

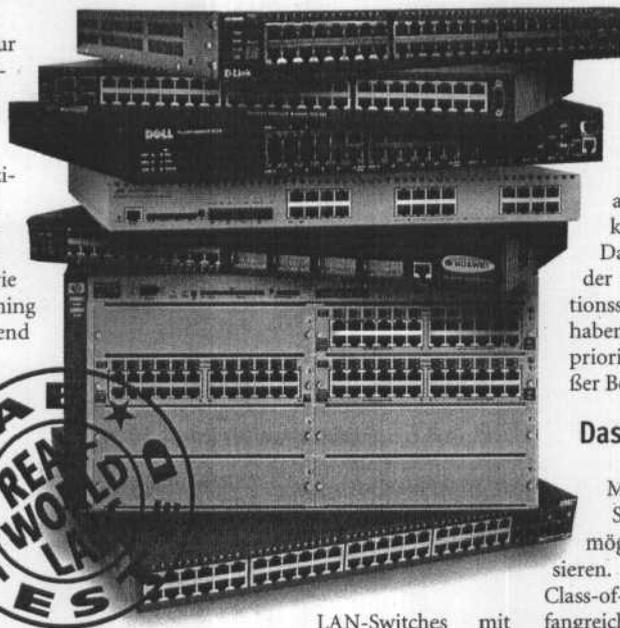
7 Gigabit-Ethernet-Switches

Intelligenter schalten

Gigabit-Durchsätze, Quality-of-Service und Uni- wie Multicast-Funktionalität – moderne Unternehmensnetze stellen hohe Anforderungen an Ethernet-Switches. In unseren Real-World Labs mussten sieben Gigabit-Ethernet-Switches zeigen, ob sie diesen Anforderungen auch gewachsen sind und intelligent genug schalten können.

Moderne LAN-Switches müssen nicht nur den verschiedensten Anwendungen eine adäquate Performance und Übertragungsqualität bieten, sondern auch diverse Kommunikationsbeziehungen beherrschen. Klassischerweise stellt ein Switch dezidierte Verbindungen zwischen jeweils zwei Endgeräten her, die an einen Port angeschlossen sind. Neben solchen Unicast-Verbindungen werden durch Anwendungen wie Video-over-IP – beispielsweise beim E-Learning oder der Video-Distribution – aber zunehmend auch Punkt-zu-mehr-Punkt-Verbindungen erforderlich. In diesem Multicast-Modus verteilt der Switch den Datenstrom, der an einem Punkt ankommt, auf eine festgelegte Anzahl von Empfängern, die an entsprechend festgelegten Ports angeschlossen sind. An welche Ports der Switch die Datenströme zu lenken hat, »lernt« er mit Hilfe des auf IP-basierenden Internet-Group-Management-Protocoll, kurz IGMP. Vom Broadcast-Betrieb unterscheidet sich der Multicast-Betrieb dadurch, dass nicht generell an alle erreichbaren Empfänger, sondern nur an eine Liste individuell festzulegender Empfänger gesendet wird.

Auch im Multicast-Betrieb in hochperformanten konvergenten Netzen ist es wichtig, dass kritische und wichtige Daten bevorzugt behandelt werden. Dafür sorgen Gigabit-Ethernet-



LAN-Switches mit ihrer implementierten Datenpriorisierung. In vielen auf Ethernet basierenden Unternehmensnetzen sind Lastspitzen dafür verantwortlich, dass es im Netz eng wird. Dann kommt es je nach Aus- und Überlastung zu teils erheblichen Datenverlusten. Aber auch andere Übertragungstechnische Parameter wie Latency oder Jitter können – wenn sie gewisse Toleranzwerte überschreiten, weil beispielsweise das Netzwerk

und seine Komponenten überfordert sind, zu Störungen einzelner Services oder auch der gesamten Kommunikation führen. Im Zeitalter konvergenter Netze, die Echtzeitapplikationen wie Voice- und Video-over-IP, aber auch Produktionssteuerungsdaten in das klassische Datennetz integrieren, führen Datenverluste und andere Störungen in der Praxis zu empfindlichen Kommunikationsstörungen und Produktionsausfällen. So haben in realen Projekten schon mangelhaft priorisierende Switches ganze Call-Center außer Betrieb gesetzt.

Das Real-World-Labs-Testszenario

Wir wollten wissen, wie gut Queuing-Mechanismen heute in aktuellen LAN-Switches arbeiten und ob es mit ihnen möglich ist, die gewünschte Policy zu realisieren. Aus diesem Grund haben wir sieben Class-of-Service-fähige LAN-Switches einer umfangreichen Testprozedur unterzogen. Um unseren Test wie gewohnt in der Vorbereitungsphase sauber zu strukturieren, haben wir zur Definition unserer Test-Spezifikation, die wir an alle einschlägigen Hersteller gesandt haben, wieder auf unser Modellunternehmen Highfair zurückgegriffen.

Das Modellunternehmen Highfair möchte neben den klassischen Datenapplikationen und Voice-over-IP weitere Real-Time-Applikationen

Report-Card / interaktiv unter www.networkcomputing.de

Multicast-Switching und Stray-Frames

| Feature | Gewichtung | D-Link DXS 3350SR | Allied Telesyn AT-9924T | Extreme Networks Summit 400-48T | SMC TigerSwitch 8648T | HP ProCurve Switch 5308x | Huawei Quidway S5024G | Dell PowerConnect 6024 |
|------------------------------------|---------------|----------------------|----------------------------|--|-----------------------------|--------------------------------|-----------------------------|------------------------------|
| Multicast-Switching Burst-Size 1 | 45,0% | 5 | 5 | 5 | 5 | 3 | 3 | 1 |
| Multicast-Switching Burst-Size 100 | 45,0% | 3 | 2 | 2 | 2 | 2 | 1 | 1 |
| Stray-Frames | 10,0% | 5 | 5 | 5 | 2 | 5 | 4 | 1 |
| Gesamtergebnis | 100,0% | 4,10 | 3,65 | 3,65 | 3,35 | 2,75 | 2,20 | 1,00 |
| | | B+ | B- | B- | C+ | C- | D | E |

Abweichungen vom Sollwert:

| | |
|----------------------|--------------|
| Multicast-Switching: | Stray-Frames |
| > 30% = 1 | > 10 000 = 1 |
| > 20% = 2 | > 5000 = 2 |
| > 10% = 3 | > 1000 = 3 |
| > 5% = 4 | > 100 = 4 |
| 5% = 5 | 100 = 5 |

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.

in ihr Unternehmensnetz integrieren. Ein geeigneter Vergleichstest soll ergeben, welche Switches für diese Aufgaben auch unter entsprechender Last geeignet sind. Dabei sollen verschiedene CoS-Queuing-Mechanismen wie Strict-Priority-Queuing, Weighted-Fair-Queuing oder Weighted-Round-Robin auf ihre Eignung für das geplante Szenario untersucht werden.

Folgende Dienste sollen im LAN integriert werden:

- ▶ Videokonferenzen (Video-over-IP, bidirektional, Unicast-Betrieb),
- ▶ Videodistribution (Multicast-Betrieb),
- ▶ Voice-over-IP (Call-Center),
- ▶ SAP-Anwendungsdaten sowie
- ▶ übrige Datenanwendungen und Updates.

Um die möglichst absolute Störungsfreiheit der Kommunikations- und Arbeitsprozesse im Unternehmen zu garantieren, sind eine vierstufige Daten-Priorisierung sowie eine intelligente Queuing-Policy erforderlich. Gefordert ist für die Switches neben der Datenpriorisierung ein intelligentes Bandbreitenmanagement. Dieses soll es ermöglichen, von vorneherein eine Überlastung des Backbones zu vermeiden. Daraus ergaben sich folgende Anforderungen an die Teststellungen:

- ▶ Layer-3-Gigabit-Ethernet-Switch,
- ▶ mindestens 20 Gigabit-Ethernet-Kupfer-Ports mit RJ45-Stecker,
- ▶ Datenpriorisierung nach IEEE 802.1p/Q auf Layer-2,
- ▶ Diffserv-Datenpriorisierung nach RFC 2474 oder
- ▶ Type-of-Service-Datenpriorisierung nach RFC 791 und/oder 1349 auf Layer-3,
- ▶ mindestens zwei CoS-Queuing-Mechanismen wie Strict-Priority-Queuing, Weighted-Fair-Queuing oder Weighted-Round-Robin, die softwareseitig konfiguriert werden können, sowie
- ▶ Multicast-Funktionalität (IGMP-Snooping).

Als Testverfahren haben wir Messungen nach RFC 2544 (Many-to-One) festgelegt, die die Parameter Performance, Packet-Loss, Latency und Jitter ermitteln. Analysiert wird dann das unterschiedliche Verhalten der Systeme in den verschiedenen CoS-Queuing-Modi.

Unsere Testspezifikation haben wir an die einschlägigen Hersteller geschickt und diese eingeladen, an unserem Vergleichstest in unseren Real-World Labs an der FH Stralsund teilzunehmen. Im persönlichen Gespräch haben wir dann den einschlägigen Herstellern unser Testprojekt erläutert. Neben Allied Telesyn, Dell, D-Link, Extreme Networks, Hewlett-Packard und SMC waren auch Cisco, Huawei und Nortel an einer Teilnahme interessiert. Als dann nach fast dreimonatiger Vorlaufphase unsere Testtermine anstanden, sah sich Nortel nicht in der Lage, eine Teststellung zur Verfügung zu stellen. Und auch Cisco hat wegen »mangelnder Ressourcen« ihre Teilnahme abgesagt. Der chinesische Hersteller Huawei, der sich gerade anschickt, den europäischen Markt zu erobern, sah sich auch nicht in der Lage, unseren Test zu unterstützen. Das Feld der Gigabit-Ethernet-Switches bildeten schließlich Allied Telesyn »AT-9924T«, Dells »PowerCon-

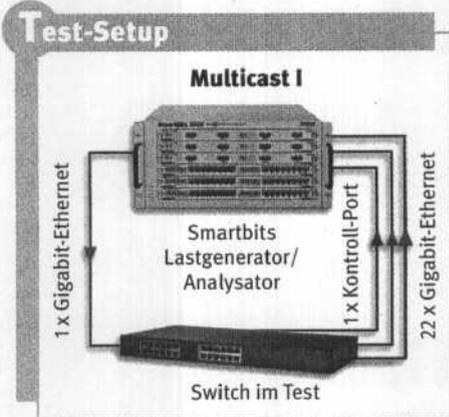
nect 6024«, D-Links »DXS-3350SR«, Extreme Networks »Summit 400-48T«, HPs »ProCurve Switch 5308x« und SMCs »TigerSwitch 8648T«. In einem Nachttest war es uns dann noch möglich, Huawei »Quidway S5024G« unserem standardisierten Testverfahren zu unterziehen. Im vorliegenden Artikel stellen wir nun dar, wie sich die Systeme in Bezug auf Multicast-Funktionalität und Stray-Frames verhalten haben. Die Messergebnisse mit Strict-Priority-Queuing und mit Bandbreitenmanagement haben wir in Network Computing 4/05, S.14 ff. sowie im Network Computing Special Networking 2005, S.4 ff. dargestellt.

Aus den Ergebnissen von Performance-Messungen wie den von uns durchgeführten ist gut zu erkennen, ob, und wenn ja, in welchem Bereich, das jeweilige System Schwierigkeiten hat. Arbeitet der so belastete Switch korrekt, muss er in allen Fällen gemäß den »Class-of-Service-Regeln« die niedrig priorisierten Daten zugunsten der höher priorisierten verwerfen. Ein Datenverlust in der höchsten Priorität dürfte bei allen unseren Strict-Priority-Tests theoretisch nicht vorkommen. Nur dann würde der jeweilige Switch die fehlerfreie Übertragung der am höchsten priorisierten Echtzeitanwendung, beispielsweise einer Video-Konferenz, garantieren.

Für die Switches sind pro Zeiteinheit um so mehr Header-Informationen auszuwerten, um so kleiner die einzelnen Datenrahmen sind. Ein Switch wird also zuerst Probleme mit 64-Byte-Datenströmen bekommen, wenn er bei der internen Verarbeitungsgeschwindigkeit an seine Grenzen stößt. Bei großen Datenrahmen können je nach Design dagegen eher Probleme mit dem Speichermanagement beziehungsweise mit der Größe des überhaupt verfügbaren Pufferspeichers entstehen.

Multicast-Betrieb

Bei den vorhergehenden Messungen hat es sich generell um Punkt-zu-Punkt-Verbindungen gehandelt, bei denen der zu testende Switch Datenströme zwischen den verschiedenen Ports vermitteln muss. Im Multicast-Betrieb sendet der Switch dagegen den Datenstrom, der an einem Port ankommt, an eine festgelegte Zahl von Ports weiter. Ein solches Switch-Verhalten ist immer



Info

Das Testfeld

Gigabit-Ethernet-Switches

- ▶ Allied Telesyn AT-9924T
- ▶ Dell PowerConnect 6024
- ▶ D-Link DXS-3350SR
- ▶ Extreme Networks Summit 400-48T
- ▶ HP ProCurve Switch 5308x
- ▶ Huawei Quidway S5024G
- ▶ SMC TigerSwitch 8648T

dann erforderlich, wenn eine festgelegte Reihe von Endgeräten gleichzeitig einen entsprechend vielfältigen Datenstrom erhalten soll. Dies ist beispielsweise bei der Videodistribution erforderlich. Für den Multicast-Betrieb haben wir das auf IP basierende Internet-Group-Management-Protocol, kurz IGMP, genutzt. Dieses ermöglicht es dem Switch im Trainingsverkehr zu »lernen«, an welche Ports er den Multicast-Traffic senden soll. Hat der Switch seine Verteilerliste gelernt, vervielfältigt er die eingehenden Datenströme und sendet sie parallel an die entsprechenden Zielports.

Im Multicast-Test haben wir zwei Multicast-Messreihen durchgeführt. Unsere erste Multicast-Messreihe diente lediglich dazu, die Multicast-Funktion zu verifizieren. Hierzu haben wir mit unserem Smartbits-Lastgenerator auf einen Port des zu testenden Switches zunächst mit 50 und dann mit 100 Prozent Last gesendet. Der jeweilige Switch im Test musste den Datenstrom vervielfältigen und an 22 Ausgangsports senden, die wiederum mit dem Smartbits-Analysator verbunden waren. Den 24-ten Gigabit-Ethernet-Port haben wir zur Kontrollmessung ebenfalls mit dem Smartbits-System verbunden, um sicher zu stellen, dass das Switch-System nicht in den Broadcast-Betrieb wechselt. In diesem Fall sendet er dann die Datenströme einfach an alle Ports. Unsere Messungen haben wir wieder nacheinander mit 64, 128, 256, 512, 1024, 1280 und 1518 Byte großen Frames durchgeführt. Dieser Testaufbau erzeugt keine Überlasten. Arbeitet ein Switch fehlerfrei und mit Wirespeed, sollte es zu keinerlei Datenverlusten kommen. Allerdings musste der jeweilige Switch im Test mit Wirespeed arbeiten. Flaschenhalse welcher Art auch immer führen hierbei zu unerwünschten Datenverlusten.

In der zweiten Multicast-Testreihe haben wir auf Layer-2 unterschiedlich priorisierte Datenströme erzeugt und jede der vier erzeugten Prioritäten an einen anderen Switch-Port gesendet. Die vier Switch-Ports, die die Datenströme jeweils einer Priorität empfangen, sollten dann den jeweiligen Datenstrom vervielfältigen und an 18 Switch-Ports senden. Die verbleibenden zwei Switch-Ports dienten wie oben der Kontrollmessung. So konnten wir ausschließen, dass der Switch im Test nicht anstatt Multicast-Broadcast-Betrieb fährt. Auch diese Messungen haben wir wieder mit 64, 128, 512, 1024 und 1518 Byte großen Frames durchgeführt. Die Last an den vier Eingangsports haben wir wie bei den vorher-

Info / alle Messergebnisse unter www.networkcomputing.de/nwc_downloads/messergebnisse_lanswitches_02_05.zip**Messergebnisse mit Strict-Priority-Switching Layer-2 / GE-GE Burst-Size 100** (Teil 1, Fortsetzung auf Seite 16)

| | Ausgangslast | Prio 7 | Prio 5 | Prio 3 | Prio 1 | Verlust gesamt |
|-----------------------|--------------|-------------|---------------|---------------|---------------|----------------|
| Allied Telesyn | | | | | | |
| Sollwert | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 64 Byte | 100% | 2,93 | 2,95 | 2,72 | 13,64 | 5,56 |
| 128 Byte | 100% | 0,00 | 0,00 | 0,00 | 1,64 | 0,41 |
| 512 Byte | 100% | 0,00 | 33,36 | 22,83 | 34,46 | 22,66 |
| 1024 Byte | 100% | 0,70 | 66,53 | 60,92 | 64,70 | 48,21 |
| 1518 Byte | 100% | 2,54 | 76,52 | 72,47 | 74,15 | 56,42 |
| Sollwert | 200% | 0,00 | 0,00 | 100,00 | 100,00 | 50,00 |
| 64 Byte | 200% | 0,17 | 1,72 | 99,73 | 100,00 | 50,40 |
| 128 Byte | 200% | 0,02 | 0,08 | 100,00 | 99,99 | 50,02 |
| 512 Byte | 200% | 0,00 | 33,39 | 71,32 | 99,99 | 51,18 |
| 1024 Byte | 200% | 0,13 | 66,58 | 61,37 | 81,10 | 52,30 |
| 1518 Byte | 200% | 2,48 | 76,47 | 72,42 | 74,31 | 56,42 |
| Sollwert | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 64 Byte | 400% | 33,27 | 66,96 | 99,89 | 100,00 | 75,03 |
| 128 Byte | 400% | 0,01 | 99,99 | 99,99 | 100,00 | 75,00 |
| 512 Byte | 400% | 0,13 | 99,99 | 100,00 | 100,00 | 75,03 |
| 1024 Byte | 400% | 3,44 | 99,99 | 100,00 | 100,00 | 75,86 |
| 1518 Byte | 400% | 0,67 | 99,99 | 100,00 | 100,00 | 75,16 |
| D-Link | | | | | | |
| Sollwert | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 64 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,11 | 0,03 |
| 128 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,05 | 0,01 |
| 512 Byte | 100% | 0,00 | 0,61 | 0,17 | 1,10 | 0,47 |
| 1024 Byte | 100% | 0,00 | 48,83 | 43,03 | 49,21 | 35,27 |
| 1518 Byte | 100% | 0,00 | 65,79 | 61,94 | 65,32 | 48,26 |
| Sollwert | 200% | 0,00 | 0,00 | 100,00 | 100,00 | 50,00 |
| 64 Byte | 200% | 0,00 | 0,12 | 99,99 | 99,99 | 50,03 |
| 128 Byte | 200% | 0,00 | 0,06 | 99,99 | 99,99 | 50,01 |
| 512 Byte | 200% | 0,00 | 0,65 | 99,39 | 99,97 | 50,00 |
| 1024 Byte | 200% | 0,00 | 48,87 | 51,44 | 99,67 | 50,00 |
| 1518 Byte | 200% | 0,00 | 65,80 | 61,72 | 72,51 | 50,01 |
| Sollwert | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 64 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 128 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 512 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 1024 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 1518 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| Extreme | | | | | | |
| Sollwert | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 64 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,13 | 0,03 |
| 128 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,07 | 0,02 |
| 512 Byte | 100% | 0,00 | 53,89 | 42,21 | 57,10 | 38,30 |
| 1024 Byte | 100% | 0,00 | 76,86 | 71,02 | 77,21 | 56,27 |
| 1518 Byte | 100% | 0,00 | 84,84 | 80,93 | 84,32 | 62,52 |
| Sollwert | 200% | 0,00 | 0,00 | 100,00 | 100,00 | 50,00 |
| 64 Byte | 200% | 0,00 | 0,13 | 100,00 | 100,00 | 50,03 |
| 128 Byte | 200% | 0,00 | 0,08 | 100,00 | 100,00 | 50,02 |
| 512 Byte | 200% | 0,00 | 53,95 | 50,24 | 95,82 | 50,00 |
| 1024 Byte | 200% | 0,00 | 76,84 | 71,05 | 77,21 | 56,28 |
| 1518 Byte | 200% | 0,00 | 84,84 | 80,93 | 84,32 | 62,52 |
| Sollwert | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 64 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 128 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 512 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 1024 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 1518 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |

gehenden Unicast-Tests wieder stufenweise von 25 bis auf 100 Prozent erhöht. Dieses Verfahren erzeugt eine maximal vierfache Überlastung der adressierten Ausgangsports am jeweiligen Switch. Bei diesem Testverfahren ist es völlig normal, dass der Switch im Test bei wachsender Überlast zunehmend Daten verwerft. In diesem Szenario kam es darauf an, dass der jeweilige Switch gemäß dem geforderten Strict-Priority-Mechanismus bei entstehender Überlast zunächst die Daten der niedrigsten Priorität, dann der zweitniedrigsten und zuletzt auch der zweithöchsten Priorität verwerft. In allen Fällen sollten die Daten der höchsten Priorität verlustfrei weitergeleitet werden. Um auch zu untersuchen, welche Rolle die Burst-Size auf das Verhalten der Switches spielt, haben wir die zweite Multicast-Testreihe zunächst mit einer Burst-Size von einem und dann von 100 Frames durchgeführt. Gemessen haben wir mit unseren Smartbits das Datenrahmenverlustverhalten sowie Latency und Jitter. Zur Auswertung herbeigezogen haben wir primär auf Grund seiner fundamentalen Bedeutung für die Übertragungsqualität das Datenrahmenverlustverhalten.

Allied Telesyns AT-9924T verhielt sich in der ersten Multicast-Messreihe recht unauffällig. Spürbare Datenverluste leistete er sich lediglich bei Volllast und großen Datenpaketen. So betrug bei der Messung mit 512-Byte-Frames der mittlere Datenverlust 3,63 Prozent, bei 1518-Byte-Paketen betrug der Mittelwert 3,84 Prozent. Den höchsten Mittelwert von 4,36 Prozent Frame-Loss erreichte der ATI-Switch bei der Messung mit 1024-Byte-Paketen. Allerdings streuten diese Verluste nicht gleichmäßig über alle Ports. Sechs Ports verloren bei der Messung mit 512-Byte-Paketen 13,31 Prozent der Daten. Bei allen übrigen Ports betrug der Frame-Loss null Prozent. Bei der Messung mit 1024-Byte-Paketen verloren die selben Flows dann jeder 15,97 Prozent der Daten, der Datenverlust der übrigen Flows blieb bei null. Und auch bei der Messung mit den noch größeren Frames bot sich das gleiche Bild. Durch diese Ungleichbehandlung der verschiedenen Datenströme würde ein Teil der Videostreams gestört. Die anderen kämen problemlos durch. Dieses Problem zeigte der AT-9924T allerdings ausschließlich bei Volllast, bei 50 Prozent Last konnten wir ihm keine Schwächen nachweisen.

Auch in unserer zweiten Multicast-Messreihe arbeitete Allied Telesyns AT-9924T recht unauffällig – so lange die Burst-Size 1 lautete. Lediglich bei der Messung mit Volllast und 1024-Byte-Paketen leistete sich der Switch einen ungebührlichen Datenverlust von 2,55 Prozent in der höchsten Priorität. Bei größeren und kleineren Frames blieb immer eine Null vor dem Komma. Betrug die Burst-Size 100 Frames, zeigte der AT-9924T die Grenzen seiner Belastbarkeit. Schon bei einer Last von 100 Prozent, als theoretisch noch gar keine Datenverluste notwendig gewesen wären, leistete sich der Switch bei den Messungen mit dem größten Frame-Format einen durchschnittlichen Datenverlust von über 56 Prozent. Das Maximum lag dabei mit 76,52 Prozent in der zweithöchsten Priorität. Lediglich die höchste Priorität blieb hier mit 2,54 Prozent relativ »sauber«. Mit kleineren Frames kam der ATI-Switch besser zurecht. Der Mittelwert der Datenverluste lag aber auch noch bei der Messung mit 512-Byte-Frames bei 22,66 Prozent. Auch bei 200 Prozent Last, als der

Info / alle Messergebnisse unter www.networkcomputing.de/nwc_downloads/messergebnisse_lanswitches_02_05.zip

Messergebnisse mit Strict-Priority-Switching Layer-2 / GE-GE Burst-Size 100 (Teil 2, Fortsetzung von Seite 14)

| | Ausgangslast | Prio 7 | Prio 5 | Prio 3 | Prio 1 | Verlust gesamt |
|-----------------|--------------|-------------|---------------|---------------|---------------|----------------|
| HP | | | | | | |
| Sollwert | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 64 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,10 | 0,03 |
| 128 Byte | 100% | 0,00 | 0,03 | 0,00 | 0,00 | 0,01 |
| 512 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1024 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1518 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| Sollwert | 200% | 0,00 | 0,00 | 100,00 | 100,00 | 50,00 |
| 64 Byte | 200% | 43,73 | 43,73 | 43,74 | 68,51 | 49,93 |
| 128 Byte | 200% | 5,36 | 5,36 | 88,96 | 100,00 | 49,92 |
| 512 Byte | 200% | 0,00 | 62,71 | 53,33 | 83,98 | 50,00 |
| 1024 Byte | 200% | 0,00 | 65,84 | 60,78 | 73,23 | 49,96 |
| 1518 Byte | 200% | 0,00 | 66,73 | 63,32 | 69,72 | 49,94 |
| Sollwert | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 64 Byte | 400% | 71,80 | 71,80 | 71,80 | 84,38 | 74,95 |
| 128 Byte | 400% | 52,69 | 52,69 | 94,41 | 100,00 | 74,95 |
| 512 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 1024 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 1518 Byte | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| Huawei | | | | | | |
| Sollwert | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 64 Byte | 100% | 92,86 | 28,57 | 28,57 | 92,86 | 60,71 |
| 128 Byte | 100% | 92,86 | 28,57 | 28,57 | 92,86 | 60,71 |
| 512 Byte | 100% | 92,86 | 28,57 | 28,57 | 92,86 | 60,71 |
| 1024 Byte | 100% | 92,86 | 49,44 | 46,51 | 92,94 | 70,44 |
| 1518 Byte | 100% | 92,86 | 66,13 | 56,01 | 95,23 | 77,56 |
| Sollwert | 200% | 0,00 | 0,00 | 100,00 | 100,00 | 50,00 |
| 64 Byte | 200% | 92,86 | 28,67 | 28,58 | 100,00 | 62,53 |
| 128 Byte | 200% | 92,86 | 28,62 | 28,58 | 100,00 | 62,51 |
| 512 Byte | 200% | 92,86 | 28,59 | 28,57 | 100,00 | 62,50 |
| 1024 Byte | 200% | 92,86 | 49,43 | 46,52 | 96,12 | 71,23 |
| 1518 Byte | 200% | 92,86 | 66,16 | 55,98 | 95,24 | 77,56 |
| Sollwert | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 64 Byte | 400% | 92,86 | 100,00 | 28,57 | 100,00 | 80,36 |
| 128 Byte | 400% | 92,86 | 100,00 | 28,57 | 100,00 | 80,36 |
| 512 Byte | 400% | 92,86 | 100,00 | 28,57 | 100,00 | 80,36 |
| 1024 Byte | 400% | 92,86 | 100,00 | 28,57 | 100,00 | 80,36 |
| 1518 Byte | 400% | 92,86 | 100,00 | 28,57 | 100,00 | 80,36 |
| SMC | | | | | | |
| Sollwert | 100% | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 64 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,12 | 0,03 |
| 128 Byte | 100% | 0,00 | 0,00 | 0,00 | 0,06 | 0,02 |
| 512 Byte | 100% | 0,00 | 29,63 | 18,52 | 33,11 | 20,31 |
| 1024 Byte | 100% | 0,00 | 64,71 | 59,26 | 65,22 | 47,30 |
| 1518 Byte | 100% | 0,00 | 76,72 | 73,09 | 76,33 | 56,53 |
| Sollwert | 200% | 0,00 | 0,00 | 100,00 | 100,00 | 50,00 |
| 64 Byte | 200% | 0,00 | 0,13 | 100,00 | 100,00 | 50,03 |
| 128 Byte | 200% | 0,00 | 0,07 | 99,99 | 99,99 | 50,01 |
| 512 Byte | 200% | 0,00 | 29,64 | 70,37 | 99,99 | 50,00 |
| 1024 Byte | 200% | 0,00 | 64,72 | 59,03 | 76,28 | 50,01 |
| 1518 Byte | 200% | 0,00 | 76,72 | 73,08 | 76,32 | 56,53 |
| Sollwert | 400% | 0,00 | 100,00 | 100,00 | 100,00 | 75,00 |
| 64 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 128 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 512 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 1024 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |
| 1518 Byte | 400% | 0,00 | 99,99 | 100,00 | 100,00 | 75,00 |

Switch die beiden niedrigen Prioritäten zu Gunsten der beiden hohen Prioritäten verwerfen sollte, leistete sich der AT-9924T deutliche Schwächen im Handling mit großen Datenpaketen. So verlor der Switch bei der Messung mit den größten Frames in den beiden niedrigen Prioritäten nicht 100, sondern lediglich gut 74 beziehungsweise über 72 Prozent. Diese zu niedrigen Verluste gingen zu Lasten der zweithöchsten Priorität. Hier waren gut 76 Prozent Frame-Loss zu messen, die nicht erforderlich gewesen wären. Lediglich die höchste Priorität kam mit 2,48 Prozent Datenverlust verhältnismäßig ungeschoren davon. Bei den Messungen mit kleineren Frame-Formaten schwächte sich aber auch hier das Problem ab. Bei den Volllastmessungen kam dann noch ein Problem mit den kleinen 64-Byte-Paketen hinzu. Hier verlor der AT-Switch rund 33 Prozent aller Daten in der höchsten Priorität.

Dells Powerconnect-6024 weigerte sich standhaft im Multicast-Modus zu arbeiten. Diese Funktionalität ließ sich zwar konfigurieren, die Kontrollmessungen haben aber in allen Fällen ergeben, dass der Dell-Switch anstatt seine Zielpriorität abzuarbeiten einfach per Broadcast-Modus an alle Ports sendete. Durch dieses Verhalten erhalten auch nicht adressierte Ports – und somit im Netzwerk Endgeräte – Datenströme, die sie gar nicht bekommen sollen. Daher mussten wir den Dell-Switch disqualifizieren und in der Report-Card entsprechend abwerten.

D-Links DXS-3350SR ließ sich dagegen bei der ersten Multicast-Messreihe keinerlei Schwächen nachweisen. Das Gerät leitete alle Flows mit Wirespeed an die richtigen Ports, ohne Datenverluste zu erzeugen. Auch in unserer zweiten Multicast-Messreihe arbeitete das D-Link-Produkt bei unseren Messungen mit einer Burst-Size von 1 völlig unauffällig. Es verwarf die Frames gemäß gewünschter Policy und leitete die übrigen Daten wie gewünscht weiter. Betrug die Burst-Size 100 Frames, dann vergaß auch der D-Link-Switch seine guten Sitten. Sein Verhalten erinnerte an den Switch von Allied Telesyn, war aber nicht ganz so stark ausgeprägt. Diese Fakt hat ihm in unserer Report-Card rechnerisch einen Vorteil gesichert. So verlor der DXS-3350SR schon bei 100 Prozent Last und den größten Frames im Mittel 48,26 Prozent aller Daten, obwohl hier noch gar keine Datenverluste erforderlich gewesen wären. Bei der Messung mit 200 Prozent Last und der Verwendung der größten Frames teilten sich die zweithöchste, die zweitniedrigste und die niedrigste Priorität wieder brüderlich die Datenverluste. Lediglich in der höchsten Priorität blieben die Daten in allen Fällen frei von ungebührlichen Datenverlusten.

Der Summit-400-48T von Extreme Networks leistete sich in der ersten Multicast-Disziplin keine Schwächen. Er verteilte alle Streams wie zuvor gelernt und frei von Datenverlusten mit Wirespeed an die vorgesehenen Ports. Auch in unserer zweiten Multicast-Messreihe zeigte sich der Summit-400-48T von seiner besten Seite und priorisierte tadellos. Diese Aussage gilt aber wie schon zuvor ausschließlich für die Messungen mit einer Burst-Size von 1. Betrug diese 100, ähnelt das Bild den Ergebnissen des Allied-Telesyn- oder des D-Link-Switches. So kam der Extreme-Networks-Switch bei der Messung mit den größten Frames und einer Last von 100 Prozent auf eine Verlustrate von nicht erforderlichen 62,52 Prozent. Mit kleineren Frame-Formaten kam der Switch

wie seine Kollegen zuvor deutlich besser zurecht. Allerdings kam das System auch bei der Messung mit 512-Byte-Paketen noch auf einen ungebührlichen Mittelwert von 38,30 Prozent Frame-Loss. Lediglich die Daten der höchsten Priorität passierten den Switch in allen Fällen völlig ungehindert. Bei der Messung mit 200 Prozent Last und großen Frames konnten wir dann wieder das Phänomen beobachten, dass der Switch die Datenverluste nicht wie vorgesehen zwischen den beiden niedrigen Prioritäten, sondern zwischen den Prioritäten 1, 3 und 5 aufteilte.

HPs Procurve-Switch-5308x erwies sich zunächst als Musterschüler. Er arbeitete wie gewünscht mit Wirespeed und leistete sich keine Datenverluste in der ersten Multicast-Messreihe. In der zweiten Multicast-Messreihe war dann schon bei unseren Messungen mit einer Burst-Size von 1 eine Schwäche im Handling mit kleinen Datenrahmen zu beobachten. Bei 200 Prozent Last und 64-Byte-Rahmen betrug der Frame-Loss in der höchsten, zweithöchsten und zweitniedrigsten Priorität jeweils 43,7 Prozent. In der niedrigsten Priorität betrug der Datenverlust hierbei gut 68 Prozent. Sendeten wir 128 Byte große Datenpakete, so kam die Verlustrate in der höchsten und zweithöchsten Priorität noch auf ungebührliche 5,38 Prozent. Bei Vollast und kleinen Frames waren die Verluste dann noch höher. So betrug die Verlustrate bei der Messung mit 64-Byte-Frames 71,81 und mit 128-Byte-Frames noch 52,7 Prozent. Verwendeten wir größere Frames, war die Priorisierungswelt wieder in Ordnung. Betrug die Burst-Size 100 Frames, so kam zu der beschriebenen Schwäche ein Problem mit der Aufteilung der Datenverluste zwischen der zweithöchsten und den niedrigeren Prioritäten zu Tage. So waren bei den Messungen mit 200 Prozent und größeren Frames durchweg unerwünschte Datenverluste in der zweithöchsten Priorität von über 60 Prozent zu messen.

Probleme mit unserem ersten Multicast-TestszENARIO hatte Huawei's Quidway-S5024G.

Der Switch aus Fernost »vergaß« immer wieder einzelne Ports zu beliefern, was sich messtechnisch an diesen Ports als 100 Prozent Frame-Loss auswirkte. Dieses Phänomen war sowohl bei 50 als auch bei 100 Prozent Last zu beobachten. Von diesem Verhalten war bei den Messungen mit 64- und 128-Byte-Frames jeder zweite Flow betroffen, was einen Mittelwert der Datenverluste von 50 Prozent bedeutet. Ab der Messung mit 128-Byte-Paketen reduzierte sich dann die Zahl der

Info

Real-Time-Switches in Network Computing

- ▶ Huawei Quidway S5024G, Schnell und intelligent im Netz, in NWC Special Networking 2005, S.4 ff.
- ▶ Gigabit-Ethernet-Switches, Schnell und intelligent im Netz, in NWC 4/05, S.14 ff.
- ▶ Gigabit-Ethernet-Switches, Schnell und intelligent im Netz – die Theorie, in NWC 4/05, S.52 ff.
- ▶ Extreme Networks BlackDiamond 10808, Der Netzwerkdiamant, in NWC 19/04, S.20 ff.
- ▶ Gigabit-Ethernet-Switches, Switch as Switch can, in NWC Special 6/04 Infrastruktur, S.4 ff.

»vergessenen« Ports auf ein Verhältnis von 2 von 18. Dabei war das Fehlverhalten wie auch schon im Fall Allied Telesyn statisch. Benachteiligt oder präziser gar nicht beliefert wurden immer wieder die gleichen Ports, die daher in unserem Szenario unbrauchbar waren.

Ein merkwürdiges Verhalten können wir Huawei's Quidway-S5024G auf Grund der ermittelten Messergebnisse attestieren. So betrug der Frame-Loss bei unseren Messungen mit Burst-Size 1 beziehungsweise 100 und 200 Prozent Last durchweg 7,14 Prozent. Bei 100 Prozent

kamen dann noch nicht erforderliche 21,43 Prozent Datenverluste in der zweitniedrigsten und 14,29 Prozent Datenverluste in der niedrigsten Priorität hinzu. Lediglich die höchste Priorität kam immer ungeschoren davon. Mit einer Ausnahme, bei unserer Messung mit Vollast und den größten Frames schnellte der Frame-Loss-Wert dann plötzlich auf 62,48 Prozent wohlgermerkt in der höchsten Priorität hoch. Völlig aus der Fassung geriet der Quidway-S5024G dann bei unseren Messungen mit einer Burst-Size von 100 Frames. So kam der Huawei-Switch hier durchgängig auf Datenverluste in der höchsten Priorität von 92,86 Prozent.

SMCs Tigerswitch-8648T zeigte sich dagegen in unserem ersten Multicast-TestszENARIO von seiner besten Seite und leistete sich keinerlei Schwächen. Auch in unserem zweiten Multicast-TestszENARIO bei den Messungen mit einer Burst-Size von 1 arbeitete der Tigerswitch-8648T ohne Fehl und Tadel. Er erzielte exakt die gewünschten Durchsatzraten. Probleme bekam aber auch der SMC-Switch bei unseren Messungen mit einer Burst-Size von 100 und größeren Datenrahmen. So gingen bei 100 wie auch bei 200 Prozent Last in der zweithöchsten Priorität bei der Messung mit 512-Byte-Pakete unnötigerweise rund 30 Prozent, bei der Messung mit 1024-Byte-Paketen rund 65 und bei der Messung mit den größten Frame fast 77 Prozent der Daten verloren. Allerdings zeigten sich die Daten der höchsten Priorität von diesem Problem stets unbeeinträchtigt. Sie passierten den Switch planmäßig und ungehindert.

Vagabundierende Daten

Stray-Frames, auf Deutsch »vagabundierende« Datenrahmen, sind vom Switch fehlgeleitete Datenpakete. Diese sendet der betroffene Switch nicht oder nicht nur zum adressierten Zielport, sondern (auch) zu anderen Ports. Für den »richtigen« Zielport sind solche vagabundierenden Daten mit Frame-Loss gleichzusetzen, sofern die-

se umgeleitet wurden. Sind die Stray-Frames auf unerwünschte Multicast-Sendungen zurückzuführen, erhöht diese Fehlfunktion die Belastung des Switches. Außerdem erhalten Endgeräte Daten, die gar nicht für sie bestimmt sind. Ab einer gewissen Fehlerrate ist mit Störungen der Kommunikation durch Stray-Frames zu rechnen. Aber auch wenn der Anteil der fehlgeleiteten Frames deutlich unter einem Prozent bleibt und es somit nicht zu spürbaren Kommunikationsstörungen kommt, bleibt ein Switch, der Daten falsch ausliefert, ein unkalkulierbares Sicherheitsrisiko.

Bei allen Messungen, die wir an den Switches im Testfeld vorgenommen haben, haben wir gleichzeitig auch mit unseren Smartbits erfasst, ob Stray-Frames zu ermitteln waren. Die folgende Auswertung bezieht sich daher sowohl auf die Messungen mit Punkt-zu-Punkt-Verbindungen (siehe NWC Special Networking 2005, S.4 ff. und NWC 4/05, S.14 ff.), wie auch auf die oben dargestellten Multicast-Tests. Die Smartbits haben die absolute Zahl der Stray-Frames ermittelt. Wir haben die Zahl der fehlgeleiteten Frames dann in das prozentuale Verhältnis zu den insgesamt transportierten Daten gesetzt.

In unserem Testfeld waren bei drei Switches Stray-Frames festzustellen. Huawei's Quidway-S5024G sandte vergleichsweise wenige Frames an die falschen Ports. Insgesamt konnten wir bei drei Messungen 432 Stray-Frames registrieren. 360 entfielen auf unsere Messungen mit Layer-2-Priorisierung, Burst-Size 1 und Bandbreitenmanagement. Jeweils 36 Stray-Frames entfielen dann auf unsere Messungen mit Layer-3-Priorisierung, Burst-Size 1 und Bandbreitenmanagement beziehungsweise Strict-Priority-Switching.

Um eine Zehnerpotenz höher lag die Zahl der fehlgeleiteten Datenpakete bei SMCs Tiger-switch-8648T mit insgesamt 5520 gemessenen Stray-Frames. Dabei waren die höchsten Fehleraten bei Messungen mit einer Burst-Size von 100 Frames festzustellen. 1308 Datenrahmen leitete der SMC-Switch bei unserer Messung mit Layer-3-Priorisierung und Strict-Priority-Switching fehl. 1287 Stray-Frames waren dann bei der Messung mit Layer-3-Priorisierung, Burst-Size 100 und Bandbreitenmanagement zu verzeichnen. Bei der gleichen Messung mit Layer-2-Priorisierung leitete der SMC-Switch noch 1269 Frames falsch. Bei den Messungen mit einer Burst-Size von einem Frame pendelte der Wert für Stray-Frames dann um die 400.

Das vergleichsweise größte Problem mit vabundierenden Daten hatte mit Abstand Dells Powerconnect-6024. Er sandte insgesamt 37 947 Frames an falsche Adressen. Dieses Problem bestand wie auch beim SMC-Switch in erster Linie bei den Messungen mit einer Burst-Size von 100 Frames. Die größte Zahl der Fehlleitungen konnten wir bei der Messung mit Layer-3-Priorisierung und Strict-Priority-Switching feststellen. Hier gingen 13 050 Datenpakete an die falsche Adresse. Bei der gleichen Messung mit Layer-2-Priorisierung erreichten immerhin noch 9739 Pakete einen falschen Port. Mit Bandbreitenmanagement und einer Burst-Size von 100 kam der Dell-Switch dann noch auf Stray-Frame-Werte von 6715 bei Layer-2- und 6855 bei Layer-3-Priorisierung. Die einzige Messung mit dem Powerconnect-6024, die Stray-Frames bei einer Burst-Size von 1 erzeugte, war die Messung mit Bandbreitenmanagement und Layer-2-Priorisierung. Prozentual ist die Anzahl der gemessenen Stray-

Frames in Relation zu den richtig verteilten recht gering. So erreichten beim Dell-Switch insgesamt rund 0,013 Prozent der in den jeweiligen Messungen mit Stray-Frames gesendeten Daten falsche Ports. Bei den Switches von Huawei und SMC war der Anteil der Stray-Frames noch deutlich geringer. Punkteabzug gab es trotz dieser geringen Zahl, weil die ermittelten Stray-Frames zwar keine nennenswerten Kommunikationsstörungen verursacht hätten aber falsch ausgelieferte Datenpakete generell ein Sicherheitsproblem darstellen.

Fazit

Es gibt weder gute noch schlechte Switches. Dies zeigt ein Vergleich der Report-Card, die die Unicast-Messungen mit Strict-Priority und Bandbreitenmanagement bewertet (NWC Special Networking, S.4 oder unter www.networkcomputing.de), mit der hier vorgelegten Report-Card, die nun die Multicast-Messungen und die Stray-Frames würdigt. Auch wenn D-Link in beiden Tests die Nase vorn hat, haben doch einige Hersteller die Plätze im Ranking gewechselt. Dies liegt daran, dass jeder Switch seine konzeptionellen Stärken und Schwächen hat und es immer noch Fehler im Design von Hard- und Software gibt. Hier haben noch lange nicht alle Hersteller ihre Hausaufgaben gemacht.

Wenn bei den Messungen mit einer Burst-Size von 1 die Ergebnisse weitgehend in Ordnung sind, bei einer Burst-Size von 100 aber deutliche Datenverluste festzustellen sind, dann liegt die Ursache zumeist in zu knapp bemessenen Pufferspeichern. Sind die Pufferspeicher zu klein dimensioniert, dann führen entsprechend hohe Lasten im Strict-Priority-Betrieb zwangsläufig dazu, dass die Daten der niedrigeren Prioritäten schon frühzeitig gar nicht mehr transportiert werden, weil die höchste Priorität absoluten Vorrang hat. Die Problematik lässt sich vermeiden, indem man ein intelligentes Bandbreitenmanagement im gesamten Netzwerk konfiguriert und so die Flaschenhalse in den Edge-Bereich verlagert. Dies setzt aber intensive Analysen und Konfigurationsarbeiten sowie ein Ende-zu-Ende-Management voraus, so dass die verfügbare Bandbreite möglichst effizient genutzt werden kann. Andernfalls kann Bandbreitenmanagement dazu führen, dass im Edge-Bereich die eigentlich verfügbare Wirespeed vielfach nicht genutzt werden kann oder für bestimmte Applikationen nicht genügend Bandbreite zur Verfügung steht. Da im realen Netzbetrieb Burst-Größen von 100 Frames keine Seltenheit sind, sollten die Switch-Hersteller ihren Systemen besser größere Pufferspeicher spendieren.

Generell gilt weiterhin der Rat, grundsätzlich dafür zu sorgen, dass es im Netzwerk erst gar nicht eng wird, und die Systeme auszuwählen, die für die gewünschte Policy auch geeignet sind. Bei größeren Netzwerkprojekten kommen die Verantwortlichen dann eigentlich nicht darum herum, die benötigten Eigenschaften der in Frage kommenden Systeme in qualifizierten Tests zu überprüfen, wenn sie auf Nummer Sicher gehen wollen. Denn Vertrauen ist gut aber Kontrolle ist besser.

Prof. Dr. Bernhard G. Stütz, [dg]

Info

So testete Network Computing

Als Lastgenerator und Analyser haben wir in unseren Real-World Labs einen »Smartbits 6000B Traffic Generator/Analyzer« von Spirent Communications eingesetzt. Das in dieser Konfiguration rund 320000 Euro teure System ist mit der Software »SmartFlow« ausgestattet und mit 24 Gigabit-Ethernet-Kupfer-Ports bestückt. Alle Ports können softwareseitig als Lastgeneratortausgang und/oder als Analyseringang 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-2 nach IEEE 802.1p/Q untersucht. In unseren Testscenarien »Gigabit-Ethernet-Switches« haben wir verschiedene priorisierte Datenströme auf die Eingangsports gesendet und auf die Ausgangsports des Switches geschickt. Dort haben wir mit den Smartbits die entsprechenden Flows erfasst und ausgewertet. Die die Priorisierung festlegenden Bits haben wir im Header der Datenrahmen mit drei Bits nach IEEE 802.1p und auf Layer-2 und nach ToS beziehungsweise Diffserv auf Layer-3 festgelegt. Durch eine gezielte Überlastung der Switches in diesen Tests ist es möglich, das genaue Datenverlustver-



halten sowie weitere Testparameter wie Latency oder Jitter zu ermitteln, das Leistungspotential der untersuchten Switches zu analysieren und deren Eignung für bestimmte Einsatzszenarien zu prüfen.

