auskünfte an Ermittlungsbehörden unterliegen den deutschen gesetzlichen Regelungen. Mit einem deutschen Provider, der deutschem Recht unterliegt, besteht auch ein höherer Schutz vor dem Zugriff der US-Behörden auf die in Deutschland gespeicherten Daten.

Wer die Kontrolle über seine Daten und die seiner Kunden behalten will, ist daher mit einer Lösung auf Basis hochsicherer deutscher Rechenzentren gut beraten. In einer aktuellen Bitkom-Umfrage im Auftrag von KPMG attestierten die meisten Befragten der Telekom und T-Systems ein hohes Niveau an Datenschutz und Compliance bei Public-Cloud-Lösungen. Auf den Plätzen zwei und drei folgten – mit deutlichem Abstand – die größten Anbieter aus den USA.

Aber nicht nur in den Bereichen Datenschutz und Compliance oder der Cloud- und Data-Center-Sicherheit sind deutsche Anbieter erste Wahl. Eine Untersuchung der Experton Group (Security Vendor Benchmark 2016) zeigt, dass hiesige Dienstleister auch in vielen anderen Sicherheitsdisziplinen über herausragendes Know-how verfügen.

Sicherheit muss praktisch denken

Sicherheitslösungen müssen heute leicht sein – leicht verständlich und ebenso leicht umzusetzen. Nur dann werden sie tatsächlich von den Anwendern genutzt. Viele scheuen noch die vermeintliche Komplexität solcher Lösungen. Daher ist es Aufgabe der gesamten deutschen IKT-Branche, Sicherheitsprodukte so anwenderfreundlich wie möglich zu machen – leicht eben.

Ein Markt für neue IT-Security-Lösungen und -Dienstleistungen made in Germany ist definitiv vorhanden. In Deutschland wird er in diesem Jahr um etwa 7 % auf 5 Milliarden Euro anwachsen. Die Experton Group schätzt, dass davon etwa die eine Hälfte auf Security Services inklusive Cloud und die andere Hälfte auf Security-Produkte entfallen wird: „Insbesondere Anforderungen in den Bereichen Schutz vor Cyberwar, DDoS-Angriffen, Data Loss Prevention etc., aber auch […] der Einfluss aktueller Trendthemen wie Digitalisierung, Cloud-Computing, Social Business, Industrie 4.0, Mobile Enterprise und Big Data werden dabei zu weiterhin steigenden Investitionen führen.“

Und dieses Geld ist sinnvoll angelegt, wie der Raub von 117 Millionen Passwörtern und viele andere aktuelle Beispiele zeigen. Denn künftig werden die Cyberangriffe noch heftiger und überraschender kommen, noch trickreicher werden und sich auch auf neue Bereiche des Wirtschafts- und Privatlebens erstrecken – zum Beispiel auf das Internet of Things. Man muss wahrhaftig kein Prophet sein, um eine Zunahme der Cyberattacken auf vernetzte Geräte und Prozesse der Industrie 4.0 vorherzusagen. Vor diesem Hintergrund ist es umso wichtiger, die eigenen Daten und die seiner Kunden zuverlässig zu schützen – am besten mit Sicherheitslösungen made in Germany.

Dr. Ferri Abolhassan,
Geschäftsführer Service Transformation Telekom Deutschland und verantwortlich für Telekom Security

Gefahrenlage in Echtzeit

In der Früherkennung zeigt sich die Qualität einer Sicherheitsschaltzentrale

Die Cyberkriminalität hat ein Ausmaß erreicht, das viele Rechenzentren, Admins und Unternehmen vor erhebliche Probleme stellt. Es fehlt ihnen schlicht an Personal und Ressourcen für die oft zeitkritischen Aufgaben der IT-Sicherheit. Genau an diesem Punkt setzen Security Operation Center an.

Für Security Operation Center (SOC) kursieren mehrere Definitionen. Die einen reduzieren den Begriff auf operative Betriebsleistungen, die anderen setzen ihn gleich mit der Erkennung von Cyberangriffen und der Einleitung von Gegenmaßnahmen. Wieder andere assoziieren damit die proaktive Risikominderung durch die Beseitigung identifizierter Schwachstellen.

Klar ist: Zu den Grundanforderungen gehört, dass ein SOC Logfiles überwachen und Angriffe erkennen muss und dass der Auftraggeber dazu ein Reporting erhält. Das ist aber lediglich die Basis und letztlich auch eindeutig old school. Aktuelle sicherheitsstrategische Entwicklungen weisen in eine andere Richtung: Neben der Prävention und Absicherung werden vor allem Früherkennung, Abwehr und Reaktion immer wichtiger.

Intern, hybrid oder extern?

Es gibt kein SOC-Standardbetriebsmodell, mit dem sich überall maximaler Nutzen erzielen lässt. Prinzipiell hat ein Unternehmen drei Möglichkeiten: Es kann ein eigenes SOC aufbauen, einen Hybrid-Ansatz verfolgen oder einen externen Service nutzen.

Ein erster Ansatz wäre, ein internes SOC aufzubauen und SOC-Services vor Ort bereitzustellen. Unternehmen, die sich für dieses Modell entscheiden, müssen auf umfangreiche interne Kompetenzen zurückgreifen können. Allerdings sind neue Mitarbeiter mit adäquaten Qualifikationen auf dem Markt schwer zu finden. Auch sind die Kosten nicht zu unterschätzen: Für den zwingend erforderlichen 24/7-Betrieb sind in der Regel mindestens fünf Mitarbeiter erforderlich.

Teilweise entscheiden sich Unternehmen auch dafür, einzelne Elemente des SOC-Betriebs auszulagern und ein SOC aufzubauen, das sie gemeinsam mit einem Partner betreiben. Die Vorteile liegen in der Zusammenarbeit mit externen Experten, die über Spezialkenntnisse verfügen, im Zugriff auf externe kontextbezogene Bedrohungsdaten bzw. in der Entlastung von Mitarbeitern. Beim Hybrid-Ansatz ist aber auf eine vernünftige Aufgabenverteilung und eine nahtlose Verknüpfung der Services zu achten. Häufig verbleibt beispielsweise das wichtige Thema Incident Response im Unternehmen selbst – ein durchaus sinnvoller Weg. Weniger Sinn macht hingegen die verschiedentlich propagierte und auch praktizierte Trennung von Tag- und Nachtbetrieb. Schnittstellenprobleme sind dabei vorprogrammiert.

Am sinnvollsten ist letztlich eine vollständige Auslagerung des SOC-Betriebs. Auch große Konzerne mit riesigen IT-Abteilungen sind diesen Weg bereits gegangen, etwa Daimler. Und für Unternehmen, die über keine große IT-Infrastruktur verfügen und ihr Kapital nicht durch Technologieinvestitionen binden möchten, ist die Nutzung eines externen SOC praktisch der einzig gangbare Weg. Das Gleiche gilt für Unternehmen mit begrenzten Personalressourcen.

Ein externes SOC punktet vor allem mit einem zentralen Vorteil: Es kann als proaktives Abwehrzentrum fungieren, das auch eine Früherkennung von Angriffen unterstützt. Auf Unternehmensseite implementierte Systeme sind dafür allein schon aufgrund der Geschwindigkeit oft nicht geeignet. Ein modernes SOC setzt dagegen intelligente Tools ein, die eine permanente, automatische Analyse des Datenverkehrs und die Korrelation unterschiedlichster Informationen leisten. Hinzu kommt noch das Experten-Know-how von Cyber-Security-Analysten.

Global aufgestellt, lokal präsent

Eine zentrale Frage lautet: Soll sich ein Unternehmen für einen lokalen oder globalen Anbieter entscheiden? Die Antwort: Die beste Wahl ist der globale Anbieter mit lokaler Präsenz. Ein wichtiger SOC-Vorteil global aufgestellter Anbieter ist der umfassende Überblick über die Sicherheitslage. Ein weltweit agierender Provider kann Meldungen und Störungen von Tausenden Kunden überwachen und analysieren – und auf dieser Datenbasis dann ein Echtzeitbild der Bedrohungslage erzeugen, das er wiederum für die Erstellung effizienter Cyberabwehrlösungen nutzt. Nur eine globale Threat Intelligence kann letztlich die zuverlässige Basis für einen umfassenden Schutz vor akuten – auch vor ganz neuen – Bedrohungen sein.

Ebenso wichtig ist aber die lokale Präsenz – allein schon hinsichtlich gesetzlicher und aufsichtsrechtlicher Aspekte in Bezug auf Datenschutz, -zugriff und -haltung. Auch lokale Anbieter können hier punkten. Zugleich darf man nicht vergessen, dass es bei der Auslagerung des SOC-Betriebs immer um Themen wie Vertrauen und Investitionssicherheit geht. Das Thema Sicherheit einem relativ kleinen Unternehmen anzuvertrauen, dürfte nicht der Weisheit letzter Schluss sein.

Passgenaue Lösungen finden

Durch die Investition in ein geeignetes SOC haben sich bereits viele Unternehmen gegen Cyberbedrohungen aller Art gewappnet. Dabei heißt „geeignet“ allerdings auch, dass es sich um ein SOC handelt, das nicht nur als zentrales Archiv für die Logdaten aus Netzwerk-, Infrastruktur- und Sicherheitsprodukten fungiert. Es muss vielmehr eine Sicherheitsschaltzentrale sein, in der proaktive Prozesse zum Bedrohungs-, Vorfalls- und Schwachstellenmanagement etabliert sind und ein vorausschauendes Risikomanagement betrieben wird. Wichtig ist vor allem eines: Nicht das Unternehmen muss zum Lösungs- und Produktangebot des Betreibers passen, sondern das SOC muss immer zum Unternehmen und seinen spezifischen Anforderungen passen.

Patrick Schraut,
Director Consulting & Governance,
Risk und Compliance (GRC) DACH bei NTT Security

Mehrwert aus der Pflichtübung

Risiko- und Kostenkontrolle sind die wirtschaftlichen Argumente für ein ISMS

Für Betreiber kritische Infrastrukturen ist ein Information Security Management System praktisch Pflicht, vielen weiteren Unternehmen ist es dringend anzuraten. Zwar ist der Aufwand gerade zu Beginn je nach Rahmenwerk spürbar bis groß, doch erst dann kann man Risiken und Kosten in den Griff kriegen.

Immer mehr Unternehmen führen ein Information Security Management System (ISMS) ein. Die Gründe dafür reichen von gesetzlichen Anforderungen (zum Beispiel durch das IT-Sicherheitsgesetz) über Corporate Governance bis hin zu Kundenwünschen. Ziel der Anstrengungen ist oft die Zertifizierung, auf internationaler Ebene nach dem Standard ISO 27001, national auch nach dem BSI-Grundschutz. Die Akzeptanz in den Unternehmen selbst bleibt jedoch oft mittelmäßig.

Schon der sperrige Name „Informationssicherheitsmanagementsystem“ lässt vermuten, dass man es hier mit einem papierlastigen Bürokratiemonster zu tun bekomme. Diese Einstellung drückt sich vom Vertriebsmitarbeiter bis hin zum Top-Manager oft in einer achselzuckenden Sachzwangergebenheit aus: „Die Kunden wollen das“, „Das liest sowieso nie wieder jemand“ oder „Hauptsache, wir haben jetzt das Zertifikat fürs Marketing“.

Trügerische Sicherheit

Sowohl die Auffassung „Wir sind sicher genug, wir brauchen kein ISMS“ als auch das sorglose „Wir haben ein Zertifikat, wir sind sicher“ kann sich als gravierender Trugschluss erweisen. Nehmen wir zum Beispiel ein Unternehmen, das einen hohen Sicherheitsstandard, aber kein ISMS hat. Der IT-Leiter hat viel Aufwand und Geld in eine Netz-

werkstruktur investiert, sich Beratung der Hersteller eingekauft und seine Administratoren vorbildlich geschult.

Solange aber nur in technische Maßnahmen investiert wird, ist es offensichtlich, dass die Organisation zum Beispiel gegen Social-Engineering-Angriffe schlecht gewappnet ist, also etwa gegen Schadcode, der per E-Mail-Anhang oder infizierte USB-Sticks eingeschleppt wird. Die größte Schwachstelle ist nun einmal der Mensch selbst. Hier setzen Schulungen und Sensibilisierungen an, die das Sicherheitsbewusstsein der Mitarbeiter stärken. Eine ebenso problematische Situation entsteht, wenn nach der ISMS-Einführung oder einer überstandenen Zertifizierung das Thema ad acta gelegt wird, Dokumente im Schrank verstauben und Prozesse nicht eingehalten werden. Das Managementsystem muss kontinuierlich betrieben werden, sonst verfehlt es seine Wirkung.

Information Security Management

Abstrakt ausgedrückt ist ein ISMS ein Steuerungsprozess für Organisationen, die ihre geschäftskritischen Informationen absichern wollen. Das wichtigste Instrument dazu ist das Risikomanagement: Fach- oder Produktverantwortliche müssen Gefahren konsequent identifizieren, bewerten und mit der Unterstützung sicherheitsgebender Bereiche wie der IT-Abteilung oder Gebäudeverantwortlichen angehen. Das schafft Risikotransparenz, sodass Kosten und Nutzen von Investitionen auf dem Gebiet der Informationssicherheit und damit Risiken selbst gesteuert und kontrolliert werden können. Um ein ISMS geeignet aufzusetzen, muss sich ein Unternehmen also zunächst über sein Umfeld, seine (schutzwürdigen) Werte und seine Ziele klar werden. Gut vorbereitet ist eine Organisation, die ihre Geschäftsprozesse versteht, im besten Fall modelliert hat und die ihre Abhängigkeiten von den benötigten Ressourcen kennt.

Die Sicherheitsziele, etwa bezüglich der Vertraulichkeit, können in diesem Fall leicht vom Geschäftsprozess übertragen werden. Vereinfacht gesagt „vererben“ sich die Sicherheitsanforderungen vom Geschäftsprozess bzw. den dazu benötigten Informationen über softwaretechnische Systeme bis zum Gebäude und auf die Umgebung. Die verbreitetste Norm ist der Standard ISO/IEC 27001, der die Anforderungen an ein ISMS formuliert. Seine Stärke und gleichzeitig seine Grenze liegt darin, dass er nur den Management-Rahmen vorschreibt, aber nicht weiter auf konkrete Sicherheitsmaßnahmen eingeht; prinzipiell ist er jedoch auf jede Organisation anwendbar.

Der Ball liegt beim Management

Damit die ISMS-Einführung gelingt, sollte eine Organisation sich über die wichtigsten Erfolgsfaktoren im Klaren sein. Ein ISMS hat mittelbare und unmittelbare Auswirkungen auf die komplette Organisation, zum Beispiel durch die Veränderung von Personaleinstellungsprozessen über Änderungen an der Gebäudesicherung bis hin zur Konfiguration von Systemen. Die Grundvoraussetzung ist, dass das Vorhaben durch das Topmanagement unterstützt wird. Dies fordert die ISO/IEC 27001 auch explizit. Der Versuch, ein ISMS von einer unteren Ebene aus einzuführen, führt in der Regel zu großen Problemen – schon mangels Durchgriff auf die beteiligten Organisationsbereiche.

Das Management wird vor allem die Risikotransparenz und die daraus folgenden besseren Steuerungsmöglichkeiten der IT-Budgets zu schätzen wissen. Die Evaluierung von Risiken ist zwar niemals leicht, da es an sinnvollen quantitativen Verfahren fehlt und Daten der Vergangenheit in vielen Fällen nicht auf die Zukunft übertragen werden können. Man begegnet diesem Umstand in der Praxis aber durch die Anwendung von qualitativen Verfahren unter Beteiligung breit gefächerter Qualifikationen und Erfahrungen. Risiken werden nicht beziffert, sondern in Klassen eingeteilt. Diesen Ansatz bezeichnet man auch als semiquantitativ. Um das Sicherheitsbewusstsein im Unternehmen zu stärken, muss das Management sich seiner Vorbildrolle bewusst sein. Aber es kommt auch maßgeblich auf jeden einzelnen Mitarbeiter an. Schon bei der ISMS-Einführung sollte man zum Beispiel die Mitarbeiter direkt oder indirekt am Projekt beteiligen, ihnen die Wirkung und Ziele des Systems erklären, sie regelmäßig über das Projekt informieren und auf diese Weise Widerständen begegnen, die größtenteils aus Unkenntnis entstehen.

Einführung und Umsetzung

Nachdem das Top-Management sein Bekenntnis zum Thema Informationssicherheit in einer Leitlinie verabschiedet hat, werden der Geltungsbereich (Scope) des ISMS festgelegt und die ISMS-Organisation aufgebaut. Je früher sie etabliert ist und ihre operative Arbeit aufgenommen hat, desto besser für den Projektverlauf. Auf diese Weise kann man bereits zur Einführung die nötige Routine und das Vertrauen in die neuen Prozesse gewinnen.

Dies unterstützt eine erfolgreiche Zertifizierung erheblich. In der Folge werden innerhalb des ISMS-Geltungsbereiches die zentralen Werte

Quelle: secunet

Ein ISMS ist als systematischer und kontinuierlicher Prozess zu verstehen, der das gesamte Unternehmen betrifft.

ermittelt, die das Unternehmen schützen möchte. Weitere Schritte sind der Aufbau des Risikomanagements einschließlich der Planung und Umsetzung der sich daraus ergebenden Sicherheitsmaßnahmen. Natürlich müssen auch die Abläufe etabliert werden, die aus dem ISMS einen kontinuierlichen, steuernden Prozess machen.

Der kritische Pfad im Projekt beginnt erfahrungsgemäß mit der Festlegung eines stabilen Geltungsbereichs. Organisationen tun sich hier oft schwer, da verschiedene Interessen aufeinanderstoßen. Beispielsweise haben Produktverantwortliche die Kundeninteressen im Blick und sehen einen bestimmten Geltungsbereich vor, während der IT-Leiter sich wegen einer angeblich mangelnden Prozess- und Dokumentationsreife am liebsten ganz aus dem Thema halten möchte und das Management nur auf die Kosten schaut. Hier kommt es auf einen vernünftigen Interessenausgleich an. Ist der Geltungsbereich definiert, so sind der Aufbau und die Durchführung des Risikomanagements von größter Bedeutung. Erst nach der Risikobewertung kann eine sinnvolle Auswahl der Maßnahmen erfolgen. Auf jeden Fall sollte eine Organisation, die eine schnelle Zertifizierung anstrebt, schon projektbegleitend den Lebenszyklus des ISMS starten, damit sie beim Audit schon die praktische Funktionsfähigkeit der beteiligten Regelungen und Prozesse nachweisen kann.

Konkret heißt das, dass zu diesem Zeitpunkt bereits (Nachweis-)Dokumente wie beispielsweise ein Management-Review vorliegen müssen, die bereits Erfahrungen und Befunde aus dem laufenden Betrieb berücksichtigen und bewerten. Für diejenigen, denen Aufbau und Betrieb eines ISMS zu kompliziert oder teuer erscheint, was zum Beispiel bei Kleinunternehmen der Fall ist, gibt es einen Ausweg: Wesentliche Bestandteile eines ISMS lassen sich auch ausgliedern. ISMS as a Service ist keine Zukunftsvision mehr, und externe Informationssicherheitsbeauftragte sind keine Seltenheit. Vorbild ist hier der Umgang mit Datenschutzsystemen.

Grundlage gezielter Steuerung

Trends wie die Digitalisierung der Energiewende, die zunehmende Kooperation der klassischen IT mit der Prozessleittechnik zum Beispiel im Umfeld der Versorger oder Auswirkungen von Industrie 4.0 erfordern in vielen Organisationen ein Umdenken, das eine Ausweitung der Informationssicherheit auch auf Bereiche zur Folge hat, die bisher nicht im Fokus standen. In jedem Fall darf Informationssicherheit sich nicht auf technische oder organisatorische Einzelmaßnahmen beschränken, sondern muss als ein unternehmensweiter, kontinuierlicher Prozess betrieben werden. Es ist nicht einfach, ein ISMS einzuführen und auch nicht, es am Leben zu erhalten – und es kostet Geld. Doch nur so ist Risikotransparenz zu schaffen, damit aus einer diffusen Bedrohungslage ein scharfes Bild wird; Kosten und Nutzen lassen sich erst dann effektiv steuern.

Dipl.-Phys. Alexander Schlensog,
Dr. rer. nat. Hans-Jürgen Heinrich,
Division Kritische Infrastrukturen, secunet AG

Intelligente Einbauschränke

Wie leicht ein späterer Umbau wird, hängt von der Art des Rack-Systems ab

Rechenzentren sind langfristig angelegte Investitionen mit Betriebszeiten von bis zu 20 Jahren und wollen entsprechend überlegt geplant sein. Wer eine Modernisierung oder eine Erweiterung der IT-Umgebung plant, tut sich mit standardisierten und flexibel ausbaubaren Rack-Systemen deutlich leichter.

Am Anfang einer Investitionsentscheidung steht die Frage, wofür der IT-Schrank genutzt werden soll. Generell lassen sich in einem IT-Rack sowohl Server- als auch Netzwerktechnik montieren. Die standardisierte Grundausstattung besteht aus einer flexiblen 19-Zoll-Montageebene, geteilten Seitenwänden mit Schnellverschluss und einer optimierten Kabeleinführung mit Bürstenleisten. Je individueller der Innenausbau sein soll, desto eher empfiehlt sich eine Schnellmontagetechnik. Damit lassen sich Zubehörkomponenten wie Geräteböden und Gleitschienen werkzeuglos durch nur eine Person innerhalb weniger Minuten montieren.

Die spätere Nutzung bestimmt schließlich die nötigen Abmessungen. In einem Netzwerkschrank werden einzelnen Komponenten seitlich umfassend verkabelt, sodass dieser meist eine Breite von 800 mm hat und eine Tiefe von bis zu 1000 mm. Für den reinen Serverschrank ist dagegen eine Breite von 600 mm ausreichend, da sich die Strom- und Netzwerkkabel in der Regel auf der Rückseite befinden. Der Schrank ist 1000 bis 1200 mm tief. Bei Mischbestückung mit Server- und Netzwerktechnologie muss die Breite 800 mm betragen, die Tiefe 1000 bis 1200 mm. So lassen sich Netzwerkverteiler, Patch-Felder und PDUs (Power Distribution Units) sowie größere Mengen an Kabeln bequem installieren. Die jeweils passende Höhe wird über die benötigten Höheneinheiten (HE) ermittelt. Ein Schrank mit 42 HE erreicht etwa 2 m; das ist heute die meist verwendete Größe.

Schutzklassen nach Standort

Über den Schutzbedarf entscheidet der Standort des IT-Schranks. Die Schutzklasse wird nach der internationalen IP-Norm (International Protection) klassifiziert. In einer Büroumgebung ist ein Zugriffsschutz gemäß IP 20 mit einer abschließbaren Tür ausreichend. Verschließbar sollte ein Schrank auch in einer Produktionshalle sein; dort müsste er außerdem IP 55 genügen, also zum Beispiel gegen Staub und Strahlwasser geschützt sein. In einem Raum mit Schloss sowie in einem geschlossenen Rechenzentrum genügt meist ebenfalls IP 20, da ohnehin nur Befugte Zutritt haben.

Quelle: Rittal

Idealerweise erlaubt das Rack-System den Einbau von PDUs im Zero-U-Space.

Der IT-Schrank ist zugleich ein wichtiges Element des Sicherheitskonzepts, denn er verhindert den physischen Zugriff unbefugter Personen und schützt die IT-Komponenten vor Gasen, Feuer und Rauch, Trümmerlasten, Staub, Wasser oder EMV-Strahlung. Diese Anforderungen erfüllt allerdings nicht jeder Schrank. Wer ein hohes Schutzniveau anstrebt, greift auf spezielle Sicherheitslösungen zurück, die eine zusätzliche Umhausung des IT-Racks bieten. Für kleine Firmen bietet schon ein sogenanntes Micro Data Center genügend Platz für Server, Switches und Speichersysteme.

Standards sparen Handarbeit

Wie häufig Umbauarbeiten an den IT-Systemen stattfinden, wird maßgeblich die Kaufentscheidung bestimmen. Schließlich können einheitliche Technologien und Produkte den Betriebsaufwand deutlich senken. Das fängt bei der Planung von neuen Schrankreihen an: Wenn nur ein Modell eines Schranksystems infrage kommt, ist klar, wie viele 19-Zoll-Racks im Raum Platz finden. Öffnungen für Kabelzuführungen und Anschlüsse sind dann durchgehend an derselben Stelle, sodass eine Erweiterung mit Serverracks frühzeitig vorbereitet werden kann.

Außerdem sind gleich aufgebaute Schrankreihen aufeinander abgestimmt, sodass man die verfügbare Fläche besser nutzen kann. Damit ergibt sich für das Rechenzentrum eine höhere Packungsdichte bei geringerem Platzbedarf. Auch bei der Inneneinrichtung hilft die Standardisierung, und wenn es an den praktischen Einbau geht, ist sie ohnehin von Vorteil: Alle Komponenten lassen sich mit den gleichen Werkzeugen oder über die gleiche Befestigungsmethode einbauen. Türen und Seitenwände passen in jedem Rack. Das Unternehmen muss auch weniger unterschiedliche Ersatzteile vorhalten und kann Lagerbestände auf das Minimum beschränken.

Da außerdem die Leistungsdichte pro Rack beständig zunimmt, stellen viele IT-Manager von der früher üblichen Doppelbodenkühlung über Kaltluft auf eine Direktkühlung durch Wärmetauscher in den Schrankreihen um. Der Wärmetauscher rückt so immer näher an die Hitzequelle heran, was einen deutlich effizienteren Betrieb ermöglicht. Einheitliche Schranksysteme sind eine der wichtigsten Voraussetzungen dafür, denn die Wärmetauscher sind entweder in die hinteren Schranktüren oder zwischen den Racks integriert.

Viele aktive IT-Komponenten verfügen über eine redundante Stromversorgung und haben dafür zwei Netzteile eingebaut. Für den Innenausbau heißt das, dass auch die PDUs entsprechend redundant aus-

gelegt sein müssen. Idealerweise erlaubt das Rack-System den Einbau von PDUs im Zero-U-Space, also dem Raum zwischen Seitenwand und 19-Zoll-Montagerahmen. Auf diese Weise werden keine Höheneinheiten blockiert, und auch bei einem voll ausgebauten IT-Rack sind noch Wartungs- und Installationsarbeiten möglich. Außerdem wird so die Luftführung nicht behindert.

Direkte Rack-Kühlung

Rechenzentren wurden früher durchweg über den Doppelboden gekühlt. Das Konzept ist simpel und günstig, aber ineffizient. Moderne Systeme arbeiten eher mit einer Gangeinhausung oder kombinieren beide Methoden. Die Gangeinhausung trennt durch einfache bauliche Maßnahmen die Kalt- und Warmluftseiten und konzentriert die Kaltluft auf dem Weg zu den Serverschränken. Dadurch wird verhindert, dass sich die kühle Luft mit ausgeblasener Abwärme vermischt. Noch effizienter arbeiten Rack-Kühlsysteme: Sie nehmen die Hitze noch innerhalb des Schranks auf, kühlen die Luft mit Wärmetauschern und blasen sie wieder an der Front ein. Die deutlich höhere Kühlleistung pro Quadratmeter ist eine wichtige Voraussetzung für moderne Serversysteme, die sehr viel Abwärme pro Flächeneinheit produzieren.

Innerhalb des Racks strömt bei Servern die kühle Luft von vorne nach hinten, bei Netzwerkkomponenten dagegen oft seitwärts. In beiden Fällen ist es wichtig, die 19-Zoll-Ebenen zu schotten, damit einmal erzeugte Kaltluft nicht ungenutzt an den Komponenten vorbeiströmt. Offene Höheneinheiten sollten darum geschlossen werden, um warme und kalte Luft zu trennen. Hier gibt es vielfältiges Zubehör, um die Luftführung zu optimieren oder eine seitliche Belüftung zu ermöglichen.

Ein Sonderfall sind Rechenzentren für High Performance Computing (HPC). Durch die extreme Leistungsdichte sind hier die Server auf engstem Raum konzentriert. So entsteht im Rack ein Verbrauch von bis zu 30 kW auf einer Fläche von weniger als 1 m^2. Bei dieser Energiedichte greifen HPC-Rechenzentren oft zu einer Rack-basierten Kühlung. Entscheidend ist, die Kühlkapazität möglichst nah an den Ort der Wärmeerzeugung zu bringen.

Monitoring und Management

Kein Rechenzentrum kommt heute ohne Überwachungssysteme aus, bei denen Sensoren die wichtigsten Werte im IT-Rack erfassen: Temperatur, Luftfeuchtigkeit, Spannung etc. Bei der Auswahl eines Rack-Systems ist daher zu prüfen, welche Module zum Monitoring der Hersteller verfügbar macht. Wer nur einige wenige Serverschränke betreibt, kommt übrigens oft auch mit der Software des Rack-Herstellers aus. Eine große DCIM-Lösung für das Monitoring kompletter Rechenzentren ist dann nicht mehr nötig. Ein entsprechendes System im Rack kann zum Beispiel Feuer und Schwelbrände automatisch erkennen. Es braucht nur eine Höheneinheit und flutet den Schrank mit Löschgas.

Moderne Rack-Systeme können viel dazu beitragen, das Servermanagement und die Inventarisierung der IT-Komponenten zu vereinfachen. Hierzu sind Managementlösungen auf RFID-Basis verfügbar. Eine hochpräzise RFID-Antenne, die auf voller Höhe neben der 19-Zoll-Ebene im Rack eingebaut ist, erkennt die montierten IT-Systeme und kann sie den Höheneinheiten zuordnen. Dann wissen Administratoren jederzeit, welche Komponenten auf welcher Höheneinheit der 19-Zoll-Ebene eingeschoben sind. Entsprechende Lösungen werden von etlichen Herstellern angeboten und sollten sich in eine Managementsoftware wie DCIM integrieren lassen.

Bernd Hanstein,
Hauptabteilungsleiter Produktmanagement IT, Rittal

Bitte eine Build-Prüfung!

Der Markt für Data-Center-Zertifizierungen gerät in Bewegung

In Deutschland ist TÜV IT einer der wichtigsten Anbieter von RZ-Überprüfungen vor Ort. Neue Konkurrenz könnte dem etablierten Zertifizierer jetzt durch das US-amerikanische Uptime Institute erwachsen. Allerdings startet der Wettbewerber ausgerechnet von Großbritannien aus auf den europäischen Markt.

Zertifizierungen werden demnächst wohl wichtiger. Das neue IT-Sicherheitsgesetz bezieht sich explizit darauf: Wer in Zukunft als Anbieter von Rechenzentrumsdienstleistungen anerkannte Zertifizierungen vorweisen kann, dessen Kunden können auch damit rechnen, dass ihnen zum Beispiel im Fall von Datenverlusten beim Provider zumindest nicht vorgeworfen wird, sie hätten ihren Dienstleistungspartner ohne die nötige Sorgfalt ausgewählt.

TÜV IT bis in die Umsetzung

Glaubwürdigkeit und Sorgfalt einer Zertifizierungsinstitution sind daher ein wichtiges Kriterium, während finanzielle Erwägungen eher zurücktreten dürften. Denn nur Zertifikate, deren Qualität überzeugt, zum Beispiel weil sie gründlich nach einer umfassenden Vor-Ort-Prüfung vergeben werden, erfüllen unter Umständen die Anforderungen der europäischen Gesetzgeber.

Eine wichtige Institution in Deutschland, die vor Ort Rechenzentren unter die Lupe nimmt, ist TÜV IT. Das Tochterunternehmen des TÜV Nord prüft Rechenzentren schon seit dem Jahr 2001. Bisher hat es 350 Zertifizierungen durchgeführt, davon rund 110 Erstzertifizierungen (Stand Sommer 2016). „85 % unserer Zertifizierungskunden kommen wieder“, sagt Joachim Faulhaber, stellvertretender Bereichsleiter IT-Zertifizierung und Product Manager Data Center bei TÜV IT. Zertifiziert wird nach der Spezifikation TSI (Trusted Site Infrastructure), und zwar immer bezogen auf den Gesamtprozess der RZ-Erstellung – von den Plänen über die Architektur bis hin zu den technischen Einrichtungen. Eine reine Design-Begutachtung bietet TÜV IT nicht an.

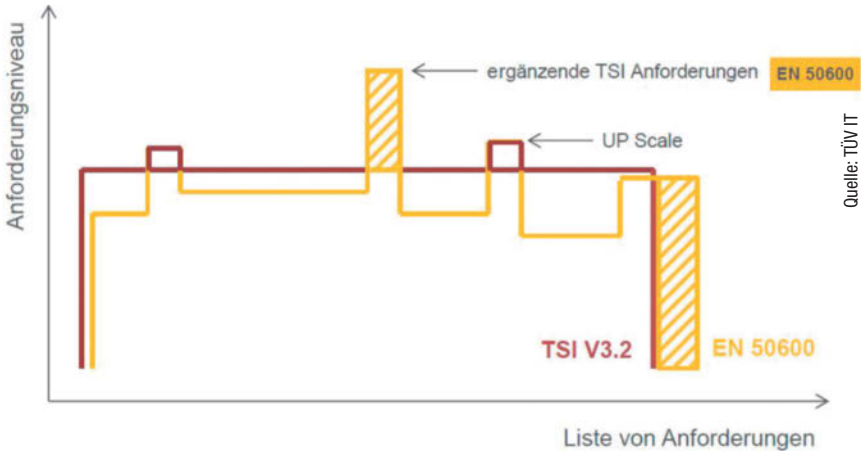
Hier liegt einer der wesentlichen Unterschiede zum Konkurrenten The Uptime Institute, das just im Jahr des Brexit-Votums seine europäischen Aktivitäten mithilfe einer europäischen Zentrale, die gerade in London eingerichtet wurde, auszuweiten sucht. Kurz vor dem Votum stellten einige Vertreter des Uptime Institutes bei einer Veranstaltung in München das Angebot des Zertifizierungs-, Beratungs- und Schulungsunternehmens vor.

Uptime in drei Varianten

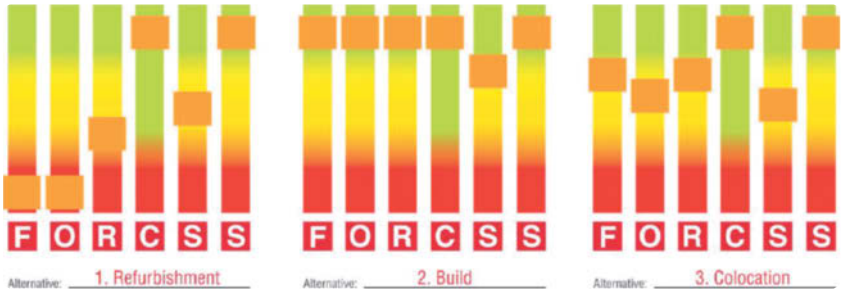
Uptime bietet mehrere Zertifizierungen an: eine Design-, eine Build- und eine Operate-Variante. Die Design-Zertifizierung wird separat und unabhängig von den anderen Typen vergeben, was schon Anlass für Kritik war. Des Weiteren berät Uptime RZ-Betreiber und hat ein Vorgehensmodell für die Entscheidung zwischen Public, Cloud, Outsourcing und Eigenbetrieb entwickelt, das Kunden ebenfalls gegen Honorar als Beratungsleistung in Anspruch nehmen können.

The Uptime Institute verzeichnet annähernd 1000 Zertifizierungen in 80 Ländern für sich, darunter viele in den USA, aber auch in arabischen Ländern. Eine geografische Übersicht im Web zeigt, wo das Unternehmen bisher aktiv geworden ist. Die Karte hat freilich gerade in Deutschland und einigen anderen europäischen Ökonomien ein aus Sicht des Uptime Institute ärgerliches Loch.

Wie viele der 1000 Zertifizierungen genau zu welchem Typ gehören, war trotz schriftlicher Nachfrage beim Presseverantwortlichen des Unternehmens, Ali Moinuddin, nicht herauszubekommen. Er legte sich in der schriftlichen Antwort auf eine diesbezügliche E-Mail-Anfrage nur darauf fest, dass es immer mehr Design-Zertifizierungen als andere gebe. Man bewerbe diese Form inzwischen zurückhaltender, obwohl sie weltweit weiterhin möglich sei. Der Grund: Manche Kunden brüsten sich zwar mit der Design-Zertifizierung, erwerben aber niemals eine Build-Zertifizierung. Das sei unerwünscht, schreibt der Sprecher des Uptime Institute, da das tatsächlich gebaute RZ oft in wichtigen Einzelheiten vom geplanten abweiche. Papier ist ja bekanntlich geduldig, und dieses Vorgehen wäre letztlich Betrug am Endkunden des RZ-Betreibers, der wahrscheinlich auch das Design-Zertifikat als Indiz für zielkonforme Ausführung ansieht. Insofern ist die Politik des Uptime Institute sicher auf dem richtigen Weg. Es fragt sich aber, ob man überhaupt eine unabhängige Design-Zertifizierung vergeben sollte. Denn letztlich geht es für den Endkunden – und an den richten sich die Prüfnachweise ja schließlich – darum, RZ-Leistungen in Anspruch zu nehmen und nicht utopische Pläne zu bewundern.



Schematische Darstellung der Erweiterungen und Veränderungen in TSI durch EN 50600.



Quelle: The Uptime Institute

Die Lebenszyklusanalyse, die Uptime als Beratungsleistung neben den Zertifizierungen anbietet, hilft, eine fundierte Entscheidung zwischen Eigenbetrieb und verschiedenen Formen der Auslagerung zu treffen.

Bleibt die Feststellung: Wie viele physisch existierende Rechenzentren tatsächlich von Uptime hinsichtlich Aufbau, Betrieb und Funktion zertifiziert wurden, lässt sich nicht genau sagen. Wahrscheinlich sind es einige mehr als die 110 Rechenzentren, die TÜV IT für sich reklamiert, aber sicher längst nicht 1000, sondern eher einige Hundert. Das M&O-Siegel (Management and Operations), soweit legt sich das Management immerhin fest, wurde bisher rund 70 Mal vergeben.

Kooperationen und Kontrolle

Weitere Daten zu Uptime: Das Unternehmen beschäftigt etwa 200 Mitarbeiter, hat rund 1800 RZ-Spezialisten ausgebildet und seine Gründer sind teils dieselben wie die von The Green Grid, einer Organisation, die sich bemüht, mehr Ökologie beim Bau und Betrieb von Rechenzentren zu erreichen, und bei der Erarbeitung entsprechender Kennziffern eine globale Vorreiterrolle einnimmt. Paradebeispiel ist der Wert PUE (Power Usage Effectiveness), der heute schon beinahe zu den Standard-Parametern bei der Bewertung von Rechenzentren gehört.

Rund 90 Mitgliedsunternehmen weltweit, davon 23 aus Europa, stehen im Hintergrund des Uptime Institute. Die Vertreter der Unternehmen treffen sich halbjährlich. Sie informieren einander über Verbesserungsmöglichkeiten und neue Technologien, wobei diese Auskünfte auch in die Zertifizierungsprüfungen von Uptime einfließen. Regelmäßig besichtigen die Mitglieder jeweils das Rechenzentrum eines anderen Mitgliedes und „sehen dort nach dem Rechten“, sprich: stellen fest und teilen dem Betreiber mit, ob es dort etwas zu verbessern gibt. Diese Treffen sind streng vertraulich.

Ein wichtiger Unterschied zwischen The Uptime Institute und TÜV IT: Während TÜV IT sich als reiner Zertifizierer versteht, ist Uptime wie schon erwähnt auch in der Beratung aktiv. Kritiker sehen darin eine unerwünschte Überschneidung zwischen Beratungs- und Prüfungsgeschäft. Andererseits gibt es Stimmen, die TÜV IT mangelnde Praxisorientierung vorwerfen. Zudem werden Fehler an Rechenzentren mit Ursachenanalyse in einer öffentlich über das Web zugänglichen, recherchierbaren Datenbank gespeichert.

Prüfkriterien – offen oder nicht?

Eine kontroverse Diskussion dreht sich um die (kostenlose) Verfügbarkeit der Prüfkriterien im Vorfeld einer RZ-Überprüfung. Sortiert man das Wortgeklingel der beiden Prüfspezialisten, stellt sich heraus, dass die Situation letztlich nicht so unterschiedlich ist, wie man meinen könnte: TÜV IT veröffentlicht einen Prüfleitfaden, der kostenlos erhältlich ist und aufführt, was geprüft wird, nicht jedoch die technischen Details und deren Realisierungsweisen, die vorliegen müssen, damit ein Kriterium als erfüllt gilt. Diese erfahren nur die Kunden, die sich auf einen Zertifizierungsprozess einlassen, in einem einleitenden Workshop. Zudem hören sie von TÜV IT in den einleitenden Prüfschritten auch, an welchen Stellen ihr bisheriges Design, ihre Bau- und Betriebsweisen noch nicht den Regeln entsprechen. Wie sie das Problem lösen, darüber äußere sich TÜV IT nicht, sagte Joachim Faulhaber anlässlich einer Präsentation in München. Den Kunden direkt und möglicherweise gegen zusätzliches Honorar Lösungsvorschläge zu unterbreiten, sei eine unzulässige Überschneidung zwischen Beratung und Prüfung, meint der Manager.

Auch bei Uptime gibt es die grundlegenden Zertifizierungsinformationen im Internet, Details dazu, was nun ein Kriterium erfüllt, aber erst im Prüfprozess. Die Begründung lautet hier: Die Zertifizierung sei an grundsätzlichen Zielen orientiert, nicht an bestimmten Implementierungsvarianten. Diese

EN-50600-KONFORME TSI-4.0-ZERTIFIZIERUNG

Bau: Zugangswege und gesicherte Anlieferung, Umsetzung eines Schutzzonen-Schalenkonzepts

Sicherheitssystem und -organisation: sichere Alarmübertragung

Verkabelung: Struktur der Kommunikationsverkabelung, redundante, separierte Telekommunikationskabelführung, Ausführung von Kreuzungspunkten, Schutz der Datenleitungen vor Störquellen, Kabelführung bei Schränken und Gestellen, WAN-Versorgung über mindestens zwei Provider

Organisation: Lifecycle Management, Kundenmanagement, Erfassung wichtiger Schlüsselindikatoren

Dokumentation: Standortgutachten bei Neubauprojekten, Dokumentation der Telekommunikationsverkabelung, Leistungsbedarfsrechnung und Auslegungsvorgaben, Klimatisierungskonzept, Regelungen für RZ-Kunden bei Fremdvermietung

grundlegenden Ziele ließen sich nun einmal nicht in Form von Checklisten oder anderen formalisierten Vorgaben prüfen. Wie die Ziele, beispielsweise ein bestimmter Sicherheitsgrad bei der Stromversorgung, erreicht werden, bleibe grundsätzlich dem Kunden überlassen, weshalb es ohnehin sinnlos sei, detaillierte Erfüllungsmöglichkeiten vorzugeben. Uptime gibt den Kunden aber nach eigenen Angaben durchaus Ratschläge, wie sie ein bestimmtes Kriterium erfüllen können. Das dürfte durch Zeit- und Beratungsaufwand die Kosten nach oben treiben, erspart aber den Weg zu anderen Beratern und Partnern.

Aufwand, Kosten und Ergebnis

Eine weitere Gemeinsamkeit ist bei allen Zertifizierungen gebauter Rechenzentren und des RZ-Betriebs, dass sie grundsätzlich persönlich und vor Ort durchgeführt werden. Auch TÜV IT sieht sich die Baupläne etc. im Zertifizierungsdurchlauf an, allerdings niemals unabhängig von der Bau-Zertifizierung. Die reine Design-Zertifizierung von Uptime dagegen ist – natürlich – eine reine Papierprüfung. TÜV IT betont ferner die Interdisziplinarität seines Prüfungspersonals – niemand könne schließlich gleichermaßen Experte in IT-Sicherheit, Kühltechnik und effizientem Betrieb der Stromversorgungsanlagen sein. Oft werde daher, so Faulhaber, ein Antrag von mehreren Spezialisten durchgesehen, die, falls notwendig, auch vor Ort erscheinen. Für eine Prüfung veranschlagt der Manager zwischen 10 und 35 Manntage, je nach Aufwand vor Ort und bei der vorherigen Dokumentenprüfung. Pro Manntag fallen branchenübliche Tagessätze an. Abweichungen im Zeitbedarf nach oben und unten sind selbstverständlich möglich.

Uptime betont die Praxiserfahrung seiner Prüfer im RZ-Bereich und die internationale Einheitlichkeit der Prüfungen, da die Prüfer weltweit eingesetzt werden. Fraglich ist aber, ob sich dieses Konzept auch dann durchziehen lässt, wenn Uptime tatsächlich in Europa stärker expandiert, oder ob man aus Effizienzgründen nicht am Ende doch eine europäische Prüfmannschaft aufbaut, die dann vielleicht einen explizit europäischen Blick auf die Rechenzentren entwickelt. Zu solchen Szenarien wollte sich Uptime bei der Präsentation in München aber nicht äußern.

Was die Kosten angeht, bleibt Uptime eher kryptisch. Der Prüfungsaufwand sei schließlich nicht unerheblich. Bei einem Rechenzentrum für 20 Millionen US-Dollar sollen die Kosten bei etwa 5 % liegen, mithin bei 100.000 US-Dollar pro Zertifizierung. Für kleinere Rechenzentren gelte nicht in jedem Fall auch ein entsprechender Abschlag, schließlich gebe es Fixkosten.

The Uptime Institute legt größten Wert darauf, dass nur es selbst die „echte“ (geschützte) Tier-Zertifizierung ausgebe, zu erkennen an der Angabe des Zertifizierungsniveaus in römischen Ziffern von I bis IV. Demgegenüber spricht TÜV IT von Level 1 bis 4. Eine Eins in beiden Schreibweisen steht für das niedrigste, eine Vier für das höchste Schutzniveau. Weitere Unterschiede ergeben sich durch die EN 50600. Aktuelle Prüfungen gibt es in drei Varianten: Für neue Data Center empfiehlt TÜV IT eine Prüfung nach TSI 4.0 (Trusted Site Infrastructure). Zudem ist die Zertifizierung nach der Vorversion 3.2 noch einige Jahre

möglich, sie deckt bereits rund 90 % der EN-50600-Anforderungen ab. Diese empfiehlt Faulhaber beispielsweise Kunden, die eine Zertifizierung neu erwerben oder eine bestehende verlängern wollen, deren Anlagen aber nicht mehr allzu lange laufen sollen.

TSI 4.0 nach EN 50600

Die EN-50600-konforme Version bietet, wie auch die zugrunde liegende Norm, die Möglichkeit, für jeden Bereich ein bestimmtes Niveau anzustreben, das die Vorgaben erfüllt. So definiert sie für Kälteversorgung, Stromversorgung und Netzverkabelung vier Verfügbarkeitsklassen, weiter gibt es vier Schutzklassen und ein Schalenschutzkonzept, sodass der Schutz je nach RZ-Bereich unterschiedlich ausfallen kann. Bei Management und Betrieb, beides ist in die Norm integriert, gibt es vier Realisierungsniveaus mit niveauangepassten Umsetzungsvorschlägen. Für das Gebäude dagegen sind keine unterschiedlichen Level vorgesehen. TÜV IT prüft bei jeder Zertifizierung im Prinzip alle Kriterien, allerdings entfallen einzelne, sofern nur ein niedriges Level angestrebt wird.

Das heißt: Die akzeptablen Umsetzungsvarianten eines bestimmten Prüfkriteriums ändern sich mit den Ansprüchen an die Lösung – je nachdem, welches Niveau das RZ in dem betreffenden Bereich erreichen will. Praxisrelevant sind in erster Linie die Level 2 und 3.

Das Rechenzentrum insgesamt gilt am Ende auf der Stufe zertifiziert, die alle Bereiche erfüllen, wobei Abweichungen nach oben, zum Beispiel im Bereich Sicherheit, extra angegeben werden. Wann welcher Level erforderlich ist, ergibt sich aus der Risikoanalyse, die der erste Teil der Norm EN 50600 verpflichtend fordert. Wie diese aber abzulaufen hat, ist den Anwendern weitgehend freigestellt. Zudem gibt es Bereiche, in denen die Norm es einzelnen europäischen Ländern überlässt, beispielsweise strengere Regeln durchzusetzen. „EN 50600 ist eher Rezeptbuch als Checkliste“, betont Faulhaber, und entsprechend komplex sind die Vorüberlegungen der Kunden und die Zertifizierungen der Prüfer. Die voll zur EN 50600 kompatible TSI 4.0 prüft unterm Strich 166 bekannte Kriterien sowie 18 neue, die mit der Europanorm hinzugekommen sind. Die bisherigen Kriterien wurden im Vergleich zur Vorgängerversion teilweise umformuliert, was aber nicht immer auf eine Verschärfung hinausläuft. Dabei geht es oft sehr ins Detail.

Und schließlich kommt in EN 50600 das Thema IT-Sicherheit in Form von Einrichtungen und Prozessen vor – von der sicheren Beleuchtung des RZ-Umfelds bis zur Datenverschlüsselung. TÜV IT hat Sicherheitsthemen allerdings schon in den Vor-Norm-Zeiten zum Prüfungsbestandteil gemacht.

Hier klafft, zumindest was die europäischen Rechenzentren angeht, eine Lücke im Prüfprogramm des Uptime Institute. Pressesprecher Moinuddin schreibt in der erwähnten Anfrage: „Sicherheit […] wird determiniert durch den Standort, die Branche, die üblichen Vorgehensweisen im Unternehmen, seine Geschichte sowie andere Faktoren. Die Einordnung in einen Sicherheitsbereich des Unternehmens diktiert meist, wie Sicherheitsmaßnahmen gehandhabt werden. Wir glauben nicht, dass Sicherheit im RZ fehlt. Es wird viel Aufmerksamkeit auf das Thema gerichtet, aber es ist keine wichtige Ursache von Rechenzentrumszwischenfällen und Betriebsunterbrechungen. In unserer Datenbank [von RZ-Zwischenfällen] haben wir nicht einen einzigen Zwischenfall registriert, der auf eine Durchbrechung von Sicherheitsmaßnahmen zurückgeht“ (Übersetzung: A. R.). Es bleibt jedem Anwender selbst überlassen, ob ihm das ausreicht – der Europanorm genügt es nicht.

Strategien nach dem Brexit

Insofern wird es wohl stark von der weiteren Brexit-Entwicklung abhängen, wie Uptime auf den kontinentaleuropäischen Märkten agiert, wo in Zukunft die Norm gilt. Will Uptime dortige Kunden überzeugen, könnte man über eine Erweiterung der Prüfkriterien spekulieren. Auch die Einrichtung einer Niederlassung auf dem europäischen Festland dürfte sich dann empfehlen.

Sucht Uptime aber angesichts der politischen Entwicklung und der Eigenheit europäischer Normen doch lieber in den angestammten Märkten und Großbritannien nach neuen Kunden, dürfte sich die europäische Aufstellung mit London als Zentrale bewähren.

Ariane Rüdiger,
freie Autorin, München

Unter den Wolken …

Cloud-Szenarien erfordern eine perfekt skalierbare Netzwerkinfrastruktur

Vor allem weil sie Anwendern viel Flexibilität und Wirtschaftlichkeit bieten, erfreuen sich Cloud-Lösungen ungebrochen großer Beliebtheit – egal ob es sich um eine Public, eine Private oder eine hybride Cloud handelt. Entscheidend ist, wie sich die Eigenschaften der Cloud im Netzdesign niederschlagen.

Der Betreiber einer Cloud, insbesondere einer Public Cloud, steht vor diversen Herausforderungen, die sich von denen klassischer Rechenzentren zum Teil deutlich unterscheiden. Sie alle sind jedoch einer einzigen Maxime geschuldet: der Skalierbarkeit. Das Netzwerk in einer Cloud-Umgebung muss skalierbar sein – und zwar in allen Belangen. Nur dann ist der Betreiber flexibel genug, um technisch und wirtschaftlich sinnvoll auf die Anforderungen seiner Kunden sowie auf eine veränderte Marktsituation zu reagieren.

Virtualisierung in Cloud-Rechenzentren

Was sind nun die Eigenschaften eines skalierbaren Netzes und wie wirken sie sich auf das Design aus? – Zunächst ist in diesem Zusammenhang ein hoher Grad an Virtualisierung zu nennen. Was sich selbst in klassischen Unternehmensnetzen schon seit geraumer Zeit als Trend beobachten lässt, gilt erst recht für die Cloud: Durch umfassende Virtualisierung werden die Dienste und Applikationen von der darunter liegenden Hardware entkoppelt. Das ergibt hohe Flexibilität und bietet vor allem die Möglichkeit, sozusagen im Baukastensystem neue Hardware zu installieren und die Gesamtlösung transparent zu erweitern.

Anders als in den meisten Unternehmensnetzen gilt dies in Cloud-Umgebungen jedoch nicht nur für die Server, sondern auch für den Storage-Bereich. Im Unternehmensnetzwerk kann man in der Regel den Speicherbedarf für einen definierten Zeitraum relativ gut vorhersagen. Daher ist es hier weiterhin opportun, mit getrennten Frontend- und Backend-Netzen zu arbeiten. Zwar werden diese mittel- bis langfristig beide auf der Basis von Ethernet betrieben werden, müssen aber nicht zwangsläufig konvergieren. Im Netz eines Cloud-Providers sieht dies ganz anders aus. Hier lässt sich der künftige Storage-Bedarf nur sehr schwer bis gar nicht vorhersagen. Will man also ein skalierbares Design für Cloud-Umgebungen schaffen, muss auch der Speicher einfach erweiterbar sein. Das geht am besten, wenn man den Storage ebenfalls virtualisiert.

Netzwerkdesign und Performance

Die Virtualisierung hat deutliche Auswirkungen auf das Netzwerk-Design. Zunächst gilt es, die zu erwartenden Verkehrsströme zu betrachten. Während im klassischen Rechenzentrum der überwiegende Teil der Pakete zwischen Client und Server verkehrt, erhöht sich in einer virtualisierten Umgebung der Verkehr im Rechenzentrum selbst und macht sogar den Löwenanteil aus. Gerade wenn man mit virtualisiertem Speicher arbeitet, wird das Netz sozusagen zum neuen Systembus. Der Server spricht mit seiner Festplatte nicht mehr über den internen Bus, sondern über das Netzwerk. Dazu kommen die Hochverfügbarkeit von virtuellen Servern oder das Wandern von virtuellen Maschinen von einer Hardware zur anderen (vMotion) und andere Merkmale. Diesen Datenverkehr bezeichnet man auch als Ost-West-Verkehr. Ein skalierbares Design in Rechenzentren muss also einem steigenden Anteil an Ost-West-Verkehr Rechnung tragen.

Darüber hinaus ist auch die Performance des Netzes insgesamt ein wichtiger Punkt. Hier ist der Anschluss der Systeme mit 10GbE mittlerweile Standard. Im Uplink sollten deshalb vorzugsweise 40GbE und in absehbarer Zeit auch 100GbE verwendet werden. Infolgedessen sollten Betreiber eines solchen Netzes bei der Auswahl der Switches auf eine hohe Port-Dichte von 40GbE und zumindest auf einen heute schon definierten Migrationspfad zu 100GbE achten. Zusätzlich etablieren Standardisierungsgremien wie IETF oder IEEE neue Geschwindigkeiten von 25GbE und 50GbE. Da diese die internen Ressourcen der Server besser ausnutzen, ist zu erwarten, dass 25GbE in absehbarer Zeit zur hauptsächlichen Anschlussgeschwindigkeit für Serversysteme werden wird. Bereits heute sind Systeme am Markt, die durch QSFP28-Optiken (Quad Small Form-factor Pluggable) 10, 25, 40, 50 und 100GbE auf demselben Interface erlauben.

Die transparente Layer-2-Wolke

Viele Rechenzentrumsnetze bzw. ihre Server und Applikationen benötigen große und schnelle Layer-2-Umgebungen. Dies erfordert Mechanismen zur Schleifenunterdrückung und für Redundanzen im Fall fehlerhafter Links. Der klassische Spanning Tree genügt hier nicht mehr den heutigen Anforderungen, da bis auf einen alle aktiven Links in einen blockierten Status versetzt werden und man eine Vollvermaschung unter Ausnutzung aller aktiven Links somit nicht realisieren kann. Um diesen Mangel zu beheben, wurden von den Gremien zur Standardisierung und den einzelnen Herstellern neue Protokolle und Technologien entwickelt.

Zunächst wären hier das TRILL-Protokoll (Transparent Interconnection of Lots of Links) der IETF und die SPB-Technologie (Shortest Path Bridging) der IEEE zu nennen. Beide Verfahren sind in der Grundidee ziemlich ähnlich. Über ein Link-State-Protokoll (in diesem Fall IS-IS) bauen die beteiligten Brücken respektive Switches eine Datenbank über alle verfügbaren Wege auf. Somit können sie im Einzelfall entscheiden, welchen Weg ein Paket nutzen soll. Dabei sind alle beteiligten Links gleichzeitig aktiv, ohne dass es zu einer Schleifenbildung kommt. Beide Verfahren sind mittlerweile standardisiert, jedoch sind die Marktakzeptanz und die Unterstützung durch die Hersteller weit unter den Erwartungen geblieben; die ersten Hersteller haben sich bereits bewusst von beiden Technologien abgewandt.

Einige Hersteller propagieren seit geraumer Zeit den Einsatz sogenannter Fabrics. Dabei werden alle beteiligten Switches zu einer einzigen logischen Einheit kombiniert und stellen sich nach außen wie ein einziger

Switch dar. Das klingt zunächst einmal verlockend: Das gesamte Netz wird über eine einzige logische Instanz verwaltet – das Marketing arbeitet an dieser Stelle gerne mit dem Begriff „Single Hop Network“. Dies ist jedoch mit Vorsicht zu genießen. Zunächst einmal handelt es sich bei allen verfügbaren Fabric-Lösungen ausschließlich um rein proprietäre Technologien. Der Betreiber legt sich also auf einen einzigen Hersteller fest. Da sämtliche Switches im Rechenzentrum von diesem einen Hersteller sein müssen, ist der Austausch einzelner Switches oder ganzer Ebenen des Netzes durch Produkte eines anderen Herstellers unmöglich.

Abgesehen davon ist der Begriff „Single Hop Network“ reines Marketing. Auch die Fabric-Netze arbeiten mit verschiedenen Ebenen im Netz. Wenn ein Paket zum Beispiel von einem ToR-Switch (Top of Rack) zu einem anderen geschickt werden soll, geht dies ebenfalls über mehrere Wege. Der Single Hop ist also rein virtuell und hat keinen positiven Einfluss auf die Überbuchung sowie die Verzögerungszeiten im Netz. Außerdem gestaltet sich die Fehlersuche in einem Fabric-Netz extrem schwierig, da nahezu keine Möglichkeit eines manuellen Eingriffs in das Verhalten der Fabric besteht und die Verkehrsströme nicht mehr klar definierten Wegen und Regeln folgen. Grundsätzlich ist von der Verwendung solch proprietärer Lösungen abzuraten.

Realisierung mit virtuellem Overlay

Aus den Reihen der Virtualisierungshersteller gibt es Lösungen, mit denen man Layer-2-Verbindungen durch ein virtuelles Overlay-Netz realisieren kann. Diese Lösungen sind in der Regel gegenüber dem darunter liegenden Netzwerk (Underlay) komplett agnostisch. Das ist der Tatsache geschuldet, dass der Hersteller eines Hypervisors seine Lösung auf jeder Art von Netzinfrastruktur betreiben können muss und sich insofern sein eigenes „virtuelles Netz“ baut. VXLAN ist hier als prominentester Vertreter solcher Technologien zu nennen.

Overlay-Netze stellen in sehr großen Umgebungen ein probates Mittel dar. Insbesondere die Layer-2-Verbindungen mehrerer Standorte lassen sich so unabhängig von der Protokollstruktur umsetzen. Insbesondere in Verbindung mit Leaf-Spine-Architekturen als Underlay können Overlay-Netze in Zukunft die Notwendigkeit von sehr großen Core-Systemen im Data Center beseitigen und zu einer kostengünstigen Scale-out-Lösung werden. Doch heutzutage sind diese Architekturen noch in der Minderheit. Das liegt zum einen an der Komplexität der Underlay-Netze in großen Umgebungen. Man muss hier zum Beispiel auf OSPF basierende Layer-3-Strukturen mit ECMP (Equal-cost Multi-path Routing) zurückgreifen. In absehbarer Zeit wird es für solche Architekturen jedoch vereinfachte Modelle geben, die quasi eine Plug-and-play-Administration ermöglichen werden.

MLAG plus Stacking

Als optimale Umsetzung von Layer-2-Strukturen im Rechenzentrum hat sich momentan eine Kombination von zwei Technologien erwiesen: MLAG (Multi-Switch Link Aggregation Group) und das Stacking mehrerer ToR-Switches. Beim Stacking werden mehrere (meistens bis zu acht) einzelne Switches zu einer logischen Einheit verbunden. Im Gegensatz zur Fabric erfolgt dies jedoch über einen dedizierten Bus mit zum Beispiel 160 GBit Bandbreite. Dieses Verfahren hat sich gerade für die Verbindung von ToR-Switches bewährt, da es zum einen die Anzahl der zu verwaltenden Switches reduziert und zum anderen den Ost-West-Verkehr zwischen benachbarten Schränken über den internen Bus und nicht über den Core abwickelt.

Die einzelnen Stacks werden dann über MLAGs mit dem Core verbunden. Eine MLAG ist dabei prinzipiell nichts anderes als eine klassische Link Aggregation Group (LAG) nach IEEE 802.1AX (früher 802.3ad). Es werden also mehrere physikalische Links zu einem logischen Link zusammengefasst, um die Gesamtbandbreite zu erhöhen. Bei MLAG werden diese logischen Links im Core jedoch auf zwei verschiedene Systeme verteilt. Nach außen erfolgt dies transparent. Der angeschlossene Server oder ToR-Switch weiß also nicht, dass er mit zwei verschiedenen Systemen verbunden ist. Für ihn handelt es sich um eine klassische Link Aggregation Group. Somit ist MLAG zwar streng genommen ebenfalls eine proprietäre Lösung, da die beiden Core-Systeme vom selben Hersteller sein müssen. Allerdings betrifft dies nur die Core-Ebene und ist nach außen nicht ersichtlich. Die ToR-Switches können also von jedem beliebigen Hersteller sein.

POD-Verbund im Praxisbeispiel

Die Voraussetzungen und technologischen Grundlagen einer Cloud-Infrastruktur lassen sich am besten an einem Beispiel aus der Praxis zeigen, das tatsächlich so für einen großen Cloud-Hosting-Provider entworfen wurde: Da das Netzwerkdesign hochskalierbar sein sollte, wurde es modular aufgebaut. Die kleinste Einheit besteht aus einem POD (Portable Optimized Data Center), wobei ein einzelnes POD aus zwei Schränken besteht. In jedem Schrank befinden sich 20 Server und etwa 100 TByte Storage. Außerdem ist in jedem Schrank ein ToR-Switch mit 48 Ports für 10GbE und vier Ports für 40GbE als Uplink untergebracht. Die Server und die Storage-Systeme sind mit MLAGs an beide ToR-Switches angebunden.

Wenn ein zweites POD hinzukommt, bilden jeweils die linken und rechten ToR-Switches einen Stack, also eine logische Einheit. Beide PODs haben damit weiterhin nur zwei logische ToR-Switches. Darüber hinaus bilden vier PODs eine Data-Center-Zone. Auch innerhalb der Zone sind die jeweils linken und rechten ToR-Switches zu einem Stack zusammengeschaltet. Deshalb hat auch die Zone nur zwei logische ToR-Switches, und die Stacking-Bandbreite zwischen den einzelnen ToR-Switches beträgt 160 GBit/s. Dieses Design hat den Vorteil, dass der Ost-West-Verkehr einer Zone diese nicht verlassen muss. Wenn die Anwender also die Storage-Virtualisierung auf eine einzige Zone beschränken, muss der Ost-West-Verkehr nicht über den Core fließen. Das wirkt sich positiv auf die Überbuchung der Uplinks zum Core aus.

Da es bei den heute verfügbaren Port-Dichten der Core-Switches in den allermeisten Fällen vollkommen ausreicht, mit einer oder zwei Netzwerkebenen zu planen, kommt den Themen End of Row oder Distributionsebene hier keine größere Bedeutung zu. So verfügen die modernen Core-Switches beispielsweise bis zu 768 Ports in 10GbE oder bis zu 192 Ports in 40 GbE auf etwa 14 Höheneinheiten. Legt man diese Port-Dichten zugrunde, kann man in einem solchen Design bis zu 22 Zonen mit jeweils 320 GBit/s mit dem Core verbinden. Die Anbindung erfolgt wiederum über MLAGs. Von jedem ToR-Switch geht ein Link mit 40 GBit/s zum Core, und alle acht Links sind zu einer logischen Verbindung zusammengefasst. Das bedeutet in diesem Fall 176 Schränke mit insgesamt 3520 Servern in einem rein zweistufigen Netz. Erst darüber hinaus könnte man über die Einführung einer weiteren Ebene nachdenken.

Das beschriebene Design zeigt deutlich, dass es möglich ist, mit heute verfügbaren Technologien ein hochskalierbares Design für Cloud-Umgebungen in Rechenzentren aufzubauen, ohne dass man auf proprietäre Lösungen zurückgreifen müsste. Die skizzierte IT-Umgebung skaliert von 40 Servern bis zu 3520 Servern und ist auch dann noch erweiterbar. Insofern ist das Design sicherlich auch für klassische Unternehmensnetze interessant.

Olaf Hagemann,
SE Director DACH, Extreme Networks

Die Physik setzt auch der Cloud Grenzen

Die Hardware hinter der Wolke muss zur Arbeitsweise der Anwendung passen

Cloud-Speicher sind nichts anderes als abstrahierte IT-Infrastrukturen, deren Grundlage aber immer noch handfeste Hardware ist. Daher gelten für sie auch die physikalischen Grundsätze der Informatik. Aufschluss über die Leistungsfähigkeit eines Cloud-Speichers liefern vor allem Bandbreite und Latenz.

Das Markforschungsunternehmen Osterman hat 2016 untersucht, nach welchen Gesichtspunkten sich Unternehmen in einer Hybrid-Cloud-Infrastruktur für eine Public oder eine Private Cloud entscheiden. Eine der wichtigsten Erkenntnisse: „Keine Diskussion über die Cloud ist vollständig, wenn die hardwareseitigen Grenzen außen vor bleiben." Weiter heißt es: „Wird ein Workload von der Private Cloud auf die Public Cloud ausgelagert, kann das zu Latenzen führen, die nicht akzeptabel sind. Im Hinblick auf reaktionskritische Anwendungen kann daher die Entscheidung für oder gegen einen Cloud-Service-Provider schwierig werden. Auch ist der Datentransfer zwischen Benutzer und Public Cloud oft langsamer als innerhalb eines Rechenzentrums."

Physikalische Datenträgheit

Bereits 2010 hatte Dave McCrory, CTO bei Basho Technologies, dieses Phänomen als „Data Gravity" beschrieben. Damit ist gemeint, dass Daten – je nach Eigenschaft und Sicherheitsklasse – eine gewisse Trägheit besitzen. Sie werden desto beweglicher, je näher sie an den Applikationen oder Services dran sind. Je weiter sie sich davon entfernen, desto träger werden sie – das trifft vor allem auf die Public Cloud zu, bei der die Daten oft auf viele Server an unterschiedlichen Standorten verteilt sind. Die Trägheit der Daten hat damit im Grunde eine physikalische Ursache – sie heißt Latenz. Viele Faktoren wirken sich direkt auf die Latenz der Daten aus – etwa die Länge der Netzwerkleitungen und die Anzahl der Knotenpunkte.

Latenz und Bandbreite sind feste Größen der Rechnerarchitektur und unter anderem für die Leistung eines Systems verantwortlich. Man darf sie aber nicht verwechseln! Andrew Tanenbaum, der Informatiker und Schöpfer des Betriebssystems Minix, verdeutlicht das an einem extremen Beispiel: „Unterschätze niemals die Datenübertragungsrate eines mit Bändern vollgeladenen Kombis, der über die Autobahn rast." Um im Bild zu bleiben: Ginge es nur um die Bandbreite, könnten viele Cloud- und Backup-Provider genauso gut die Datenmassen auf Festplatten speichern und dann per Post verschicken.

Bei der gesamten Diskussion rund um die Cloud wird immer die Latenz vergessen. Sie ist die Zeit, die zwischen Anfrage und Auslieferung der Daten verstreicht. Die Latenz ist die Summe aller Verzögerungen, die jede Komponente bei der Verarbeitung einer Anfrage hinzufügt. Da die Latenz für jedes einzelne Datenpaket gilt, das durch das System wandert, ist sie mindestens genauso wichtig wie die oft überschätzte Bandbreite. Bandbreite bedeutet im Grunde nur, dass die Daten auf einer breiten Autobahn dahinfließen und nicht auf einer holprigen Landstraße. Auf die Latenz übertragen heißt das: Werden die Daten in einem alten Pickup oder in einem Rennwagen transportiert?

Latenz je nach Kontext

Schon die Latenzen zweier Systeme zu vergleichen, ist nicht ganz einfach, weil die Werte stark von der konkreten Konfiguration und von dynamischen Variablen wie der Systemlast abhängen. Darüber hinaus verwenden Anbieter zum Teil unterschiedliche Definitionen, ohne dass man als Kunde im Einzelfall wüsste, auf welches Szenario sich die Angaben beziehen.

Im modernen Cloud-Kontext ist die Bedeutung, die der Latenz zukommt, jedenfalls sehr unterschiedlich. Einerseits greifen moderne Applikationen vorwiegend lesend auf die Datenbestände zu – Schreibprozesse finden hingegen vergleichsweise selten statt; außerdem arbeiten diese modernen Applikationen vorwiegend parallel und kommen auch mit älteren Datenzuständen zurecht. Das heißt: Die Bedeutung der Latenz nimmt in diesem Zusammenhang ab, und es ist dadurch recht unwahrscheinlich, dass sie zu einem Flaschenhals wird.

Andererseits gibt es aber auch viele Applikationen, für die Latenzen essenziell sind. Durchgehende Latenzanforderungen – wenn beispielsweise Berechnungen oder andere Operationen linear abgearbeitet werden müssen – können voraussetzen, dass alle Daten in ein und demselben Rechenzentrum gespeichert sein müssen; das kann auch bei einem Public-Cloud-Provider sein. Noch höhere Anforderungen können dazu führen, dass bei Private und Public Clouds die Hardware mit SSDs oder Speicherpools optimiert werden muss. Reicht das immer noch nicht aus, können in bestimmten Situationen spezielle Netzwerke und Workload-Verteilungen in einer On-premise-Umgebung erforderlich sein.

Faustregel auf Bewährung

Die pauschale Aussage also, dass die typische Anwendung von heute toleranter gegenüber reduzierter Bandbreite und höherer Latenz ist als eine typische Anwendung, die vor 15 Jahren auf einem großen Unix-Server lief, kann daher immer nur eine Verallgemeinerung und keine feste Regel sein. Es zeigt sich: Die Cloud ist im Grunde nur eine Abstraktion. Dahinter steckt immer reale Hardware, die an physikalische Gesetze gebunden ist. Daher spielen auch für die Cloud Bandbreiten und Latenzen entscheidende Rollen.

Gordon Haff,
Senior Cloud Product Manager, Red Hat

Eine Sorge weniger

Unternehmen mit eigenem RZ stellen auf externe Wartungsdienste um

Gute Admins sind weder einfach zu finden noch billig. Kostendruck und Personalknappheit haben dazu geführt, dass mehr und mehr Rechenzentren Aufgaben wie Monitoring, Verfügbarkeitskontrolle und Security an Managed Services auslagern. Das erfordert jedoch klare Vereinbarungen, auch für Notfälle.

Auch bei nur geringer oder mittlerer Auslastung übersteigen allein die elektrischen Betriebskosten eines Rechenzentrums oft die Investitionskosten in deutlich weniger als zehn Jahren. Hinzu kommen auf der Ausgabenseite im laufenden RZ-Betrieb aufeinander abgestimmte Instandhaltungsmaßnahmen sowie verschiedene Management-Disziplinen, die ebenfalls einen beträchtlichen Aufwand darstellen. Sie erfordern ein zunehmend hohes Maß an Erfahrung und Fachkompetenz und binden wertvolle, meist knappe personelle Ressourcen.

Nicht umsonst suchen RZ-Betreiber händeringend nach Möglichkeiten, solche Kosten einzugrenzen oder besser noch ganz außen vor zu lassen. Dabei greifen sie immer häufiger auf Outsourcing-Strategien und Managed Services zurück. Partner, die hier ansetzen, sollten ihre Lösungen auf die Anforderungen und Eigenschaften individueller Kundenprojekte anpassen und das Data Center als ganzheitlichen Service betrachten.

Vor-Ort-Service für Data Center

Im englischen Sprachraum ist DCaaS (Data Center as a Service) oft gleichbedeutend mit Serverhousing bzw. Collocation; hier ist aber das Gegenteil gemeint: ein unternehmensspezifisches Dienstleistungskonzept für das eigene Rechenzentrum des Unternehmens vor Ort. Die Basis bildet das klassische Outsourcing, bei dem alle Aufgaben, die keinen unternehmerischen Mehrwert bilden, an Servicepartner abgegeben werden. Das setzt rund um das Data Center viel Vertrauen voraus und bedeutet, dass die Partner drei wichtige Elemente immer im Fokus behalten: Verfügbarkeit, Sicherheit und Wirtschaftlichkeit.

Wichtig ist zunächst die vereinbarte Verfügbarkeit, die sich in der Ausfallsicherheit abbildet. Nur ein individuelles Betriebs- und Störfallmanagement kann hier den Verfügbarkeitsansprüchen gerecht werden. Damit geplante oder ungeplante Ausfälle von Infrastrukturkomponenten, also Wartungen und Störungen, nicht die Redundanzen gefährlich verringern oder sogar zum Totalausfall führen, sind proaktive Maßnahmen notwendig. Dazu gehören unbedingt vorausschauende Instandhaltungsstrategien, die Downtimes von vornherein verhindern. Kommt es dennoch zu einer Beeinträchtigung des RZ-Betriebs, müssen klare Regelungen getroffen sein, um die Störung so schnell wie möglich zu beseitigen und den normalen Zustand wiederherzustellen.

SLAs (Service Level Agreements) definieren solche besonderen Bedingungen und sichern den Betrieb zum Beispiel nach einer festgelegten Zeit. Dazu ist es unabdingbar, den genauen Bestand aller Infrastrukturkomponenten und die Konfigurationen zu kennen und regelmäßig zu aktualisieren. Haben die Partner dann noch die Kapazität und Leistungsreserven im Blick, sind schon die wichtigsten Voraussetzungen geschaffen.

Ein sicheres Rechenzentrum bedeutet, dass es frei von unvertretbaren Risiken und Gefahren ist. Das setzt aber deren Kenntnis voraus. Darum müssen die Sicherheitsanforderungen aus der Planungsphase regelmäßig auf die des laufenden RZ-Betriebs abgestimmt werden, und zwar aus ganz unterschiedlichen Perspektiven: Die physische Sicherheit bezieht sich auf interne und externe Einflüsse, die Lage des Gebäudes sowie auf das Verhalten und die Organisation im Rechenzentrum. Die betriebliche Sicherheit dagegen legt den Fokus auf die gesetzlichen und regulatorischen Anforderungen. Sie unterstützt beispielsweise bei (Re-)Zertifizierungen und kümmert sich um Betriebsführungs- und Notfallhandbuch.

Damit der RZ-Betrieb wirtschaftlich ist, muss zudem der energetische Zustand der IT-Systeme sowie der Infrastrukturkomponenten so transparent wie möglich sein. In der Praxis heißt das: Es braucht immer ein übergeordnetes Monitoring für diese beiden Bereiche. Neben der Verfügbarkeitskontrolle können so auch wichtige Faktoren wie Energieflüsse, Temperaturen oder Drücke visualisiert werden.

Zu einem guten Monitoring gehört ferner eine professionelle Auswertung der erhobenen Daten. Hier kommt allein aus Personalgründen am besten ein Servicepartner zum Einsatz, der die Daten interpretiert, energetische Kennzahlen bildet und die Betriebsweise der Komponenten überwacht und auswertet. Sinnvollerweise kann er bei schlechten Kennzahlen oder unguter Betriebsweise geeignete Optimierungsvorschläge unterbreiten. Im Idealfall realisiert der Servicepartner auch die praktische Umsetzung und hält damit ein kontinuierlich hohes energetisches Niveau im Betrieb.

Planung und Implementierung

Grundsätzlich sollte DCaaS eine sehr individuell gestaltete Maßnahme sein, damit der RZ-Betrieb in allen Bereichen maßgeschneidert passt. Hier ist eine relativ enge Partnerschaft zwischen RZ-Betreiber und Servicedienstleister notwendig, die mit klaren Vereinbarungen am besten funktioniert. Beide Seiten sollten sich im Vorfeld detailliert Gedanken machen, welche Dienstleistungen tatsächlich notwendig sind und wer sie liefert.

Beim Vorgehen haben sich sogenannte Integrationsmodule bewährt. Sie bestehen aus einer anfänglichen Strategiephase, die Kundenbedürfnisse und Erwartungen definiert, reichen über eine Entwicklungsphase, die konkrete Anforderungen und Lösungen liefert, und enden in einer Implementierungsphase, welche die vereinbarten Leistungen einführt und in standardisierte Prozesse übergeht.

In der längsten Phase, im Rechenzentrumsbetrieb, müssen dann die vereinbarten Leistungen gelebt und kontinuierlich verbessert werden. Hierzu ist wieder eine sehr enge Zusammenarbeit der Partner gefragt, die in regelmäßigen Abständen die Qualität der Dienstleistung überwacht und mit Blick auf die Zukunft die derzeitigen Anforderungen an den RZ-Betrieb auf künftige Entwicklungen abstimmt.

Frank Neubauer,
Business Development Manager, RZ Services GmbH

Platz sparen, umstecken, erweitern

Wie ausbaufähig ein Rechenzentrum ist, zeigt sich auf den Datenstrecken

Die passive Infrastruktur steht bei den Planern von Rechenzentren meist sehr weit unten auf der To-do-Liste – zu Unrecht. Denn gerade hier könnten sie die Weichen für hochverfügbare, skalierbare und investitionssichere Netzwerke stellen. Modulare Systeme sind flexibel einsetzbar und leicht zu erweitern.

Die Praxis zeigt immer wieder, dass Unternehmen, die in ein Rechenzentrum investieren, der passiven Infrastruktur zu wenig Beachtung schenken. Das liegt zum Teil daran, dass RZ-Betreiber nicht ausreichend informiert sind, was in diesem Bereich überhaupt möglich ist und welche Technologien der Markt bietet. Außerdem stehen in der Planungs- und Projektphase vor allem die aktiven Komponenten im Fokus; an die Kabel wird erst zuletzt gedacht. Ein weiterer Grund: Die Investitionen in die passive Infrastruktur bewegen sich im unteren einstelligen Prozentbereich. Gemessen an der gesamten Projektsumme fallen die Einnahmen aus der Verkabelung somit eher gering aus. Entsprechend gering ist dann auch ihr Stellenwert. Das geht manchmal soweit, dass Netzwerkplaner auf alte Ausschreibungstests zurückgreifen – mit der Folge, dass die Planungen nicht mehr der neuesten Technologie entsprechen.

Die Anbieter von Netzwerkkomponenten haben noch viel Pionierarbeit zu leisten, Unternehmen davon zu überzeugen, wie wichtig eine gut durchdachte passive Infrastruktur und deren frühzeitige Planung sind. Dann allerdings wird auch rasch klar, dass die Qualität der Komponenten die Performance und Verlässlichkeit des gesamten Rechenzentrums entscheidend mitbestimmt. Für die Umsetzung stehen am Markt verschiedene Produkte und Systeme bereit. Generell wird man angesichts der rasant steigenden Anforderungen an die Datennetze eher zu Verkabelungslösungen greifen, die eine hohe Packungsdichte, Modularität und Skalierbarkeit aufweisen und daher auch flexibel einsetzbar sind.

Weniger Platz, weniger Kosten

Kompakte Verkabelungssysteme sparen nachweislich Kosten: Rechenzentren lassen sich insgesamt kleiner auslegen, wenn das Kabelvolumen gering und die Packungsdichte in den Netzwerkschränken hoch ist. Zugleich sinkt der erforderliche Klimatisierungsaufwand. Ein sinnvoller Indikator ist hier die „Packungseffizienz“. Besonders praktisch sind Verkabelungssysteme, die nicht gleich von Anfang an mit Vollbestückung und höchster Packungsdichte installiert werden müssen, sondern eine Nachbestückung je nach Bedarf ermöglichen. Es gibt bereits Systeme, deren Leistung sich auf diese Weise verdoppelt lässt.

Dabei spielt vor dem Hintergrund steigender Übertragungsraten die Wahl der Kabelart eine zentrale Rolle: Glasfaser bietet auf unterschiedlichen Ebenen Vorteile und ist die erste Wahl, wenn Übertragungsraten von 40GbE, 100GbE oder mehr erreicht werden sollen. Vor allem wird für die Datenübertragung über LWL-Kabel mit zunehmender Übertragungsgeschwindigkeit erheblich weniger Strom als beispielsweise über Kupferkabel benötigt. So sind bei 10GbE über Kupfer bereits 10 W nötig, damit das ankommende Nutzsignal nicht im Rauschen untergeht. Demgegenüber fallen bei Übertragungen von 10GbE über LWL lediglich 2 W an. LWL-Kabel sind zudem frei von Störeffekten wie Übersprechen oder dem Skin-Effekt, bei dem die äußeren Bereiche der Kupferkabel überbelastet werden.

Quelle: Berthold Steinhilber – MPI für biologische Kybernetik

Forschungsarbeit im Panolab des MPI für biologische Kybernetik

Da LWL-Kabel außerdem kleine Durchmesser haben, ist auch ihr Volumen geringer. Das hat zwei Vorteile: Die Belüftungswege in den Netzwerkschränken sind weniger blockiert, was Energieeinsparungen bedeutet, und zu-

gleich reduzieren sich die Brandlasten in Rechenzentren, je kleiner die Kabeldurchmesser sind.

Passende Module kombinieren

Eine vorausschauende Konzeption von Rechenzentren setzt heute stärker auf eine strukturierte Verkabelung. Deshalb geht der Trend von klassischen Einzelverkabelungen hin zu vorgefertigten, modularen Systemen. Sie sind flexibler, geben mehr Investitionssicherheit und helfen, Ressourcen einzusparen und die Umwelt zu schonen. Vorkonfektionierte und getestete Systemkomponenten lassen sie sich per Plug-and-play innerhalb kürzester Zeit installieren. Netzwerktechniker können danach jederzeit Änderungen vornehmen und dieselben Komponenten wiederverwenden.

Welchen Unterschied eine kompakte Verkabelung in der Praxis macht, hat sich zuletzt am Max-Planck-Campus Tübingen gezeigt. Das bisherige Rechenzentrum der Forschungseinrichtung hatte seine Kapazitätsgrenze erreicht und war zu klein geworden. Um alle Server und Speichersysteme der Institute unterzubringen, musste die IT-Abteilung bereits weitere Räumlichkeiten auf dem Campus nutzen. „Das war natürlich kein Dauerzustand – es musste etwas passieren“, schildert Stefan Tauber, EDV-Mitarbeiter am Max-Planck-Institut für Entwicklungsbiologie, die damalige Situation. „Weitere Punkte waren vor allem die immer wieder auftretenden Probleme durch uneinheitliche Kabelspezifikationen und die große Hitze im Sommer, bedingt durch das Flachdach.“ Aufgrund der besonderen Raumsituation standen Packungsdichte, Modularität und Skalierbarkeit im Fokus dieser Neuverkabelung.

Max-Planck-Campus Tübingen

Nachdem alle Entscheider zugestimmt hatten, zog die IT-Abteilung eine Etage tiefer ins Untergeschoss des bisherigen Rechenzentrums. Dort war genügend Platz, um alle bisher verteilten Server zu vereinen und auch auf künftiges Wachstum zu planen. Im alten erdgeschossigen Rechenzentrum liefen jedoch alle Netzwerkverbindungen des gesamten Campus zusammen. „Diese Verkabelung konnten wir nicht ohne Weiteres in das Untergeschoss verlegen. Der Aufwand wäre zu groß gewesen und hätte zu viele Unterbrechungen nach sich gezogen“, erinnert sich Stefan Tauber. Das IT-Team beschloss daher, die für die Anbindung nach draußen nötigen Netzwerkkomponenten im Erdgeschoss zu lassen. Da in erster Linie die passiven Komponenten betroffen waren, würde die Kühlung kein Problem sein.

„Für die Anbindung in den neuen Serverraum im Untergeschoss war jedoch eine zukunftssichere Lösung notwendig. Dies musste ein modulares System sein“, sagt Stefan Tauber. „Unser Anspruch war, nicht so viel Platz im Rack zu verschenken und für künftige Ansprüche flexibel zu sein.“ Den Zuschlag bekam die tde – trans data elektronik GmbH als einer der wenigen Netzwerkanbieter mit modularen Systemen im Portfolio. Sie lieferte die vorkonfektionierten Kabel und alle weiteren Netzwerkkomponenten exakt nach den Vorgaben der Elektroplaner.

Quelle: tde

Das modular aufgebaute Verkabelungssystem tML besteht aus den drei Kernkomponenten Modul, Trunk-Kabel und Modulträger. Es erlaubt eine extrem einfache und schnelle Migration auf 40GbE, 100GbE und höhere Übertragungsraten.

Patch-Feld für Kupfer und LWL

Um Hochverfügbarkeit zu garantieren und die früheren Kabelprobleme, die von uneinheitlichen Spezifikationen herrührten, in den Griff zu bekommen, wurden im Vorfeld alle Kabelspezifikationen genau festgelegt und die Komponenten durch einen herstellerzertifizierten Installateur eingebaut. Im ersten Raum des neuen Rechenzentrums stehen jetzt zwei Rack-Reihen mit je sechs Serverschränken. Im daran angrenzenden zweiten Raum bilden vier Schränke das Cluster (der Max-Planck-Campus hatte es zuvor ausgelagert und es im Zuge der Neukonzeption ebenfalls in das neue Rechenzentrum integriert). Die Racks sind mit dem modularen tML-System bestückt, das eine optimale Packungsdichte mitbringt und die Integration von Glasfaser- und Kupferkabeln in einem Panel innerhalb einer Höheneinheit ermöglicht.

Am Ende des Projekts war die Resonanz der Max-Planck-Institute Tübingen durchwegs positiv. Auch für die Zukunft muss sich die Forschungseinrichtung keine Sorgen machen. Die neue, hochskalierbare Verkabelungslösung kann mit steigenden Anforderungen Schritt halten und im Fall der Fälle schnell und einfach auf 40GbE migrieren.

André Engel,
Geschäftsführer tde – trans data elektronik

Quelle: MPI Tübingen

Server-Racks im Max-Planck-Institut Tübingen

Viele Regeln, wenig Neues

Ab 25. Mai 2018 gilt das gemeinsame europäische Datenschutzrecht

Die EU-Datenschutz-Grundverordnung soll die Digitalisierung nicht bremsen, sondern eine Grundlage dafür schaffen. Das deutsche Recht ist zwar schon relativ nahe an den neuen Regelungen, aber auch hier muss man Forderungen wie der nach einem lückenlosen Herkunftsnachweis von Daten erst gerecht werden.

Nach jahrelangem Gezerre wurde im Mai 2016 die EU-Datenschutz-Grundverordnung (EU-DSGV) verabschiedet, die das europäische Datenschutzrecht vereinheitlicht und damit für sichere Geschäftsgrundlagen auf diesem Gebiet in allen europäischen Ländern sorgt. Nun haben deutsche Unternehmen und Rechenzentrumsbetreiber zwei Jahre Zeit, die Regelungen umzusetzen. Dies dürfte deutschen Institutionen vielfach leichter fallen, da sie bereits die relativ strengen Regeln des deutschen Datenschutzrechts gewohnt sind. Allerdings konnte sich der „deutsche Standard“ nicht überall durchsetzen. So wurde die Größe der Unternehmen, die nun einen Datenschutzbeauftragten brauchen, erheblich heraufgesetzt – was jedoch nicht bedeutet, dass deutsche Firmen, die nach dem bis Juni 2018 geltenden Recht einen Datenschutzbeauftragten bestellt haben, diesen abschaffen sollten.

Geltungsbereich

Die EU-DSGV gilt letztlich für jeden, der europäischen Kunden Güter oder Dienste über das Internet anbietet, und zwar auch dann, wenn in Europa gar keine Niederlassung oder Vertretung besteht. Solche Unternehmen müssen für die europäischen Behörden einen Ansprechpartner benennen. Daneben gilt die EU-DSGV natürlich für alle Unternehmen und Institutionen, die in Europa aktiv sind, und sei es auch nur über einen einzigen Vertreter oder Handelspartner.

Einbezogen in die Regulierung sind folgende Daten: Rasse/ethnische Zugehörigkeit, politische Ansichten, Religionsangehörigkeit oder Weltanschauung, Mitgliedschaft in Gewerkschaften, Daten zu Gesundheit, sexueller Identität oder Orientierung. Für diese Daten gelten auch jetzt schon nach der EU-Datenschutzrichtlinie spezielle Regeln. Neu hinzu kommen in der EU-DSGV biometrische Daten, soweit sie die Identifikation einer Person erlauben, und genetische Daten. Videos und Fotos sind dann erfasst, wenn sie die Identifikation von Personen erlauben. Zusätzliche Regeln in Mitgliedsländern sind möglich.

Einwilligung

Hier orientiert sich die EU-Verordnung stark am deutschen Recht und ähnlichen Rechtsordnungen: Ab 2018 ist zum Sammeln und verarbeiten von Daten eine individuelle, aktive Einwilligung gefordert. Sowohl die gesammelten Daten als auch die jeweils geplanten Verarbeitungszwecke sind dabei vollständig anzugeben. Eine allgemeine Zustimmung reicht nicht aus, genauso wenig wie ein irgendwo verstecktes Kästchen, das man anklicken soll, wenn man mit der Datenverarbeitung nicht einverstanden ist. Anwender müssen ihre Einwilligung außerdem genauso einfach und klar widerrufen können, wie sie sie erteilt haben, und müssen über dieses Recht auch informiert werden. Kommen später weitere Einsatzzwecke für die Daten in den Blick, ist eine weitere Einwilligung für den neuen Zweck nötig.

Wer also Daten zunächst nur für die eigene Statistik erhebt und dafür eine Einwilligung erhält, muss eine weitere Einwilligung einholen, wenn später geplant ist, die Daten zu verkaufen, und auch dann, wenn vorgesehen ist, weitere Daten in Erfassung und Analyse einzubeziehen. Auch diese Einwilligung kann verweigert oder zurückgezogen werden.

Wird eine Einwilligung zurückgezogen, gilt dies nicht nur für den, der die Einwilligung zuvor eingeholt hat, sondern auch für alle, die die entsprechenden Daten später von dieser Institution erworben haben und so weiter. Im Datenhandel braucht man also zukünftig das, was im Lebensmittelrecht über Jahrzehnte mühsam durchgesetzt wurde: einen lückenlosen Herkunftsnachweis der Daten, die in irgendwelche Analysen einfließen. Denn anders lässt sich im Ernstfall das europäische Recht nicht einhalten. Man darf gespannt sein, wie sich die Regeln auf so manches abenteuerliche Geschäftsmodell im Bereich Datenservices auswirken. Gerade diese Bestimmungen könnten angesichts der vielen neuen Technologien zur Datenanalyse und regem Datenhandel den Gerichten Arbeit bescheren.

Überhaupt nicht online einwilligungsfähig werden Kinder unter 13 sein. Hier ist auf jeden Fall die elterliche Einwilligung nötig. Zwischen 13 und 15 Jahren darf jeder Mitgliedsstaat gegebenenfalls eigene Regeln festlegen, ab dem vollendeten 16. Lebensjahr entscheidet der oder die Jugendliche generell allein.

Recht auf Löschung

Neu in der europäischen Rechtsordnung ist das Recht auf Löschung, das lange als „Recht auf Vergessenwerden“ durch die Gazetten und Debatten geisterte. Dieses Recht greift, sobald ein Individuum seine Zustimmung zur Datenverarbeitung widerruft oder die Verarbeitung nicht nach den festgelegten Regeln erfolgt ist. Im Fall von Rechtsunsicherheit können Individuen verlangen, dass ihre Daten nur eingeschränkt verwertet werden. Außerdem müssen Daten auch gelöscht werden, wenn sie den ursprünglich vorgesehenen Zweck erfüllt haben. Wer also Daten nur für eine einmalige Marketing-Aktion erhoben hat, muss sie anschließend löschen und darf sie nicht ohne zusätzliche Einwilligung in die dauerhafte Unternehmensstatistik eingliedern.

Die Löschpflicht gilt auch dann, wenn Daten weiterverkauft oder veröffentlicht wurden. Dann muss derjenige, der die Daten erhoben hat, alle nachgelagerten Nutzer über die Löschpflicht informieren. Mitgliedsstaaten können allerdings Ausnahmeregeln erlassen. Liegen bestimmte legitime Interessen vor (amtliche Statistik, Beweiszwecke im Rechtsverfahren, rechtliche Verpflichtungen etc.), gilt das Löschverlangen nicht.

Verwaltung und Schutzmaßnahmen

Unternehmen sind verpflichtet, Datenschutzmaßnahmen zu ergreifen. Die Pflicht, einen Datenschutzbeauftragten zu ernennen, erstreckt sich laut Verordnung auf öffentliche Einrichtungen, alle Organisationen, die Datensubjekte regelmäßig und systematisch überwachen oder in großem Umfang sensitive Daten oder Strafregister verarbeiten, und Institutionen, die das Sozialrecht dazu verpflichtet. Die Wahrnehmung der Aufgabe durch Externe ist möglich. Das deutsche Recht geht hier erheblich weiter. Weil deutsche Unternehmen ohnehin Mitarbeiter mit den umfangreichen Aufgaben rund um die EU-DSGV betrauen müssen, sollten sie sich gut überlegen, ob sie einen vorhandenen Datenschutzbeauftragten abschaffen. Der Datenschutzbeauftragte wird den zuständigen Behörden gemeldet.

Stellt eine Form der Datenverarbeitung ein hohes Risiko für die persönlichen Rechte und die Freiheit betroffener Personen dar, weil beispielsweise ihre persönlichen Daten ständig automatisch und umfassend bewertet werden, ist eine Datenschutz-Risikoabschätzung – ein Privacy Impact Assessment (PIA) – erforderlich. Dies dürfte bei vielen Big-Data-Geschäftsmodellen zutreffen, zum Beispiel bei Kredit- oder Versicherungs-Scorings, aber auch bei der externen Videoüberwachung zum Beispiel des Bereichs vor der Eingangstür von Geschäften. Eine PIA muss die Verarbeitungsaktivitäten und ihren Zweck beschreiben, die Angemessenheit des Umfangs in Relation zum angestrebten Zweck darstellen, die Risiken, die sich für die Rechte und Freiheiten der erfassten Personen ergeben, sowie die Maßnahmen, um diese Risiken gering zu halten, insbesondere Schutzmaßnahmen für die Personendaten in Übereinstimmung mit der Verordnung. Ein vorhandener Datenschutzverantwortlicher muss hinzugezogen werden. Die PIA wird einer zuständigen Behörde vorgelegt, bevor die Institution mit der Datenverarbeitung beginnt. Diese erklärt ihr Einverständnis oder gibt beratende Hinweise, wo noch Änderungen erforderlich sind und wie diese unter Umständen umgesetzt werden können.

Grundsätzlich muss Datenschutz schon in die Systeme eingebaut werden (Privacy by Design) und sich in den Voreinstellungen der Systeme widerspiegeln (Privacy by Default). Die aktive Zustimmung per Ankreuzen zur Datenverarbeitung ist dafür nur ein spezifisches Beispiel. Dazu gehört auch die sorgfältige Auswahl von Geschäftspartnern und Mitarbeitern. Eine weitere Standardmethode ist die Pseudonymisierung persönlicher Daten. Dabei werden die Daten so verändert, dass der Name als Zuordnungsfaktor eines Datensatzes durch ein anderes, anonymes Merkmal ersetzt wird, zum Beispiel eine laufende Nummer. Die Liste der Zuordnung zwischen Namen und Nummern wird getrennt von den eigentlichen Datensätzen aufbewahrt, sodass im Regelfall nicht aus den Daten auf eine Person geschlossen werden kann. Allerdings lassen sich die beiden Listen wieder zusammenführen, sodass eine Personalisierung der Daten nicht komplett ausgeschlossen ist. Das unterscheidet die Pseudonymisierung von der Anonymisierung, bei der eine Wiederherstellung der persönlichen Zuordnung nicht mehr möglich sein sollte.

Umgang mit Zwischenfällen

Kommt es zu Diebstahl, Verlust, Veränderung, Zerstörung oder unerlaubten Zugriffen auf persönliche Daten, haben Datenverantwortliche und Datenverarbeiter Meldepflichten. Bemerkt ein Verarbeiter einen Zwischenfall, muss er diesen unverzüglich an den Verantwortlichen melden. Dieser wiederum muss den zuständigen Behörden berichten, und zwar ebenfalls unverzüglich, womit maximal 72 Stunden nach dem Zwischenfall gemeint sind, ansonsten ist eine gesonderte Begründung für die Verspätung nötig. Die Behörde kann die Verantwortlichen anweisen, die Personen, deren Daten von dem Zwischenfall betroffen sind, persönlich zu benachrichtigen. Diese Meldung darf nur unterbleiben, wenn die Daten verschlüsselt oder anderweitig so geschützt waren, dass Missbrauch ausgeschlossen ist, wenn keine Risiken für die Datensubjekte zu befürchten sind und wenn eine persönliche Adressierung unverhältnismäßig hohen Aufwand fordern würde, sofern es andere Möglichkeiten der Benachrichtigung gibt. Die Institutionen müssen darüber hinaus eine interne Liste von Datenzwischenfällen unterhalten. Die bisher in der EU geltenden Regeln für TK-Provider gelten wohl weiter.

Die Meldung an die Behörden muss die Art des Zwischenfalls enthalten, die Kategorie und Menge der betroffenen Daten und Dateien, die Kommunikation mit den betroffenen Personen, Name und Kontaktinformationen des Verantwortlichen oder Datenschutzbeauftragten, die vorhersehbaren Konsequenzen des Zwischenfalls sowie die ergriffenen Maßnahmen.

Verstöße und die Folgen

Nach der EU-DSGV sind bei Zuwiderhandlungen gegen festgelegte Regeln empfindliche Strafen möglich. Davon betroffen sind Datenverantwortliche und Datenverarbeiter. Es gibt dabei zwei Level.

Die höchsten Strafzahlungen, nämlich bis zu 20 Millionen Euro oder 4 % des Jahresumsatzes (jeweils das höhere) werden unter anderem fällig, wenn die grundsätzlichen Bedingungen für die Datenverarbeitung einschließlich der Einwilligungsregeln verletzt werden, die Rechte der Datensubjekte verletzt, Daten rechtswidrig in andere Länder transferiert oder Anordnungen der Behörden missachtet werden. Eine geringere Höchststrafe, nämlich bis zu 10 Millionen Euro oder 2 % vom Umsatz, wird verhängt, wenn die Einwilligungsregeln für Kinder missachtet, Datenverarbeiter nicht sorgfältig ausgewählt, Subdatenverarbeiter ohne Zustimmung des Datenverantwortlichen beauftragt und Datenzwischenfälle nicht berichtet werden, wenn kein Datenschutz bei Design/Default eingebaut, keine schriftliche Dokumentation aufbewahrt und nicht mit den Behörden

OHNE EXPORT NACH GROSSBRITANNIEN?

Großbritannien gehört zu den wichtigsten EU-Märkten. Doch ob die EU-Verordnung dort jemals gelten wird, steht derzeit in den Sternen. Es liegt nun an den Ergebnissen des Austrittsprozesses, der aber voraussichtlich erst im Januar 2017 beginnt, wenn die britische Premierministerin May die EU-Austrittsklausel durch ihr Austrittsgesuch aktiviert. Dann könnte Großbritannien frühestens 2019 die EU verlassen – genau ein halbes Jahr nach Inkrafttreten der EU-Verordnung. Die Verordnung könnte im Rahmen bilateraler Handelsverträge auch nach dem Austritt gelten, sofern sie in diese eingeschlossen wird, doch ob das geschieht, weiß heute keiner.

Der aus Unternehmenssicht denkbar umständlichste Fall wäre der, dass sich die EU und Großbritannien nicht über eine Paketlösung hinsichtlich Handel und Freizügigkeit einigen. Das könnte passieren, wenn Großbritannien auf Einschränkungen bei der Freizügigkeit von EU-Bürgern pocht, während die EU Handelsprivilegien nur dann gewähren will, wenn im Gegenzug auch Freizügigkeit besteht. Dann müssten für den Datenschutz ähnliche Vereinbarungen getroffen werden wie mit den USA (Privacy Act), ansonsten wäre der Transfer von Daten nach Großbritannien unter Umständen verboten.

kooperiert wird. Die Summe stellt nur den Maximalwert dar, geringere Strafen sind wie üblich nach Bewertung der Schwere des Zwischenfalles, des Maßes von Verschulden etc. möglich.

Dazu kommen gegebenenfalls Schadenersatzregelungen. Intern haftet der Datenverarbeiter lediglich für den Schaden, der durch die Verletzung seiner spezifischen Pflichten zustande gekommen ist, während der Datenverantwortliche den gesamten Schaden verantwortet, der auf unrechtmäßige Datenverarbeitung zurückgeht. In Hinblick auf die betroffenen Personen haften aber beide (Verantwortlicher und Verarbeiter) jeweils für den kompletten Schaden, damit eventuell betroffene Individuen mit hoher Wahrscheinlichkeit zu ihrem Recht kommen, beispielsweise wenn eine der beiden Parteien insolvent ist. Untereinander können sie dann später Geld voneinander verlangen, wenn das die Verteilung der Verantwortung nahelegt. Das Unternehmen haftet also auch dann gegenüber den Endkunden, wenn Daten verloren gegangen sind, weil der Rechenzentrumsbetreiber sein Data Center nicht ordnungsgemäß gegen unbefugtes Eindringen geschützt hat und deswegen Daten entwendet wurden.

Zertifizierungen und Vereinbarungen

Eine weit größere Rolle als heute sollen in Zukunft Zertifizierungen und bindende Verhaltensregeln spielen. Entsprechende Codes of Conduct können beispielsweise Industrievereinigungen formulieren, müssen sie dann aber der Aufsichtsbehörde zur Prüfung, Registrierung und Zertifizierung vorlegen. Die Einhaltung solcher Verhaltensregeln, die für viele Themen formuliert werden können, spricht für die Einhaltung der entsprechenden Vorschriften der EU-DSGV. Dass die Regeln eingehalten werden, sollen akkreditierte Gremien überwachen. Wer gegen solche Regeln verstößt, kann aus dem Teilnehmerkreis ausgeschlossen werden und müsste im Zweifel (also zum Beispiel bei gerichtlichen Verfahren) jeweils individuell die Einhaltung aller vorgeschriebenen Regeln nachweisen.

DIE WICHTIGSTEN ÄNDERUNGEN AUF EINEN BLICK

Die folgenden Themen wurden gegenüber dem geltenden Recht vieler EU-Länder gravierend geändert und bedürfen daher unbedingt größerer Beachtung; in einigen Fällen allerdings ähneln speziell die in Deutschland geltenden strengen Regeln schon dem nun kodifizierten Recht, beispielsweise bei der Einwilligung. Diese Themen werden hier trotzdem angeführt, zumal Unternehmen häufig in mehreren Ländern Geschäfte machen und unter Umständen in nicht-deutschen Märkten die laxeren Auslandsregeln verwenden:

- Einwilligungsregelungen
- Datenbestände, auf die sich die Regulierungen beziehen
- Meldepflichten bei Zwischenfällen
- Verantwortung für Datenbestände
- Datenschutzkonzepte
- Mehr Rechte für die Kunden/Datensubjekte
- Strafzahlungen und andere Sanktionen

Den Volltext der EU-Datenschutz-Grundverordnung gibt es auf www.datenschutz-grundverordnung.eu.

Gerade bei der Auswahl von Datenverarbeitern, aber auch für Datenverantwortliche selbst, könnten Siegel, Zertifikate und Ähnliches sehr viel wichtiger werden als heute. Der Erwerb soll freiwillig sein, ihr Besitz aber für die Einhaltung der Regeln der Verordnung sprechen. Zertifikate sollen drei Jahre gültig sein und von entsprechend qualifizierten Zertifizierungsgremien vergeben werden. Eine zentrale, öffentlich zugängliche Liste derartiger Qualifikationen wird beim European Data Protection Board geführt.

Neue und alte Gremien

Die meisten Regelungen erfordern kein aktives Handeln von Unternehmen oder Institutionen, man sollte sie aber trotzdem kennen, zum Beispiel um die richtigen Berichts- oder Beschwerdewege einzuhalten.

Die nationalen oder gegebenenfalls regionalen Datenschutzbehörden (wie in Deutschland) bleiben bestehen, sie agieren unabhängig voneinander, müssen aber miteinander und mit der europäischen Ebene kooperieren und die Anwendung der EU-DSGV überwachen. Gibt es mehrere nationale Datenschutzbehörden wie in Deutschland, muss eine dieser Behörden bestimmt werden, die die Schnittstelle zu den EU-Behörden bildet. Konflikte untereinander lösen die Aufsichtsbehörden durch ein in der Verordnung festgelegtes formelles Verfahren. Außerdem sollen die nationalen Behörden ihre Regierungen beraten. Sie sollen die Ausbildung von Zertifizierungssystemen fördern und die Akkreditierungskriterien festlegen und publizieren.

Neben den nationalen gibt es eine übergeordnete EU-weit zuständige Datenschutzbehörde für länderübergreifende Angelegenheiten: das European Data Protection Board (EDPB). Die EU-Kommission ist im EDPB durch ein nicht stimmberechtigtes Kommissionsmitglied vertreten, ebenso alle beteiligten Länder. Das EDPB ist unabhängig, hat eine eigene Rechtspersönlichkeit und einen eigenen Vorsitzenden, zwei Stellvertreter und ein eigenes Sekretariat. Sie organisieren die Arbeit der Behörde und verwalten die Mechanismen für die Lösung zwischen Konflikten unter den nationalen Behörden. Die Behörde entscheidet mit einfacher Mehrheit, Vorgehensregeln und bindende Entscheidungen mit Zweidrittelmehrheit.

Die Behörden können zum Beispiel Datenverarbeiter oder -verantwortliche um Informationen ersuchen, Audits durchführen, Zugang zu Einrichtungen und Daten verlangen, Warnungen und Rügen herausgeben und Bußgelder verhängen, die Einhaltung der Verordnung verlangen, die Verarbeitung von Daten und grenzüberschreitenden Datentransport außerhalb der EU verbieten sowie Standardvertragsbedingungen und bindende Geschäftsregeln formulieren. Ihre Maßnahmen sind gerichtlich überprüfbar. Auch die Behörden selbst können die Gerichte anrufen.

Sind mehrere Länder betroffen, weil eine Institution oder ein Unternehmen in mehreren Ländern aktiv ist, bekommt in der Angelegenheit diejenige nationale Aufsichtsbehörde die Leitungsfunktion, in deren Zuständigkeitsbereich sich die Haupt- oder einzige Niederlassung dieser Einrichtung befindet. Bei Beschwerden ist das Land zuständig, in dem die Beschwerde stattfand oder die Regelverletzung ihre Wirkungen zeigt, sofern die Wirkungen vorwiegend auf das Land beschränkt sind. Die leitende Behörde entscheidet im Zweifel, ob sie eine Angelegenheit an sich zieht oder nicht.

Die leitende Behörde kann andere Behörden zur Mitarbeit auffordern, zum Beispiel für gemeinsame Ermittlungen. Sie muss die übrigen Behörden hinsichtlich ihrer Entscheidungen aber auch konsultieren. In dringlichen Fällen können die Informations- und Kooperationsregeln umgangen werden, etwa um Schaden abzuwenden.

Ariane Rüdiger,
freie Autorin, München

Die Aussichten für 2025: heiß und sonnig

Wer heute ein Data Center plant, sollte wissen, wohin die Entwicklung geht

Die Planer und Architekten von Rechenzentren dürfen sich auf einen noch schnelleren Wandel einstellen: Das kommende Jahrzehnt wird die Data Center gegenüber heute kräftig verändern – zum Beispiel im Umgang mit alternativen Energiequellen und Management-Tools, aber auch im Bereich der Leistungsdichte.

Eine große Herausforderung bleibt nach Meinung der Rechenzentrumsspezialisten, die sich für die Zukunftsstudie „Data Center 2025: Exploring the Possibilities“ geäußert haben, das Management hochkomplexer und dynamischer Rechenzentrumsumgebungen. Zu den wichtigsten Zielen zählen dabei die Aufrechterhaltung der Verfügbarkeit, die Steigerung der Effizienz sowie die Kostensenkung.

Energieversorgung und Effizienz

Doch woher werden die Rechenzentren im Jahr 2025 ihre Energie beziehen und wie hoch wird ihr Bedarf sein? Der Großteil der Branchenkenner (64 %) ist sich sicher, dass das Data Center 2025 bei gleicher Leistung weniger bzw. deutlich weniger Energie verbrauchen wird. Dieses Ergebnis deckt sich damit, dass 84 % der Befragten mit einer höheren Effizienz des Infrastruktur-Equipments (Technik für Stromversorgung und Kühlung etc.) und 67 % mit einer höheren Effizienz der IT-Ausstattung (Server- und Speichersysteme etc.) rechnen. Interessant ist dabei, aus welchen Quellen die Energie voraussichtlich kommen wird.

Die Experten gehen davon aus, dass Rechenzentren künftig über einen Mix aus Energiequellen mit Strom versorgt werden. Solarenergie wird dabei den größten Teil ausmachen, gefolgt von einem jeweils etwa gleich großen Anteil an Kern-, Erdgas- und Windenergie. Man rechnet bei der Versorgung durch Solarenergie mit einer durchschnittlichen Steigerung auf rund 21 % in den nächsten zehn Jahren, wobei die Erwartungen in Westeuropa und den USA mit 18 bzw. 15 % etwas niedriger liegen als in der Region Asien-Pazifik und in Lateinamerika (jeweils 25 %).

Wenn jedoch alternative Energiequellen künftig wichtiger werden, rückt ihre Wirtschaftlichkeit in den Vordergrund. Die Mehrheit der Studienteilnehmer rechnet daher damit, dass der Strom für Hyperscale-Rechenzentren (in denen eine Vielzahl an Kleinstservern installiert ist) in Zukunft in eigenen Anlagen erzeugt wird. Insgesamt gehen 65 % davon aus, dass dies sicher oder zumindest sehr wahrscheinlich der Fall sein wird.

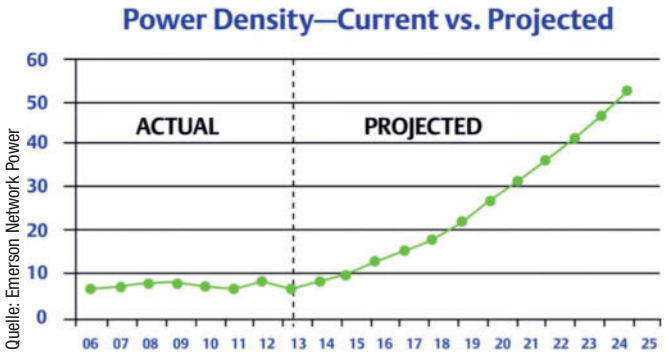
Klimatisierung und Cloud-Anteil

Wie sich Cloud-Computing auswirken wird, ist eine Schlüsselfrage der Branche. Die meisten Fachleute gehen davon aus, dass Cloud-Lösungen auch in Zukunft weiter wachsen und, über alle Rechenzentren hinweg gesehen, einen wesentlichen Bestandteil der Rechen- und Speicherkapazität ausmachen werden: Rund zwei Drittel der Experten nehmen an, dass im Jahr 2025 für 60 % oder mehr der Rechenvorgänge im Data Center Cloud-Dienste genutzt werden. Bei der Hälfte dieser Experten ist der Trend sogar noch deutlicher, sie veranschlagen sogar 80 bis 100 %.

Unternehmenseigene Rechenzentren werden in naher Zukunft zwar nicht verschwinden, doch sie müssen bei Kapazitätsverteilung und Auslastung noch flexibler werden. Und sie werden vermutlich deutlich schrumpfen: 58 % glauben, dass Rechenzentren 2025 nur die Hälfte oder noch weniger der Fläche heutiger Anlagen benötigen werden, 10 % wollen sogar nur ein Zehntel der Fläche gelten lassen.

Die Klimatechnik ist bereits in den vergangenen Jahren deutlich präziser und effizienter geworden. Diese Entwicklung führt dazu, dass der Begriff „Kühlung“ nicht mehr angemessen ist und immer mehr durch ein umfassendes „Wärmemanagement“ ersetzt wird. Die meisten Umfrageteilnehmer rechnen auf diesem Gebiet mit einer Verschiebung hin zur Flüssigkeitskühlung: 41 % stellen sich eine Kombination aus Luft und Flüssigkeit vor, weitere 11 % sehen sogar eine reine Flüssigkeitskühlung als dominierende Methode im Jahr 2025. Insgesamt knapp 40 % nennen Luftkühlung (mit etwa gleicher Verteilung auf Umgebungsluft und Kaltluft), und nur 9 % setzen auf Immersionskühlung, also Temperierung durch Eintauchen in eine Hitze absorbierende Flüssigkeit, die die Wärme aufnimmt und abtransportiert.

Grundsätzlich ist es sehr wahrscheinlich, dass die Bedeutung der Wärmeabfuhr für das Equipment in Rechenzentren zukünftig an Bedeutung weiter zunehmen wird. Dies entspräche den Erwartungen, die



Bis 2013 blieb die Leistungsdichte nach Zahlen der Data Center Users' Group relativ konstant. Wenn die Experten der Data-Center-2025-Studie Recht behalten, wird sie in den kommenden Jahren gewaltig steigen.

einen deutlichen Anstieg der Leistungsdichte pro Rack vorhersagen. Allerdings wird modernes IT-Equipment künftig auch in der Lage sein, immer höhere Temperaturen zu vertragen. Der Bedarf an Kühlung im engeren Sinn wird damit nach und nach geringer werden.

Leistungsdichte und Selbstheilung

2001 lag die Leistungsdichte pro Rack bei 1 kW, sie steigerte sich bis 2014 auf 6 kW. Bis zum Jahr 2025 rechnet der Großteil der Experten (29 %) mit einem Entwicklungssprung, nämlich einer deutlichen Steigerung auf durchschnittlich 40 kW pro Rack. Andere Schätzungen gehen sogar noch weiter: 26 % prognostizieren 80 kW und 15 % sogar 100 kW pro Rack. Ob ein derart massiver Anstieg tatsächlich realistisch ist, bleibe dahingestellt – er würde gravierende Umbrüche in der Architektur und der Kühlung von Rechenzentren mit sich bringen.

Verbesserungen im Data Center Infrastructure Management (DCIM) können helfen, den Herausforderungen von Verfügbarkeit, Skalierbarkeit und Effizienz zu begegnen. Um die Fortschritte und Ergebnisse in diesem Bereich zu messen, wurden für die Data-Center-2025-Studie drei DCIM-Stufen definiert: Transparenz in allen Systemen und Schichten, Selbstoptimierung und Selbstheilung. Die Ergebnisse bestätigen die Beobachtung, dass Rechenzentren immer automatisierter werden: 43 % der Befragten erwarten, dass Data Center in naher Zukunft selbstständig Fehler identifizieren und beheben können (Selbstheilung), 25 % rechnen mit einem sich selbst optimierenden Rechenzentrum, und 29 % sehen für die Zukunft eine umfassende Transparenz über alle Systeme und Schichten hinweg. Anders betrachtet: Nur 3 % glauben, dass DCIM-Lösungen im Jahr 2025 so aussehen werden wie heute.

Innovationsantreiber und Bedarfsmodelle

Eine der wichtigsten und interessantesten Fragen betrifft die Innovationstreiber. Die Branchenexperten wurden gefragt, wer oder was die Motoren von Neuerungen sein würden: Softwareanbieter, unternehmenseigene Rechenzentren, RZ-Infrastrukturhersteller, Hyperscale Data Center oder Hersteller von IT-Equipment? Die Mehrheit sieht diese Rolle bei den letzteren: IT-Equipment-Hersteller werden als wahrscheinlichste Quelle von Innovationen gesehen (28 %), dicht gefolgt von Hyperscale-Rechenzentren (26 %) und RZ-Infrastrukturherstellern (24 %). Von unternehmenseigenen Rechenzentren erwarten nur 13 % der Experten Neues, von Softwareanbietern sogar nur 9 %.

Sicher ist, dass es im Jahr 2025 nicht das eine typische Rechenzentrum geben wird. Je nach Verwendungszweck werden klassische Rechenzentren neben Colocation-Standorten, Superrechenzentren mit einem Strombedarf von einigen zig Megawatt und Supercomputing-Standorte auf Basis schnellster Hochleistungscomputer existieren. Modulare Rechenzentren, die durch ihre Containerbauweise mehr Flexibilität und Skalierbarkeit bieten, werden an Bedeutung gewinnen. Und wo Bandbreite und Geschwindigkeit der Netzwerke nicht ausreichen, wird die IT auch immer mehr in dezentrale Standorte bzw. Mini-Rechenzentren abwandern – eine Entwicklung, die sich auch an parallelen Trends wie höheren Sicherheitsanforderungen, Industrie 4.0 und dem Internet der Dinge ablesen lässt. Und wo große Data Center zu komplex sind, greift eine weitere zeitgleiche Entwicklung: Die IT verlagert sich immer mehr in die Endgeräte.

Dr. Peter Koch,
Vice President Solutions, Emerson Network Power EMEA

Impressum
Themenbeilage Rechenzentren und Infrastruktur

Redaktion just 4 business GmbH
Telefon: 08061 34811100, Fax: 08061 34811109,
E-Mail: tj@just4business.de

Verantwortliche Redakteure:
Thomas Jannot (v. i. S. d. P.), Ralph Novak; Florian Eichberger (Lektorat)

Autoren dieser Ausgabe:
Ferri Abolhassan, André Engel, Gordon Haff, Olaf Hagemann, Bernd Hanstein, Hans-Jürgen Heinrich, Peter Koch, Frank Neubauer, Ariane Rüdiger, Alexander Schlensog, Patrick Schraut

DTP-Produktion:
Enrico Eisert, Kathleen Tiede, Matthias Timm, Hinstorff Media, Rostock

Korrektorat:
Kathleen Tiede, Hinstorff Media, Rostock

Titelbild:
vladimircaribb, fotolia

Verlag
Heise Medien GmbH & Co. KG,
Postfach 61 04 07, 30604 Hannover; Karl-Wiechert-Allee 10, 30625 Hannover;
Telefon: 0511 5352-0, Telefax: 0511 5352-129

Geschäftsführer:
Ansgar Heise, Dr. Alfons Schräder

Mitglieder der Geschäftsleitung:
Beate Gerold, Jörg Mühle

Verlagsleiter:
Dr. Alfons Schräder

Anzeigenleitung (verantwortlich für den Anzeigenteil):
Michael Hanke (-167), E-Mail: michael.hanke@heise.de, www.heise.de/mediadaten/ix

Leiter Vetrieb und Marketing:
André Lux

Druck:
Dierichs Druck + Media GmbH & Co. KG, Frankfurter Straße 168, 34121 Kassel

Die Inserenten

iX Cloud- & OpenStack-Tag

Cloud-Dienste bauen, nutzen & rechtssicher betreiben

15. November 2016, Köln

Foto: © Oleksiy Mark – Fotolia.com

OpenStack, Kubernetes, Mesos & Co. in der Praxis

Die Cloud Computing-Konferenz für Anwender, Entwickler und Experten – offen, flexibel, hersteller- und plattformübergreifend.

Auszug aus dem Programm:

- Quo vadis Linux? Cloud-Betriebssysteme im Vergleich
  // Udo Seidel, Amadeus Data Processing //
- Kubernetes gehosted auf Openstack oder doch besser umgekehrt?
  // Burkhard Noltensmeier, teuto.net //
- OpenStack im Large Enterprise am Beispiel der Metro.Cloud
  // Thomas Lunkwitz, Metro Systems //
- IT-Vertragsrecht und Datenschutz in der Cloud,
  // Dr. Axel Frhr. von dem Bussche, LL.M. (LSE), Partner Taylor Wessing //
- u.v.m.

Teilnahmegebühr (inkl. MwSt.): 399,00 Euro

Sponsoren: rackspace the #1 managed cloud company, redhat, SUSE We adapt. You succeed., BROCADE

Eine Veranstaltung von: iX

Organisiert von: heise Events

Weitere Informationen unter:

www.heise-events.de/ix_cloudopenstack2016

DE CIX

Der sichere, direkte, flexible Weg in die Cloud via DE-CIX

Mehr als 1.000 Netzbetreiber, Internet Service Provider, Content-Anbieter, Distributoren sowie Cloud-Provider aus 60 Ländern vertrauen unseren Premium Peering- und Interconnection-Services.

Werden auch Sie Teil der Erfolgsgeschichte! Für mehr Informationen wenden Sie sich an Ihren Provider.

www.de-cix.net