

Wettlauf im LAN

Vergleichstest Gigabit-Ethernet-Switches – Im Zeitalter von Unified-Communications liefern sich die verschiedenen Anwendungen häufig einen Wettlauf im LAN. Priorisierungsmechanismen sollen dafür sorgen, dass zeitkritische Daten zur rechten Zeit am rechten Ort eintreffen.

Netzwerkbandbreiten werden auch im Zeitalter von Gigabit- und 10-Gigabit-Ethernet immer wieder zur knappen Ressource. Hierfür sorgen bandbreitenhungrige und anspruchsvolle Anwendungen wie Voice- und Video-over-IP, insbesondere in der HDTV-Variante. Aber auch das Datenaufkommen der »klassischen« Anwendungen wächst munter weiter. Damit moderne LAN-Switches diesen Anforderungen gerecht werden, statten die Switch-Hersteller ihre Systeme mit immer mehr Funktionalität und Leistungsvermögen aus. Ein Vergleichstest aktueller LAN-Switches in den Real-World Labs an der FH Stralsund sollte klären, wie gut diese Systeme ihren Anforderungen gewachsen sind. Dabei sollten insbesondere die Priorisierungsmechanismen untersucht werden, die zeitkritischen Daten die Vorfahrt im LAN einräumen sollen, wenn es im Netzwerk eng wird.

Network Computing Musterfirma

Für die Testausschreibung haben wir auf unsere Network Computing Musterfirma zurückgegriffen. Unsere Musterfirma ist ein innovatives Unternehmen, das im Bereich der Automobilzubehörindustrie tätig ist. Die Musterfirma verteilt sich auf mehrere Standorte:

- Stralsund: Firmenhauptsitz mit den Abteilungen
- ◆ Forschung & Entwicklung (250 PC-Arbeitsplätze),
 - ◆ Marketing (150 PC-Arbeitsplätze),
 - ◆ Sales (200 PC-Arbeitsplätze),
 - ◆ Verwaltung (80 PC-Arbeitsplätze),
 - ◆ Rechenzentrum (Serverfarm, SAN, Administration, 5 PC-Arbeitsplätze),
 - ◆ Geschäftsführung (20 PC-Arbeitsplätze).
- Rostock: Produktionsstandort mit
- ◆ Produktion in 4 Betrieben mit insgesamt 300 PC-Arbeitsplätzen,



- ◆ Backup-Rechenzentrum (Serverfarm, SAN, Administration, 5 PC-Arbeitsplätze).

Vier Niederlassungen in Frankfurt, Berlin, München und Passau mit jeweils 30 PC-Arbeitsplätzen.

Zwei Auslandsniederlassungen in New York und Hongkong mit jeweils 40 PC-Arbeitsplätzen.

Die Network Computing Musterfirma möchte alle Standorte sowie Partnerfirmen in einem Intranet auf IP-Basis integrieren. Neben den klassischen Datenanwendungen soll über dieses Intranet auch Telefonie und Videoübertragung realisiert werden. Dabei soll das Unternehmensnetz in Segmente unterteilt werden, die den verschiedenen Abteilungen an den Hauptstandorten beziehungsweise den einzelnen Niederlassungen zugeordnet werden sollen. Die Segmente sollen hochperformant miteinander verbunden werden aber zugleich auch durch die entsprechenden Sicherheitstechnologien gegeneinander abgesichert werden.

Das Testszenario

Die Network Computing Musterfirma möchte neben den klassischen Datenapplikationen und Voice-over-IP weitere Real-Time-Applikationen in ihr Unternehmensnetz zunächst am Firmenhauptsitz integrieren. Ein geeigneter Vergleichstest soll evaluieren, welche Switches für diese Aufgaben auch unter entsprechender Last geeignet sind.

REPORTCARD POLICY-BASED SWITCHING

	Gewichtung	Extreme Networks Summit X450a-24t	Alcatel-Lucent Omniswitch 6850	Enterasys B3G124-24 SecureStack B3
Unicast Intern Burst-Size 1	40 %	5,00	5,00	5,00
Unicast Intern Burst-Size 100	10 %	5,00	3,00	3,00
Multicast Intern Burst-Size 1	40 %	5,00	5,00	5,00
Multicast Intern Burst-Size 1	10 %	2,00	3,00	1,00
Gesamtergebnis	100 %	4,70	4,60	4,40
		A	A	A-

A > 4,3; B > 3,5; C > 2,5; D > 1,5; E < 1,5; Die Bewertungen A bis C enthalten in ihren Bereichen + oder -; Gesamtergebnisse und gewichtete Ergebnisse basieren auf einer Skala von 0 bis 5. Abweichung vom Sollwert: > 30 Prozent = 1; > 20 Prozent = 2; > 10 Prozent = 3; > 5 Prozent = 4; < / = 5 Prozent = 5

DAS TESTFELD

- ◆ Alcatel-Lucent Omniswitch 6850
- ◆ Enterasys B3G124-24 SecureStack B3
- ◆ Extreme Networks Summit X450a-24t

Folgende Dienste sollen im LAN integriert werden:

- ◆ Videokonferenzen (Video-over-IP, bidirektional, unicast),
- ◆ Videodistribution (Multicast-Betrieb),
- ◆ Voice-over-IP (Call-Center),
- ◆ SAP-Anwendungsdaten,
- ◆ CAD- und CAM-Anwendungsdaten,
- ◆ übrige Datenanwendungen sowie Updates.

Die Marketing-Abteilung produziert auch Filme zu Werbungs- und Schulungszwecken und hat ein großes Videoarchiv. Weiterhin unterhält die Marketing-Abteilung ein Callcenter für das Direktmarketing.

Die Abteilungen Forschung & Entwicklung wie auch die Produktion arbeiten mit CAD- und CAM-Applikationen. Der Fertigungsbereich soll auf Industrial-Ethernet umgerüstet und in das Intranet integriert werden. Um die möglichst absolute Störungsfreiheit der Kommunikations- und Arbeitsprozesse im Unternehmen zu garantieren, ist mindestens eine mindestens vierstufige Daten-Priorisierung sowie eine intelligente Queuing-Policy erforderlich. Gefordert ist für die Edge-Switches eine effektive Datenpriorisierung.

Daraus ergeben sich folgende Anforderungen an die Teststellung Edge-Switch:

- ◆ Layer-3-Gigabit-Ethernet-Switch (Routing),
- ◆ mindestens 24 Gigabit-Ethernet-Kupfer-Ports (Steckertyp RJ45),
- ◆ Link-Aggregation zur Trunk-Bildung für mindestens vier Ports,
- ◆ Diffserv-Datenpriorisierung nach RFC 2474 oder
- ◆ Type-of-Service-Datenpriorisierung nach RFC 791 und/oder 1349 auf Layer-3,
- ◆ CoS-Queuing-Mechanismen (Strict-Priority-Queuing) sowie
- ◆ Multicast-Funktionalität (IGMP-Snooping).

Die Testreihen

Insgesamt haben wir im Rahmen unseres Vergleichstests Gigabit-Ethernet-Switches vier Messreihen durchgeführt. Im Betriebsmodus »Unicast intern« haben wir mit unserem Lastgenerator/Analysator »Smartbits 6000C Traffic Generator/Analysor« von Spirent Communications Last erzeugt und auf 16 Eingangsports des jeweils zu testenden Switches gesendet. Die Datenströme haben wir an vier Ports des selben Switches adressiert und mit den Smartbits analysiert.



Extreme Networks Summit X450a-24t



Als zweite Messreihe haben wir dann die einzelnen Switches im Multicast-Betrieb einer Zangenmessung unterzogen. Dabei haben wir mit unserem Lastgenerator Datenströme erzeugt und auf vier Eingangsports des jeweiligen Switches gesendet. Jeder dieser vier Ports hat dann die Datenströme an 16 Ports weiter geleitet. Dort wurden sie vom Analysator gelesen und ausgewertet.

In der dritten Zangenmessung haben wir dann noch einmal den Betriebsmodus Multicast eingesetzt. Dabei haben wir hier wieder zwei baugleiche Switches mit einem Vier-Port-Trunk verbunden und die Smartbits mit den vier Ausgangsports des zweiten Switches verbunden, um die Durchsätze zu analysieren.

Im Betriebsmodus »Unicast Trunk« haben wir zwei baugleiche Switches in Reihe geschaltet. Dabei haben wir wieder mit unserem Lastgenerator Last erzeugt und auf 16 Eingangsports des jeweils zu testenden Switches gesendet. Über vier Ausgangsports des ersten Switches, die einen Trunk bildeten, haben wir dann vier Ports des zweiten Switches adressiert. Die vier Eingangsports des zweiten Switches haben dann die Datenströme an vier Ausgangsports des selben Switches gesendet.

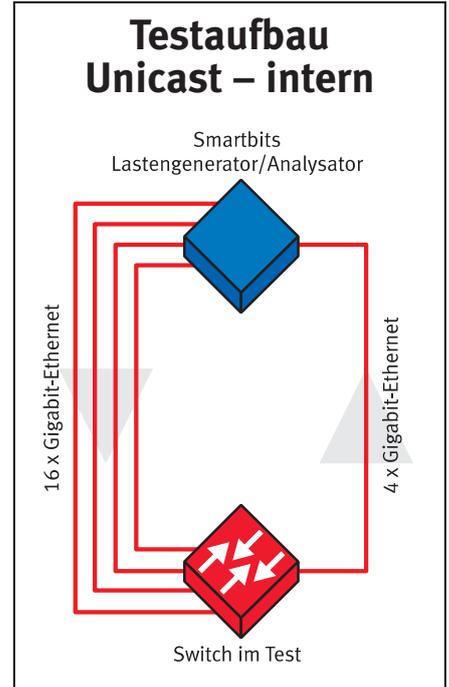
Alle Messungen haben wir nacheinander mit Frame-Formaten von 64, 256, 512, 1024 und 1518 Byte durchgeführt. Alle Messungen haben wir jeweils mit einer Burst-Size von einem und von 100 Frames durchgeführt. Beträgt die Burst-Size einen Frame, bedeutet dies, dass der Lastgenerator die erzeugte Last kontinuierlich auf der Zeitachse verteilt sendet. Der belastete Switch kann so die erzeugte Last kontinuierlich abarbeiten. Im Betrieb mit einer Burst-Size von 100 Frames sendet der Lastgenerator dagegen

immer in einem gleichen Verhältnis zueinander standen.

Bei den Messungen, die wir alle mit Strict-Priority durchgeführt haben, sollten die Switches folgendes idealtypisches Verhalten zeigen: Bei 100 Prozent Last an den Ausgangsports sollten alle Datenströme ungehindert weiter geleitet werden. Bei zunehmender Last sollten die Switches dann jeweils möglichst viele Daten der niedrigsten Prioritäten verlieren und möglichst viele Daten der hohen Prioritäten unbeschadet passieren lassen. Bei vierfacher Überlast sollte der jeweilige Switch folglich alle Daten mit Ausnahme denen der höchsten Priorität verwerfen.

Test Unicast intern

Alcatel-Lucent Omniswitch-6850 zeigte bei allen unseren Messungen mit einer Burst-Size von einem Frame eine mustergültige Priorisierung. Entsprechend den Regeln der Strict-Priority verwarf er bei entsprechenden Überlasten die Daten der niedrigen Prioritäten zu Gunsten der hö-



Alcatel-Lucent Omniswitch 6850

mit Leitungsgeschwindigkeit 100 Frames nacheinander. Dann folgt eine Pause, so dass der Generator im Mittel die jeweils konfigurierte Durchsatzleistung erzeugt. Während der Sendephase muss der belastete Switch die Datenströme im Pufferspeicher zwischenspeichern und in der folgenden Pause abarbeiten.

Die Last an den Eingangsports startete jeweils bei 25 Prozent und wurde dann in entsprechenden Schritten auf 100 Prozent erhöht. Das bedeutete an den Ausgangsports eine Last zwischen 100 und 400 Prozent – dabei waren die Lastanteile für die vier Prioritäten immer gleich groß, so dass die Daten der verschiedenen Prioritäten

her priorisierten Datenströme. Unter vierfacher Überlast konnten die Datenströme der höchsten Priorität weiterhin ohne messbare Verluste das System passieren. Verwendeten wir ein Burst-Format von 100 Frames, was einer echten Leitungsgeschwindigkeit ohne Pausen zum Abarbeiten der Daten in den Queues entspricht, dann kam der Switch im Betrieb mit großen Frames an seine Grenzen. Spürbare Datenverluste waren ab den Frame-Größen 512 beziehungsweise 1024 Byte messbar. Im Betrieb mit den größten Frames konnten wir irreguläre Datenverluste von maximal über 50 Prozent verzeichnen. Allerdings schaffte es der Alcatel-Lu-

cent-Switch, die höchste Priorität immer von Datenverlusten frei zu halten. Unter Vollast war deshalb auch bei der Messung mit einer Burst-Size von 100 die Priorisierungswelt wieder in Ordnung.

Enterasys B3G124-24-Securestack-B3 stand dem Mitbewerb in dieser Disziplin in nichts nach. Es arbeitete im Betrieb mit einer Burst-Size von einem Frame absolut korrekt und verwarf keinerlei Daten außerplanmäßig. Nicht ganz so einwandfrei war sein Verhalten dann bei unseren Messungen mit einer Burst-Size von 100 Frames. Unter Teillast kam es hier zu ein paar außerplanmäßigen Datenverlusten, die mit der Frame-Größe anstiegen und in niedrigeren Prioritäten bis über 50 Prozent betrug. Ab einem Frame-Format von 1024 Byte waren von außerplanmäßigen Datenverlusten auch bis zu rund sechs Prozent der Daten der höchsten Priorität betroffen. Unter Vollast schaffte es der Enterasys-Switch dann aber wieder, die Daten der höchsten Priorität trotz vierfacher Überlast verlustfrei zu transportieren.

Auch Extreme Networks Summit-X450a-24t verhielt sich bei unseren Messungen mit einer Burst-Size von einem Frame absolut mustergültig. Er leistete sich keinerlei nennenswerten außerplanmäßigen Datenverluste und verwarf, sofern das von der Lastsituation her notwendig war, die niedriger priorisierten Daten zu Gunsten der jeweils höher priorisierten. Verwendeten wir eine Burst-Size von 100 Frames, dann ähnelte das Verhalten des Switches dem der anderen Switches im Testfeld. Unter Teillast kam es zu außerplanmäßigen Datenverlusten, wenn wir Frames der Größen 1024 und 1518 Byte verwendeten. Dabei gingen maximal um die 38 Prozent der Daten in den einzelnen Prioritäten verloren. Dabei leistete sich der Extreme-Networks-



Enterasys B3G124-24 SecureStack B3

Switch bei 100 und 133 Prozent Last und 1518-Byte-Frames den Ausrutscher, jeweils gut 14 Prozent der Daten in der höchsten Priorität zu verlieren. Unter höheren Lasten beherrschte der Switch dann aber wieder die Strict-Priority-Spielregeln und leistete sich auch bei der hohen Burst-Size keine weiteren Ausrutscher.

Test Multicast intern

Alcatel-Lucent's Omniswitch-6850 beherrschte den Multicast-Betrieb mit einer Burst-Size von einem Frame ebenso souverän wie den Unicast-Betrieb. Er leistete sich auch hier keinerlei außerplanmäßigen Datenverluste und verwarf getreu den Regeln der Strict-Priority bei entsprechenden Überlasten die niedriger priorisierten Daten zu Gunsten der höher priorisierten. Verwendeten wir ein Burst-Format von 100 Frames, kam es zu außerplanmäßigen Datenverlusten, wenn die Frames 1024 Byte groß oder größer waren. Aber auch bei diesen Messungen kam es zu keinerlei außerplanmäßigen Datenverlusten in der höchsten Priorität. Aus diesem Grund arbeitete der Alcatel-Lucent-Switch dann auch unter Volllast tadellos und ließ sich zu keinen weiteren außerplanmäßigen Datenverlusten mehr verleiten.

Enterasys B3G124-24-Securestack-B3 arbeitete auch im Multicast-Betrieb mit einer Burst-Size von einem Frame absolut überzeugend. Wir konnten ihm hier keinerlei außerplanmäßigen Datenverluste nachweisen. Wiederholten wir die Tests mit einem Burst-Format von 100 Frames, so kam auch der Enterasys-Switch an seine Grenzen. Dabei teilte er die Datenverluste recht gleichmäßig unter den Prioritäten 1, 3 und 5 auf. Absolut stiegen die irregulären Datenverluste mit dem Frame-Format an und erreichten deutliche Höchstwerte. Die Datenströme der höchsten Priorität blieben dabei aber von allen Verlusten frei und gelangten – auch unter vierfacher Überlast – ungehindert an Ziel.

Auch Extreme Networks Summit-X450a-24t ließ sich im Betrieb mit einer Burst-Size von einem Frame keinerlei Schwächen nachweisen. Er arbeitete absolut korrekt und verwarf Datenströme wenn nötig entsprechend den geforderten Regeln. Wechselten wir zu einer Burst-Size von 100 Frames, waren auch hier wieder ein paar Unregelmäßigkeiten zu registrieren. Allerdings waren die Datenströme der höchsten Priorität hier in keinem Fall von unerwünschten Datenverlusten betroffen.

Tests im Trunking-Betrieb

Die Trunking-Technologie ermöglicht das Bündeln von mehreren Ports zu einer logischen Verbindung zwischen zwei Switches. Diese Methode verspricht die erforderliche Uplink-Bandbreiten einzurichten ohne gleich auf teurere 10-Gigabit-Ethernet-Uplinks ausweichen zu müssen. Es ist also eine kostengünstige Methode, Edge-Switches performant ans Netzwerk anzubinden. Den Trunk-Mechanismus realisieren die Switch-Hersteller heute so, dass alle Datenströme, die über den Trunk gesendet werden sol-

FRAME-LOSS IN PROZENT MIT TRUNK

		1 IP-Adresse	100 IP-Adressen	Sollwert
Eingangslast	25 %	75	0	0
Eingangslast	33 %	81	25	25
Eingangslast	50 %	87	50	50
Eingangslast	100 %	94	75	75

len, in jeweils eine Queue pro Port laufen. Jeder Trunk-Port hat seine eigene Hardware-Queue. Die einzelnen Datenpakete, die über den Trunk gesendet werden sollen, werden mittels proprietärer Hash-Algorithmen den einzelnen Queues zugewiesen, ohne Rücksicht auf die Auslastung des jeweiligen Ports zu nehmen. Die Hash-Algorithmen verwenden Teile der IP-Adressen und Teile der Portnummern der sendenden und / oder der empfangenden Systeme. Die Hersteller halten ihre Algorithmen geheim und wir dürfen diese aus rechtlichen Gründen hier auch nicht veröffentlichen.

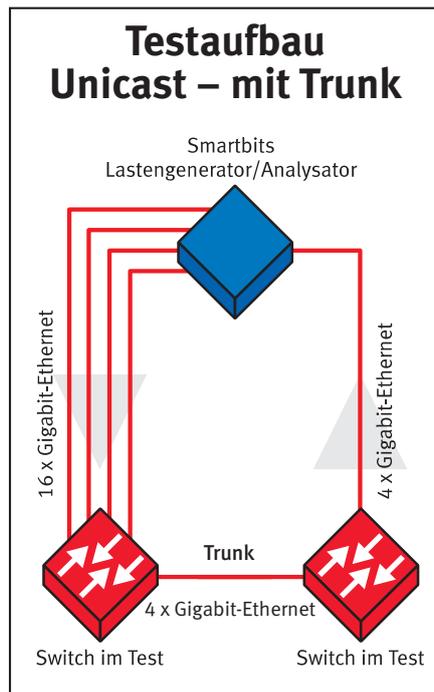
Im Test stellte sich nun heraus, dass für alle Switches Situationen erzeugbar waren, in denen die Algorithmen für eine gleichmäßige Verteilung der Datenströme sorgten. Andererseits kam es aber auch bei allen Systemen zu Situationen, in denen die Verteilung der Datenströme sehr unsymmetrisch erfolgte. Jede Leitung hat ihre eigene Queue, daher werden die Datenströme un-

abhängig voneinander abgearbeitet. Das hatte dann deutliche Performance-Einbußen zur Folge, weil der »verstopfte« Port einen Rückstau in der davor geschalteten Queue verursachte. Dabei verhielten sich die Switches nicht eindeutig für den Administrator vorhersagbar, da dieser den Algorithmus nicht kennt und zudem die Absenderportnummern in der Regel dynamisch vergeben werden, was er in der Regel nicht kontrollieren kann.

Generell gilt allerdings bei diesem Verhalten, dass bei steigender Anzahl der IP-Adressen im Subnetz auch die Verteilung der Datenströme im Trunk gleichmäßiger wurde. Da es sich hierbei um ein konzeptionelles Problem handelt, das allen aktuellen Switches – nicht nur denen im aktuellen Testfeld – gemeinsam ist, haben wir einen Test durchgeführt, der das Problem exemplarisch darstellt. Ansonsten haben die Tests gezeigt, dass für alle Switches im Testfeld hier Szenarien möglich sind, in denen die Systeme die Datenströme gleichmäßig verteilen, und solche, in denen sie sehr asymmetrisch arbeiten und somit auf sehr schlechte Durchsatzleistungen kommen. Daher haben wir darauf verzichtet, hier zwischen den Systemen Vergleiche anzustellen. Durch eine willkürliche Festlegung der Rahmenbedingungen ständen Sieger und Verlierer automatisch im voraus fest.

Statt dessen haben wir ohne Ansehen des Herstellers ein charakteristisches Switch-Verhalten dargestellt. Wir haben zwei baugleiche Switches über vier Ports miteinander verbunden und so einen Trunk mit vier Gigabit-Ethernet-Verbindungen etabliert. Dann haben wir in einer Zangemessung auf 16 Eingangsports Datenströme mit unseren Smartbits gesendet. Diese wurde über den Trunk auf vier Ausgangsports des zweiten Switch adressiert und dort mit Hilfe der Smartbits analysiert. Die Eingangslast haben wir wieder schrittweise von 25 über 33 und 50 Prozent bis auf 100 Prozent Last gesteigert. Hierdurch entsteht wieder eine maximale Überlast von 400 Prozent. Diese Messung haben wir einmal mit nur einer Absender-IP-Adresse und dann mit 100 Absender-IP-Adressen durchgeführt.

Je mehr Rechner an der Erzeugung der Datenströme beteiligt sind, um so wahrscheinlicher ist es, dass die Links innerhalb eines Trunks gleichmäßig ausgelastet sind. Wenige Rechner an



TESTVERFAHREN POLICY-BASED SWITCHING

Als Lastgenerator und Analysator haben wir in unseren Real-World Labs einen »Smartbits 6000C Traffic Generator/Analysor« von Spirent Communications eingesetzt. Das System ist mit der Software »SmartFlow« ausgestattet und mit 24 Gigabit-Ethernet-Kupfer-Ports sowie zwei 10-Gigabit-Ethernet-Fibre-Ports bestückt. Alle Ports können softwareseitig als Lastgeneratorausgang und/oder als Analytoreingang eingesetzt werden. Die Class-of-Service-Eigenschaften der Switches im Testfeld haben wir in verschiedenen Testreihen gemäß RFC 2544 (vgl.: www.ietf.org/rfc/rfc2544.txt) gemessen. In diesen Tests haben wir die Priorisierung auf Layer-3 untersucht. In unseren Test-Setups haben wir verschieden priorisierte Datenströme von den Eingangsports auf die Ausgangsports gesendet. Durch eine gezielte Überlastung der Switches in diesen Tests ist es möglich, das genaue Datenverlustverhalten sowie weitere Testparameter wie Latency oder Jitter zu ermitteln, das Leistungspotential der untersuchten Switches zu analysieren und deren Eignung für bestimmte Einsatzszenarien zu prüfen.



Alle Messergebnisse finden Sie im Internet unter www.networkcomputing.de/nwc_downloads/switchtest_08.zip

einem Edge-Switch sorgen dagegen schnell für eine einseitige Verteilung der Datenströme auf einzelne Ports. Dieses Verhalten kann beispielsweise zu Engpässen führen, wenn ein Server über einen Switch sein nächtliches Backup fährt. Das »Trunk-Phänomen« ist nicht konfigurierbar, sondern vom dynamischen Geschehen im Netzwerk abhängig. Somit besteht keine Planungssicherheit für die Netzwerkverantwortlichen. In der Praxis ist es daher häufig besser, einen Trunk durch einen 10-Gigabit-Ethernet-Uplink zu ersetzen.

Fazit

Die LAN-Switches der aktuellen Generation sind spürbar besser geworden. Auch wenn sie in Volllast-Situationen im Betrieb mit homogenen Datenströmen bestehend aus sehr großen Frames an ihre Leistungsgrenzen kommen: Im realen Betrieb mit typischen Lastmustern sollte es in der Regel zu keinen spürbaren Performance-Problemen und Kommunikationsstörungen mehr kommen.

Nichtsdestotrotz ist ihr Einsatz auch heute noch nicht ohne Risiken und Nebenwirkungen. Wer Trunks einsetzt, um teure 10-Gigabit-Ethernet-Uplink-Ports zu sparen, kann unter Umständen deutliche Leistungseinbußen erleben. Grund dafür sind die nicht optimalen Algorithmen, nach denen die Switches die Datenströme auf die verschiedenen Ports verteilen. Dabei ist das Verhalten der Switches hier nicht wirklich vorhersagbar. Und das gilt ohne Ausnahme für alle Switches, die wir in letzter Zeit in den Real-World Labs an der FH Stralsund untersucht haben.

Vor Überraschungen sind nur die IT-Verantwortlichen geschützt, die genau wissen, wie die Systeme arbeiten und dies dann gegebenenfalls auch messtechnisch überprüfen lassen. Ansonsten drohen nach wie vor Gefahren wie die der asymmetrischen Auslastung von Trunks, die durchaus beispielsweise das nächtliche Backup vereiteln können.

Dipl.-Ing. Thomas Rottenau,
 Prof. Dr. Bernhard G. Stütz,
dg@networkcomputing.de