



Pilottest Class-of-Service-Queuing

Vorfahrt im Netz

QoS- beziehungsweise CoS-Verfahren sollen besonders wichtigen und spezielle Übertragungsqualitäten fordernden Anwendungen die Vorfahrt im Netzwerk einräumen. In einem Pilottest an unseren Real-World Labs an der FH Stralsund haben wir die Implementierung der verschiedenen Queuing-Strategien näher untersucht.

Wird es im Ethernet-LAN einmal eng, dann kommt es je nach Aus- und Überlastung zu teils erheblichen Datenverlusten. Im Zeitalter konvergenter Netze, die Echtzeitanwendungen wie Voice- und Video-over-IP aber auch Produktionssteuerungsdaten in das klassische Datennetz integrieren, führen solche Datenverluste in der Praxis zu empfindlichen Kommunikationsstörungen und Produktionsausfällen. Um solchen Problemen vorzubeugen stellen die Ethernet-Hersteller ihre Switches mit einer zusätzlichen Funktionalität aus, die es ermöglichen soll, bestimmten Applikationen die Vorfahrt im Netzwerk einzuräumen, wenn es einmal eng wird. Diese Priorisierungs-Mechanismen werden allgemein als Class-of-Service oder – verfälschend in Anlehnung an ATM – als Quality-of-Service bezeichnet. Üblich ist eine achsstufige Priorisierung auf Layer-2 nach IEEE 802.1p/Q oder auf Layer-3 nach RFC 1349/2474/2475. Die jeweils zugeordnete Priorisierung lesen die Systeme aus den Headern der Datenpakete aus.

Wie Switches die Datenpakete dann gemäß ihrer Priorität behandeln, hängt von den jeweils implementierten Queuing-Mechanismen ab. So besteht beispielsweise die Möglichkeit, bestimmten Daten absolute Vorfahrt einzuräumen oder auch für niedrigere Prioritäten Mindestdurchsatzraten zu garantieren. Eine gute Queuing- oder Scheduling-Strategie sollte folgende Voraussetzungen erfüllen:

- ▶ Sie muss die faire Verteilung der Bandbreite auf die verschiedenen Serviceklassen unterstützen. Dabei sollte auch die Bandbreite für besondere Dienste berücksichtigt werden, so dass es zu bestimmten Gewichtungen bei der Fairness kommen kann.
- ▶ Sie bietet Schutz zwischen den verschiedenen Serviceklassen am Ausgangsport, so dass eine Serviceklasse mit geringer Priorität nicht die anderen Serviceklassen anderer Queues beeinflussen kann.
- ▶ Wenn ein Dienst nicht die gesamte Bandbreite verwendet, die für ihn reserviert ist, dann sollte diese Überkapazität auch anderen Diensten zur Verfügung stehen, bis der eigentliche Dienst diese Kapazitäten wieder benötigt.

▶ Ein schneller Algorithmus, der hardwaremäßig implementiert werden kann, muss für die Strategie existieren. Nur dann kann diese Strategie auch auf Switches eingesetzt werden, die mit hoher Geschwindigkeit arbeiten. Algorithmen, die nur softwareseitig implementiert werden können, eignen sich lediglich für sehr langsame Switches.

Die Testkandidaten

Wir wollten wissen, welche Queuing-Mechanismen heute in aktuellen LAN-Switches wirklich implementiert sind und wie gut sie derzeit arbeiten. Aus diesem Grund haben wir ein Feld von Class-of-Service-fähigen LAN-Switches in unseren Real-World Labs an der FH Stralsund einer umfangreichen Testprozedur unterzogen. Für die verschiedenen Tests standen insgesamt fünf verschiedene Testkandidaten zur Verfügung, die zwei Geräteklassen angehören. Bei den Geräten von D-Link und Netgear handelte es sich um Switches mit Gigabit-Ethernet-Ports. Die Geräte von Extreme Networks, Hewlett-Packard und BATM waren mit Fast-Ethernet-Ports ausgestattet.

Bei dem »D-Link DES 6300 Layer3 Chassis Switch« handelt es sich um einen Layer-3-Switch mit verschiedenen Modulen. Das Testgerät hatte ein Modul mit 24 Ports `a 100 MBit/s, ein Modul mit vier Gigabit-Ports im RJ45-Format (Cu-Kabel) und drei Module mit je zwei Gigabit-Ports, die Lichtwellenleiter für die Datenübertragung nutzten. Für die Tests wurden ausschließlich fünf Ports der LWL-Module mit der Bezeichnung »DES 6306 2 Port 1000 Base-SX Switch Module« verwendet. Die Konfiguration des Systems wurde über das Modul »System Modul DES-6302« mit der seriellen Schnittstelle und einmal via 100-MBit/s-Modul erstellt. Die Konfiguration mittels serieller Schnittstelle erfolgte über das Programm Hyperterminal. Für die Konfiguration über Ethernet kam das Programm ConfigMaster 8.4.1 (Build 30-09-2002) zum Einsatz. Die Firmware des Switches lag in der Version 1.1.31 vor. Bei diesem Testgerät wurden nach Aussage des Herstellers vier Queues implementiert, die

entweder ein WRR-ähnliches Verfahren oder Strict-Priority-Queuing verwenden. Weiterhin sicherte D-Link Deutschland zu, dass pro Modul 16 MByte Speicher zu Verfügung stehen. Die Ports eines Moduls sollten sich diesen Speicher teilen und dynamisch verwalten.

Der »Netgear 24-Port L3 Managed Gigabit Switch GSM 7324« ist ein Switch mit 24 Kupfer-Ports (RJ45), die eine Übertragungskapazität von je einem GBit/s bieten. Die letzten vier der 24 Kupfer-Ports konnten auch über je einen GBIC-Port mittels Adapter auf Lichtwellenleiter gelegt werden. Bei

den Tests haben wir vier Kupferkabel und einen Lichtwellenleiter verwendet. Der Testkandidat lief mit der Firmware 1.0.0.12. Er wurde auch via Hyperterminal auf Kommandozeilenebene und direkt über das integrierte

Web-Interface konfiguriert.

Nach Angaben des Herstellers werden für den GSM-7324 vier

Hardware-queues verwendet, die nach einem WFQ-Algorithmus arbeiten. Für jede Queue wurden dafür 256 kByte Speicher

reserviert. Bei 24 Ports, die in dem Gerät verbaut waren, ergibt das einen Speicher von insgesamt 24 Megabyte.

Erster Testkandidat im Fast-Ethernet-Segment war der »T5C-L3« der Firma BATM. Dieser Switch ist ein Gerät mit 24 Kupfer-Ports (RJ45), die mit 100 MBit/s arbeiteten. Die Untersuchungen für Layer-2 haben wir mit der Firmware 4.6.0.5.0.0 und der Java-Version 3.44 gemacht. Für einen Nachttest, den Frame-Loss-Tests auf Layer-2 stand dann die Firmware in der Version 4.6.0.6.4.1 zur Verfügung. Die Untersuchungen auf Layer-3 konnten nicht durchgeführt werden, da im Anschluss an die Untersuchungen von Layer-2 ein Fehler in der Software des Switches festgestellt wurde. Dieser Fehler konnte bis zum Redaktionsschluss noch nicht behoben werden, weswegen dieser Test erst zu einem späteren Zeitpunkt nachgeholt werden kann. Die Konfiguration des Switches hat ein Mitarbeiter von BATM über Kommandozeile vorgenommen. BATM bescheinigte ihrem Switch eine Speichergröße, die 16 Pakete pro Queue aufnehmen kann. Bei der maximalen Framelänge von 1518 Byte ergibt das einen Wert von 23,7 Kilobyte. Da der Switch pro Port 8 Queues verwalten konnte und der Switch 24 Ports besaß, ergibt das einen gesamten Speicher von 4,45 MByte. Die Queues konnten dann mit einem veränderten WRR und Strict-PQ als Queuing-Strategie eingesetzt werden.

Auch der »Summit 48 SI« von Extreme Networks konnte nur mit 100 MBit/s getestet werden. Er war mit 48 100-MBit/s-Ports ausgestattet und lief mit der Firmware-Version 7.0.1b13. Die Konfiguration des Testteilnehmers haben wir wieder via Hyperterminal auf der Kommandozeilenebene sowie via Ethernet durchgeführt. Für den Zugriff via



Info

Das Testfeld

- ▶ BATM T5C-L3
- ▶ D-Link DES 6300 Layer3 Chassis Switch
- ▶ Extreme Networks Summit 48 SI
- ▶ HP Procurve Switch 5304 XL j4850A
- ▶ Netgear 24-Port L3 Managed Gigabit Switch GSM 7324

Ethernet war aber kein zusätzliches Programm notwendig. Hier konnten wir direkt auf den internen Web-Server des Gerätes zugreifen, indem wir einen normalen Web-Browser verwendeten. Das Gerät arbeitete sowohl mit WFQ, als auch mit Strict-PQ und besaß insgesamt 4 MByte Speicher, die – laut Hersteller – dynamisch verwaltet werden. Werden also alle Ports und Queues verwendet, so ergibt das einen Speicher von etwa nur 21,3 kByte für jede Queue. Das entspräche also dem Platz, den 14 Pakete mit einer Framelänge von 1518 Byte benötigen. Beim Test haben wir aber nur die vier Queues des Ausgangsports verwendet. Damit ergibt sich, dass beim Test nur Queues verwendet wurden, die durchschnittlich 1 MByte Speicher reserviert hatten.

Auch der HP-Switch ist ein Testkandidat, der nur mit 100-MBit/s-Ports ausgestattet war. Die ausführliche Bezeichnung des Gerätes lautet »Procurve Switch 5304 XL j4850A«. In dem Switch waren zwei identische Einschub-Module mit je 24 Ports mit der Bezeichnung »Procurve 10/100-TX XL-Module j4820A« integriert. Auch hier haben wir die Konfiguration wieder über Kommandozeile und Web-Interface des Switches vorgenommen. Zum Zeitpunkt der Tests unterstützte dieser Test-Switch nur WFQ als Queuing-Strategie. Ein Firmware-Update, das erst im März 2004 veröffentlicht werden soll, soll dann aber auch Strict-Priority-Queuing unterstützen. Informationen zu den Speichergrößen der einzelnen Queues waren bisher noch nicht zu beziehen.

Queuing-Mechanismen

Die Hersteller von Switches verwenden oft eigene Namen für die Queuing-Strategien oder ändern die eigentliche Strategie nach ihren Wünschen ab. Oft werden auch verschiedene Strategien miteinander kombiniert, um die Ergebnisse zu verbessern. Die ursprünglichen Queuing-Strategien sind:

- ▶ First-In First-Out (FIFO),
- ▶ Priority-Queuing (PQ),
- ▶ Fair-Queuing (FQ),
- ▶ Weighted-Fair-Queuing (WFQ),
- ▶ Weighted-Round-Robin-Queuing (WRR), auch als Class-Based-Queuing (CBQ) bezeichnet, und
- ▶ Deficit-Weighted-Round-Robin-Queuing (DWRR).

Für unser Testfeld waren nur Strict-Priority-Queuing und Weighted-Fair-Queuing beziehungsweise Weighted-Round-Robin-Queuing relevant, weshalb diese Strategien hier kurz beleuchtet werden sollen.

Priority-Queuing (PQ) ist die Grundlage aller Scheduling-Strategien, die eine Unterteilung von verschiedenen CoS-Klassen ermöglichen. Dazu werden bei PQ für jede CoS-Klasse eine Queue angelegt und verwaltet. Die einzelnen Queues werden nach dem FIFO-Prinzip abgearbeitet. CoS wird dabei erreicht, indem die Frames einer Queue erst dann abgearbeitet werden, wenn die Queues höherer Priorität keine Frames mehr speichern. Strict-PQ sichert ab, dass die Queue höherer Priorität immer vor den Queues geringerer Priorität abgehandelt werden. Dabei besteht natürlich immer die Gefahr, dass ein relativ großer Datenstrom dieser Priorität die Datenströme geringerer Priorität ausbremst und sogar ganz »verhungern« lässt. Das kann aber auch durchaus gewünscht sein, wenn zum Beispiel Echt-

zeitdienste oder Dienste zur Netzwerkverwaltung auf der hohen Priorität liegen. Diese Dienste müssen dann auf jeden Fall weitergeleitet werden, egal, wie belastet das Netzwerk ist. Strict-PQ verwendet dabei auch Bandbreite für die hohe Priorität, die eigentlich für andere CoS-Klassen reserviert ist.

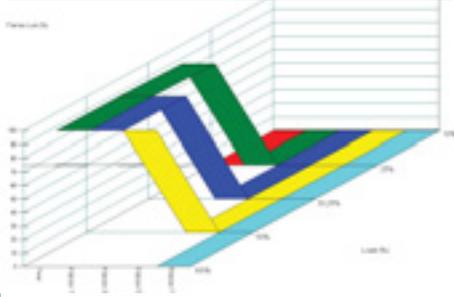
Weighted-Fair-Queuing (WFQ) unterstützt Datenströme mit verschiedenen Bandbreiten, indem es jeder Queue eine Gewichtung zuteilt, die ein Prozentsatz der gesamten Bandbreite des Ausgangsports ist. Außerdem unterstützt WFQ auch Datenströme mit verschiedenen langen Frames. Dadurch bekommen Datenströme mit langen Frames nicht mehr Bandbreite als Datenströme mit kurzen Frames. WFQ ist an den Endpunkten von Netzwerken angesiedelt, um eine faire Verteilung von Bandbreite auf verschiedene Serviceklassen zu sichern. WFQ kann auch auf verschiedene Verhalten eingestellt werden. Die erste Möglichkeit ist, WFQ so einzustellen, dass es eine große Anzahl von Queues verwalten kann. Mit der zweiten Einstellungsmöglichkeit kann man eine begrenzte Anzahl von Queues, die mehrere Datenströme beinhalten, verwalten. Bei dieser Einstellung werden CoS-Regeln oder die drei Prioritätsbits von IP-Type-of-Service verwendet, um Frames den Queues zuzuordnen. Für jede der Queues wird ein bestimmter Prozentsatz der Bandbreite des Ausgangsports allokiert.

Weighted-Round-Robin (WRR) ist eine Gruppe von Scheduling-Strategien, die die Nachteile von FQ und PQ beheben sollen. Beim WRR werden die Frames zunächst klassifiziert und in verschiedene Serviceklassen aufgeteilt. Anschließend werden diese Frames in eine Queue eingefügt, die für die entsprechende Serviceklasse reserviert ist. Jede dieser Queues wird in Round-Robin-Reihenfolge abgearbeitet. Dabei werden, genau wie bei FQ und PQ, leere Queues übersprungen. WRR unterstützt die Reservierung von verschiedenen Bandbreiten für verschiedene Serviceklassen, indem es entweder mehr als ein Frame aus Queues mit höherer Bandbreitenreservierung abarbeitet oder, indem es immer nur einen Frame pro Besuch des Schedulers abarbeitet, aber dann die Queue pro Round-Robin-Durchlauf mehrfach besucht wird. Um die Menge der reservierten Netzwerkressourcen für jede Serviceklasse zu regulieren, kann eine Anzahl von Parametern verändert werden. Der Delay eines Frames in einer bestimmten Queue ist abhängig von der Rate, mit der die Frames in die Queue platziert werden, von der Größe der Queue sowie von der Anzahl der Frames, die bei jedem Round-Robin-Durchlauf abgearbeitet werden, als auch von der Anzahl der anderen Queues, die auf den gleichen Ausgangsport konfiguriert sind.

Die Hersteller von Netzwerkkomponenten sind heute bestrebt, das beste Verhältnis zwischen guter Verteilung und einfacher Umsetzung bei der Implementierung der Queuing-Strategien zu finden. Um bei extrem überlasteten Netzen auch eine Qualitätsklasse zu haben, die von den anderen unbeeinflusst geschwächt wird, bietet sich beispielsweise die Zusammenarbeit des WRR beziehungsweise WFQ mit einer Strict-PQ-Strategie an. So können die Daten zur Netzwerkkontrolle, wie sie in der CoS-Klasse 7 in dem Standard 802.1D gefordert, jederzeit ihr Ziel erreichen. Da vier verschiedene Qualitätsklassen untersucht werden sollten, bietet es sich an, drei

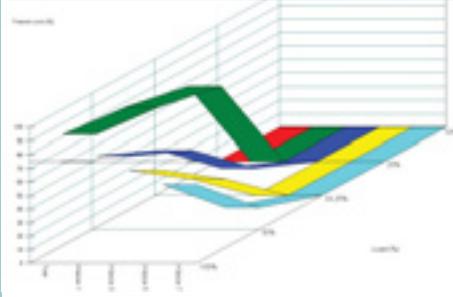
Messergebnisse

Strict-Priority-Queuing



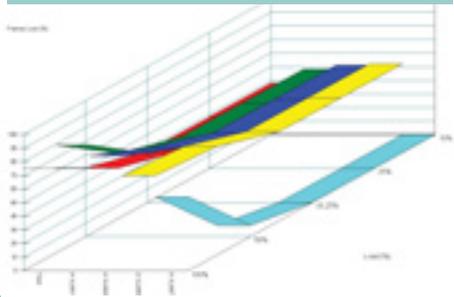
Erwarteter Frame-Loss bei korrekter Funktion unter Überlast

WFQ-ähnliche Strategien



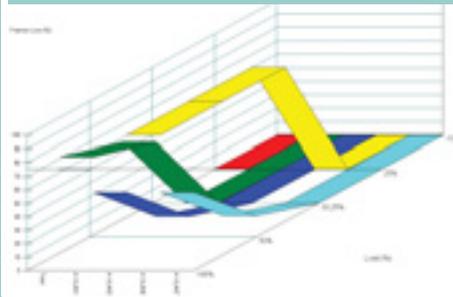
Erwarteter Frame-Loss bei korrekter Funktion unter Überlast

HP-Procurve-5304-XL



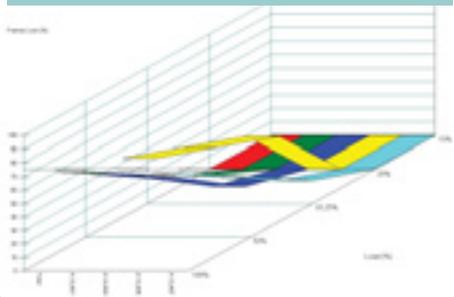
Frame-Loss Layer-2, 512-Byte-Pakete

D-Link-DES-6300



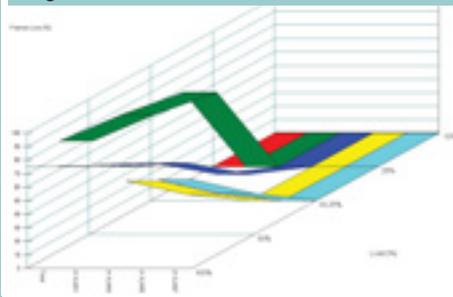
Frame-Loss Layer-2, 64-Byte-Pakete

D-Link-DES-6300



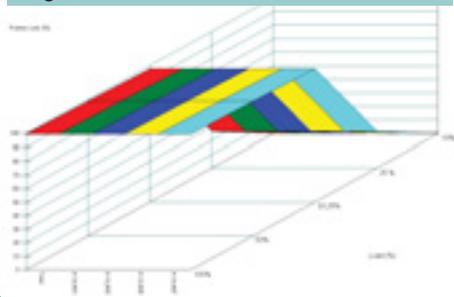
Frame-Loss Layer-2, 1518-Byte-Pakete

Netgear-GSM-7324



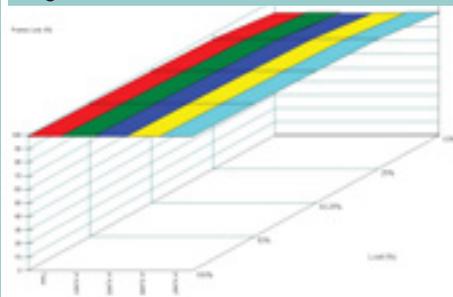
Frame-Loss Layer-2, 512-Byte-Pakete

Netgear-GSM-7324



Frame-Loss Layer-2, 1024-Byte-Pakete

Netgear-GSM-7324



Frame-Loss Layer-2, 1280-Byte-Pakete

x-Achse von links nach rechts: Mittelwert, niedrigste Priorität, zweitniedrigste Priorität, zweithöchste Priorität, höchste Priorität; y-Achse: Datenverlustrate in Prozent; z-Achse: Last an den Ausgangsports in Prozent.

Queues mit WRR / WFQ zu versehen und eine Queue mit Strict-PQ. Die mit WRR / WFQ verwalteten Queues werden für die CoS-Klassen 1, 3 und 5 verwendet. Sie entsprechen Daten für Hintergrundverkehr mit geringer Priorität, Excellent-Effort für hervorragende Übertragung und Videodaten, die in Echtzeit mit höchster Priorität übertragen werden sollten. In diesem Aufbau für die Scheduling-Strategien können dann verschiedene Gewichtungen eingestellt werden. Dabei sollte die Gewichtung des Strict-PQ nie verletzt werden. Bei Überlast in allen Prioritäten kann es in den anderen Queues theoretisch dazu kommen, dass diese ihre gewünschten Gewichtungen nicht einhalten können.

Realisierung des Testaufbaus

Um die Netzwerk-Switches zu testen, haben wir die Switches mit vier Eingangs-Ports und einem Ausgangs-Port belegt. Eine optimale Konfiguration der Switches für das gewünschte Szenario war eine Kombination von einer Queue mit Strict-Priority-Queuing für die höchste Priorität und mit drei Queues, die Deficit-Weighted-Round-Robbin oder WRR / WFQ verwendeten. Ließ sich diese Konfiguration nicht einstellen, haben wir ausgehend von den vom Werk eingestellten Konfigurationen ein möglichst ähnliches Verhalten konfiguriert. Für Layer-3 ließ sich bei den meisten Switches nur eine Bandbreitenverteilung via Diffserv einstellen. Somit war eine Abweichung vom gewünschten Testszenario auch hier unvermeidbar. Für Switches, die beim Diffserv Bandbreiteneinstellungen gestatten, haben wir eine Konfiguration gewählt, bei der pro Queue 10, 20, 30 beziehungsweise 40 Prozent der maximalen Bandbreite belegt werden durften. Je nach getestetem Switch haben wir aber auch andere Konfigurationen verwendet, die an entsprechender Stelle dargestellt sind.

Die Eingangsports und der jeweilige Ausgangsport hatten jeweils einen maximalen theoretischen Durchsatz von 1 GBit/s beziehungsweise von 100 MBit/s. Wenn alle vier Eingangsports mit 100 Prozent ihrer Kapazität überlastet, ist der Ausgangsport also vierfach überlastet. Den Netzwerkverkehr haben wir mit vier verschiedenen CoS-Typen – 1, 3, 5 und 7 – in verschiedenen Kombinationen versehen, um die Auswirkungen von CoS im Zusammenhang mit den verschiedenen Queuing-Strategien zu ermitteln. Die Tests haben wir jeweils einmal für Layer-2 via VLAN-Tagging sowie einmal für CoS auf Layer-3 via Diffserv-Signalisierung durchgeführt.

Datenverlustraten

Für die Ermittlung der Datenverlustraten haben wir vier Eingangsports auf einen Ausgangsport geschaltet. Die Burstsize haben wir auf 100 Frames pro Burst eingestellt, was aus Erfahrung ein durchschnittlicher Wert für Burst-Größen im normalen Netzwerk-Verkehr ist. Jeden Eingangsport haben wir mit vier Datenströmen belastet, wobei jeder Flow eine unterschiedliche Priorität beziehungsweise einen unterschiedlichen DSCP-Wert hatte. Sowohl für die Prioritäten als auch für die DSCP-Werte wurden jeweils die Werte 1, 3, 5 und 7 verwendet. Zum Einsatz kamen bei dem Test dann der Jumbo-Test sowie der Throughput-Test mit der Smartflow-Software. Wir haben dann die Tests für

Vergleichstest

CoS-Priorisierung bei Ethernet-Switches im Edge- und Core-Bereich

Network Computing plant einen neuen LAN-Switch-Vergleichstest. Getestet werden Ethernet-Systeme auf ihre Tauglichkeit für den Einsatz in Unternehmensnetzen mit verschiedenen Real-Time-Applikationen und entsprechenden »Quality-of-Service«-Anforderungen. Alle Hersteller/Anbieter solcher Systeme, die sich dieser Herausforderung stellen möchten, melden sich bitte per Mail unter dglogau@fox-verlag.de.

Test-Spezifikation

Testzeitraum:

12.01.2004 bis 06.02.2004

Die Testgeräte sollten für den gesamten Testzeitraum zur Verfügung stehen. Nach den notwendigen Vormessungen vereinbaren wir mit jedem Hersteller einen individuellen Messtermin in unseren Labs an der FH Stralsund. Für jeden Hersteller steht das Lab exklusiv an einem Tag zur Verfügung. An diesem Tag soll der jeweilige Hersteller die Konfiguration seines Systems vor Ort übernehmen und seine Teststellung vor Ort begleiten.

TestszENARIO

Das Modellunternehmen HighFair möchte neben den klassischen Datenapplikationen und Voice-over-IP weitere Real-Time-Applikationen in ihr Unternehmensnetz integrieren. Ein geeigneter Vergleichstest soll evaluieren, 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)
- ▶ Voice-over-IP (Call-Center)
- ▶ SAP-Anwendungsdaten
- ▶ übrige Datenanwendungen und Updates

Um die möglichst absolute Störungsfreiheit der Kommunikations- und Arbeitsprozesse im Unternehmen zu garantieren, ist eine vierstufige Daten-Priorisierung sowie eine intelligente Queuing-Policy erforderlich. Gefordert ist für die Edge-Switches neben der Datenpriorisierung ein intelligentes Bandbreitenmanagement, das es ermöglicht, von vornherein eine Überlastung des Backbones zu vermeiden.

Daraus ergeben sich folgende Anforderungen an die Teststellungen

Teststellung Core-Switch

- ▶ Layer-3-Ethernet-Switch
- ▶ mindestens 5 Gigabit-Ethernet-Ports (Multimode-LWL mit SC-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 2 CoS-Queuing-Mechanismen, wie Strict-Priority-Queuing, Weighted-Fair-Queuing oder Weighted-Round-Robin, die softwareseitig konfiguriert werden können

Teststellung Edge-Switch

- ▶ Layer-3-Ethernet-Switch
- ▶ mindestens 24 Fast-Ethernet-Ports und 2 Gigabit-Ethernet-Ports (Multimode-LWL mit SC-Stecker oder Kupfer)
- ▶ 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 2 CoS-Queuing-Mechanismen, wie Strict-Priority-Queuing, Weighted-Fair-Queuing oder Weighted-Round-Robin, die softwareseitig konfiguriert werden können
- ▶ Bandbreitenmanagement mit einstellbaren Maximalbandbreiten je Priorität

Testparameter und Verfahren

Testverfahren nach RFC 2544 (One-to-Many, Many-to-One, Many-to-Many)

Gemessen werden

- ▶ Performance
- ▶ Packet-Loss
- ▶ Latency und
- ▶ Jitter

Analysiert wird das unterschiedliche Verhalten in den verschiedenen CoS-Queuing-Modi

Testequipment

Spirent Smartbits 6000B



alle in der RFC 2544 beschriebenen Framelängen von 64 bis 1518 Byte durchgeführt. Dabei haben wir die Last stufenweise von 10 Prozent bis 100 Prozent erhöht. Die Zwischenschritte lagen bei 25, 33,25 und bei 50 Prozent.

Wenn die Ergebnisse eines Tests mit Burstsize 100 nicht nachvollziehbar waren, haben wir den Test mit einer Burstsize von 1 wiederholt. Waren die damit erzielten Ergebnisse besser, haben wir die Burstsize stufenweise erhöht, bis die Anomalien, die wir bei einer Burstsize von 100 ermittelt haben, zum ersten Mal zu sehen waren. Die Unterscheidung zwischen Layer-2- und Layer-3-Priorisierung führte zu einer Aufspaltung des Tests in zwei Testreihen, so dass der Test einmal für Prioritäten auf Layer-2 (VLAN) und einmal für Prioritäten oder auch Bandbreitenverteilungen via DSCP-Signalisierung auf Layer-3 (Diffserv) durchgeführt wurde. Für die Ergebnisse haben wir dann die für CoS wichtigen Parameter Frame-Loss, Latency, Jitter und Throuput ausgewertet.

Da im Test die Last auf dem Switch von vier auf einen Port gesendet wurde, kam es bei einer Last von 100 Prozent auf den Eingangsports zu einer vierfachen Überlastung des Ausgangsports. Dadurch ist es natürlich normal, dass die Switches im Test viele Frames nicht switchen konnten und somit viele Frames verloren. Anhand der Verteilung der einzelnen Prioritäten oder der einzelnen resultierenden Bandbreiten konnten wir dann erkennen, ob und wenn ja wo der jeweilige Testkandidat Probleme hatte.

Bei den einzelnen Laststufen wurden folgende Prämissen berücksichtigt: Bei 10 Prozent Eingangslast tritt keine Überlast am Ausgangsport auf. Somit sollten alle Datenströme durchgeleitet werden. Bei 25 Prozent Eingangslast, entsteht am Ausgangsport eine Last von 100 Prozent. Erwartet haben wir hier, dass gerade noch keine Frames verlogen gehen. Bei einer Eingangslast von 33,25 Prozent sollten die drei Klassen mit der hohen Gewichtung noch komplett geschwächt werden. Die Klasse mit der zweit schwächsten Gewichtung hat bei dieser Eingangslast aber auch schon ihre Grenzen erreicht. Bei einer Eingangslast von 50 Prozent sollte der Switch schon bei den beiden Klassen mit der geringeren Gewichtung Frames verlieren. Die zweitstärkste Klasse sollte aber bei dieser Lastverteilung auch im Grenzbereich liegen. Bei 100 Prozent Eingangslast sollte nur noch die Klasse mit der stärksten Gewichtung verlustfrei geschwächt werden

Ein solches Verhalten ist aber nur zu erwarten, wenn die verwendeten Scheduling-Strategien nach Strict-Priority-Queuing arbeiteten. Arbeiteten die Switches mit anderen Verfahren, so war zu erwarten, dass die geringeren Prioritäten nicht zu 100 Prozent verworfen werden, bevor die nächst stärker gewichtete Queue Frames verliert. In diesem Fall sind Ergebnisse zu erwarten, die bei Volllast auch zu Verlusten in der höchsten Priorität führen.

Das galt besonders für die Ergebnisse in dem Teil, in dem wir die Switches auf Layer-3 mit einer Bandbreitenverteilung getestet wurden. Hier waren auch für Datenströme mit geringer Gewichtung minimale Bandbreiten zugewiesen. Dadurch verloren sie selbst bei maximaler Eingangslast nicht alle Frames und der Datenstrom mit der höchsten Priorisierung musste auch Datenframes verlieren, damit

die eingestellten Bandbreiten resultieren konnten. Zu beachten ist hier, dass wir bei diesem Vergleichstest nur das Verhalten der einzelnen Queuing-Strategien untersucht haben. In unserem Szenario sollte nach Möglichkeit der jeweils zu testende Switch so konfiguriert sein, dass drei Queues mit WRR oder WFQ mit einer Queue nach Strict-PQ zusammen arbeiten. Dabei sollte die Strict-Priority-Queue die höchste Priorität haben.

Layer-2-Priorisierung im Test

Aus der Reihe der Testkandidaten hatten alle Switches, mit Ausnahme des Netgear-Switches und des T5C von BATM nach dem Einspielen einer neueren Firmware-Version (4.6.0.6.4.1) mit verbessertem Buffermanagement, Probleme mit einer Burstsize von 100 Frames pro Burst. Dabei zeigte sich, dass die Switches auf einigen Datenströmen schon bei einer geringen Eingangslast von 10 Prozent eine nicht unerhebliche Anzahl von Frames verwarfen. Ein typischer Vertreter dabei war zum Beispiel der Procurve-Switch von HP bei der Messung mit 512-Byte-Paketen.

Die Messergebnisse brachten hier bei allen gemessenen Frame-Längen in etwa gleiche Resultate. Diese Ergebnisse ließen nur einen Schluss zu. Die Buffergrößen bei diesen Geräten ist zu klein. Denn wenn die Frames nacheinander in die Queue kommen, konnten sie auf Grund des großen Ansturms bei einem Burst nicht sofort abgearbeitet werden. Waren die Buffer für die Queues dann zu klein, wurden die Frames am Ende eines Bursts verworfen, obwohl der Ausgangsport noch nicht seine maximale mögliche Kapazität ausgeschöpft hatte. Dieses Problem hat aber nicht unmittelbar etwas mit dem Verwenden von CoS zu tun. Deshalb haben wir bei diesen Testteilnehmern die Burstsize auch noch einmal auf eine Größe von einem Frame pro Burst verringert und dann manuell ermittelt, bei welcher Burstsize die einzelnen Kandidaten die ersten Frames schon bei einer geringen Last von zehn Prozent verlieren.

Der DES-6300 von D-Link erzeugte bei diesem Test sehr merkwürdige Ergebnisse. Prioritäten mit dem Wert 5 behandelte er durchgängig am schlechtesten. Bei steigender Framelänge änderte sich das Ergebnis zwar, aber nur die Auswirkungen waren nicht mehr ganz so stark zu sehen. Bei langen Frames neigte der DES-6300 dazu, alle Datenströme gleichmäßig zu verwerfen. Aber selbst hierbei war die ungewöhnliche Verteilung der Prioritäten noch erkennbar.

Der Grund für dieses Verhalten war leider nicht nachvollziehbar, da der D-Link-Switch keine Konfigurationsmöglichkeiten für CoS auf Layer-2 bereit hielt. Dadurch wären alle Aussagen zu diesem Verhalten reine Spekulation. Interessant war bei diesem Testkandidaten auch, dass bereits bei einer Burstsize von 4 Frames pro Burst die ersten Frames bei zehn Prozent Eingangslast verloren gingen.

Der Extreme-Summit-48-SI arbeitete bei einer Burstsize von einem Frame pro Burst exakt so, wie es nach Standard zu erwarten ist. Diese Aussage lässt sich für alle untersuchten Framelängen verallgemeinern. Da wir für den Test den Switch so konfiguriert haben, dass er für alle Prioritäten Strict-Priority-Queuing verwendet, war das somit ein gutes Ergebnis. Die Messergebnisse blieben bis zu

einer Burstsize von 64 Frames pro Burst gut, dann fing auch dieser Switch an, schon ab 10 Prozent Eingangslast Datenframes zu verlieren.

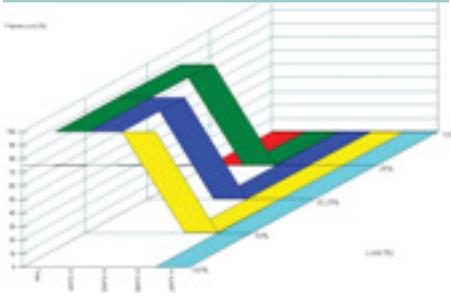
Auch HPs Procurve-5304-XL arbeitete bei kleineren Bursts wie erwartet. Bei diesem Gerät fand jedoch eine andere Queuing-Strategie (WFQ) ihren Einsatz. Auch hier gilt die Aussage gleichbedeutend für alle Framelängen. Die Schwierigkeiten des HP-Procurve begannen dann mit einer Burstsize von 49 Frames pro Burst.

Der Netgear-GSM-7324 arbeitete bei einer Burstsize von 100 Frames pro Burst zwar genauso wie bei einer Burstsize von 1, er hatte aber ein ganz

anderes Problem. Für kleine Frames ergaben die Messwerte ein gleiches Bild wie in der Abbildung, die den erwarteten Fall für WFQ-ähnliche Scheduling-Strategien darstellt. Bei einer Framesize von 512 Byte unterschied der Switch jedoch nicht mehr nachweisbar zwischen den Prioritäten fünf und sieben. Für Frames mit einer Framelänge von 1024 Byte gab es gar keine messbare Unterteilung zwischen den verschiedenen Prioritäten mehr. Außerdem wurden ab einer Eingangslast von 33,25 Prozent fast alle Datenframes verloren. Ab einer Framelänge von 1280 Byte verlor der Netgear-Switch schon ab einer Eingangslast von 10 Prozent

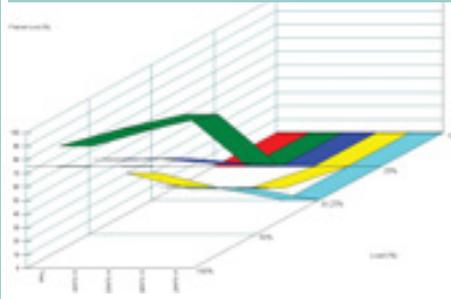
Messergebnisse

BATM-T5C



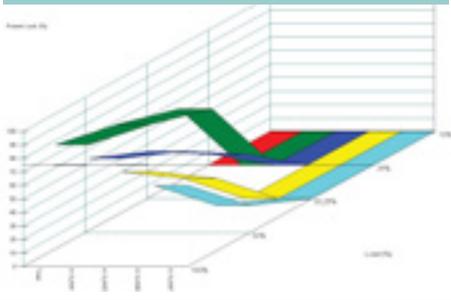
Frame-Loss Layer-2, 64-Byte-Pakete (Firmware 4.6.0.6.4.1)

BATM-T5C



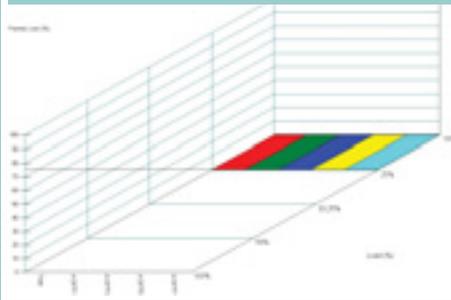
Frame-Loss Layer-2, 64-Byte-Pakete, WRR-Betrieb

BATM-T5C



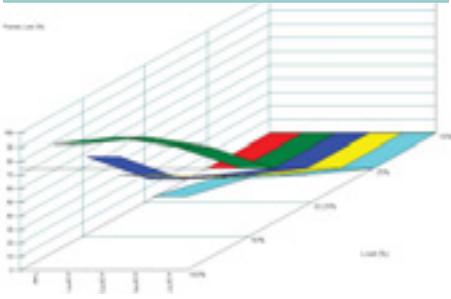
Frame-Loss Layer-2, 1518-Byte-Pakete, WRR-Betrieb

D-Link-DES-6300



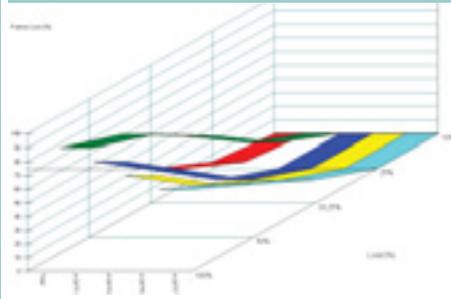
Frame-Loss Layer-3, 64-Byte-Pakete

Netgear-GSM-7324



Frame-Loss Layer-3, 64-Byte-Pakete

Netgear-GSM-7324



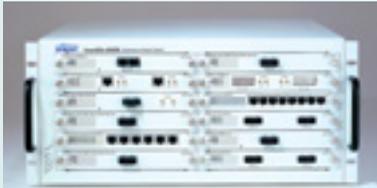
Frame-Loss Layer-3, 1518-Byte-Pakete

x-Achse von links nach rechts: Mittelwert, niedrigste Priorität, zweithöchste Priorität, zweithöchste Priorität, höchste Priorität; y-Achse: Datenverlustrate in Prozent; z-Achse: Last an den Ausgangsports in Prozent.

Info

So testete Network Computing

Als Lastgenerator und Analysator haben wir in den Real-World Labs der Network Computing einen »Smartbits 6000B Traffic Generator/Analyser« von Spirent eingesetzt. Das rund 250 000 Euro teure System ist mit der Software »SmartFlow« ausgestattet und mit 24 Fast-Ethernet-Ports, fünf Gigabit-Ethernet-LWL-Ports sowie vier Gigabit-Ethernet-Kupfer-Ports bestückt. Alle Ports können softwareseitig als Lastgeneratortausgang und/oder als Analysatoreingang konfiguriert werden. Die Class-of-Service (»Quality of Service«) der Switches im Testfeld haben wir in verschiedenen Mess- und Testreihen nach dem Testverfahren gemäß RFC 2544 gemessen. In diesen Tests haben wir zunächst Rahmen der Ebene 2 untersucht. Die vorliegenden Class-of-Service-Tests analysieren im ersten Schritt die Prioritätsmechanismen entsprechend IEEE 802.1p/Q. Im zweiten Schritt haben wir dann die Layer-3-Priorisierung nach Diffserv untersucht. Alle Messun-



gen haben wir mit Smartbits durchgeführt und analysiert. In unseren Testszenarios haben wir verschieden priorisierte Datenströme von einem oder mehreren Eingangsports auf einen oder mehrere Ausgangsports gesendet. Hierbei sind in unserem Testszenario Fast-Ethernet- sowie auch Gigabit-Ethernet-Interfaces zum Einsatz gekommen. Die die Priorisierung bestimmenden Bits wurden im Header der Datenrahmen mit drei Bits nach IEEE 802.1p auf Layer 2 sowie nach ToS beziehungsweise Diffserv auf Layer 3 festgelegt.

fast alle Frames, und das zudem auch noch recht gleichmäßig in allen Prioritäten.

Der insgesamt verfügbare Speicher dieses 24-Port-Switches schien durchaus in der Lage zu sein, viele Pakete auf einmal aufzunehmen. Mit 256 kByte Speicher pro Queue bestätigt sich die Vermutung, dass der Switch genügend Speicher besitzt. Immerhin passen bei dieser Größe 172 Frames mit einer Länge von 1518 Byte in jede Queue. Der Grund für die Probleme mit langen Frames ist also nicht in den Speichern zu suchen. Unsere Vermutung ist, dass der Switch eine fehlerhafte Umsetzung in der Queuing-Strategie hatte.

Der BATM-Switch hatte mit der zuerst verwendeten Firmware (4.6.0.5.0.0) noch Probleme bei Bursts von mehr als 16 Frames hintereinander. Stellt man bei dieser Firmware dann eine geringere Burstsize ein, arbeitet der Switch aber nahezu ideal. Hier gab es dann nur kleine Abweichungen, wobei bereits minimale Verluste bei geringerer Last in

den höheren Prioritäten auftraten. Dieses Verhalten galt für alle Framegrößen. Nach dem Einspielen einer neuen Firmware-Version (4.6.0.6.4.1) war das Verhalten für alle Framegrößen gut. Auch bei der eingestellten Burstsize von 100 Frames gab es dabei keine weiteren Probleme.

Der BATM-Switch war nicht nur für Strict-Priority-Queuing konfigurierbar, sondern auch noch für ein erweitertes WRR. Diese Konfiguration konnten wir nur mit der älteren Firmware-Version testen. Dadurch waren aussagekräftige Ergebnisse nur unter Verwendung von einer Burstsize mit nur einem Frame zu erreichen. Der T5C wurde dazu so eingestellt, dass die Frames mit einer VLAN-Priorität von 1 gerade 10 Prozent, mit 3 20 Prozent, mit 5 30 Prozent und mit 7 40 Prozent der Bandbreite bekamen. Dieses Verhältnis stellte sich bei den einzelnen Tests dann auch immer ein. Nur bei 64 und bei 128 Byte großen Datenframes kam es bei einer Eingangslast von 50 Prozent zu einer Gleichverteilung für die Prioritäten 5 und 7. Ab einer Framelänge von 512 Byte war die Arbeitsweise dieses Switches dann absolut in Ordnung.

Layer-3-Priorisierung im Test

Beim Frame-Loss-Test mit Layer-3-Priorisierung verhielten sich die Test-Geräte im Grunde genauso wie bei den Layer-2-Messungen. Auch hier war der Switch von Netgear der einzige Vertreter, der keine Probleme mit einer Burstsize von 100 Frames erkennen ließ. Der Summit von Extreme Networks zeigte in der Disziplin der Layer-3-Priorisierung exakt das selbe Verhalten wie auf Ebene 2. Insgesamt konnten die Aussagen über die Burstsizes, bei denen die Vertreter mit Problemen schon bei geringer Last anfangen Datenframes zu verlieren, problemlos übernommen werden.

Interessant ist das Verhalten des DES-6300 von D-Link bei einer Burstsize von einem Frame pro Burst. Trotz eingeschalteter CoS für Layer-3 mit verschiedenen Bandbreitenreservierungen für verschiedene Diffserv-Werte konnten wir keinerlei Unterscheidung bei der Behandlung der verschiedenen Diffserv-Klassen messen. Die ermittelten Messergebnisse sehen für alle gemessenen Framelängen aus, wie bei der Messung mit 64-Byte-Paketen. Bisher konnte D-Link Deutschland keine Erklärung für die gemessenen Werte geben.

Der Switch von HP arbeitete bei kleinen Bursts auch auf Layer-3 so gut, wie auf Layer-2. Hier konnten wir wieder Ergebnisse ermitteln, die die Erwartungen an einen Switch mit WFQ-ähnlichen Strategien darstellt.

Erstaunlich war das Verhalten des Vertreters von Netgear. Bei kurzen Frames unterschied er die drei größeren Diffserv-Klassen erst bei recht hohen Lasten. Bei wachsender Framelänge für die Datenframes kam hinzu, dass der Verlust in der Diffserv-Klasse 1 bei 25 Prozent Eingangslast schon relativ hoch war. Bei 100 Prozent Eingangslast haben sich aber dann bei diesem Kandidaten für alle Framelängen die im Bandbreiten-Management eingestellten Werte – 10, 20, 30, sowie 40 Prozent der gesamten Bandbreite – eingependelt.

Fazit

Mit Ausnahme der Switches von BATM und Netgear hatten alle Systeme im Test mehr oder weniger

deutliche Probleme mit einer Burstsize von 100 Frames. Dabei kam es teils schon bei einer Eingangslast von 10 Prozent zu erheblichen Frame-Verlusten. Das Verhalten des D-Link-DES-6300 ließ sich weder auf Layer-2 noch auf Layer-3 mit einer der geforderten Priorisierungstechniken korrelieren. Netgears GSM-7324 arbeitete dagegen nachweisbar mit einem WFQ-Verfahren. Allerdings auch nicht in allen Fällen fehlerfrei, so konnte er beispielsweise bei einer Framegröße von 512 Byte und Layer-2-Priorisierung nicht mehr zwischen den beiden höheren Prioritäten unterscheiden. Eine WFQ-Priorisierung war auch dem HP-Switch nachzuweisen. Aber der HP-Switch Procurve-5304-XL arbeitete nur bei vergleichsweise kleinen Burstsizes fehlerfrei. BATMs T5C-L3 hatte Probleme mit kleinen Datenrahmen und Extreme Networks Summit-48-SI zeigte sich nur bei Burstsizes von maximal 64 Frames von seiner besten Seite. Insgesamt ist keiner der Switches im Testfeld ohne nennenswerte Probleme geblieben – ein eher schwaches Bild, wenn man realisiert, dass die CoS-Priorisierung schon seit einigen Jahren in LAN-Switches implementiert wird und mit der fortschreitenden Verbreitung konvergenter Anwendungen in der Praxis immer wichtiger wird.

Unser Pilottest in unseren Labs an der FH Stralsund hat gezeigt, dass die geforderten CoS-Queuing-Mechanismen keinesfalls immer so funktionieren, wie von der Papierform her zu erwarten gewesen wäre. Probleme zeigten sich insbesondere bei größeren Bursts und kleinen Frame-Formaten. Steht die Investition in eine CoS-Queuing-fähige Infrastruktur an, sind IT-Verantwortliche gut beraten, wenn sie ihre Anforderungen an die Priorisierungsmechanismen genau spezifizieren und auf gründliche Tests setzen, bevor eine Investitionsentscheidung getroffen wird. Und um CoS als Ende-zu-Ende-Verbindung zu nutzen, ist es natürlich notwendig, dass alle auf diesem Weg befindlichen Netzwerkknoten dieses Feature besitzen. Befindet sich auf diesem Weg eine aktive Netzwerk-Komponente, die keine Unterstützung bietet, kann das so genannte »Service-Level-Agreement« nicht gegeben werden. Der SLA ist aber wichtig, um die verschiedenen Parameter des CoS für einen Datendienst, für den Provider und den Endteilnehmer festzulegen. Es gilt somit als eine Art von Vertrag.

Dipl.-Ing. Toralf Runge, [dg]

Anzeige

NETWORK COMPUTING TECHNOLOGY SECURITY-DAYS TOUR 2003

27. November 2003

SECURITY-DAY STUTTGART

Weitere Informationen und Anmeldung:
www.networkcomputing.de/technology-tour/