

MPLS im Test

Testszenarien und Standards für Multiprotocol Label Switching

Kai-Oliver Detken

Die Aufgabe, die Anzahl der nötigen Router-Hops auf einer Übertragungsstrecke so gering wie möglich zu halten, ist heute noch nicht optimal gelöst. So wird derzeit beispielsweise in IP-Netzen jedes einzelne Paket von einem Router zum nächsten weitergereicht. Die Router müssen dazu die Header der Datenpakete analysieren und anhand der darin abgelegten Informationen eine Wegewahl treffen (Hop-by-hop-Routing). Das Verfahren ist relativ aufwendig und erhöht die Verzögerungszeiten.

Deshalb wurde das Label-Switching-Verfahren entwickelt, das als Grundlage für einen gemeinsamen – Multiprotocol Label Switching (MPLS) genannten – Standard gilt.

Die MPLS Working Group treibt die Standardisierung stark voran. Wesentliche Punkte des Protokolls MPLS wie seine Architektur oder das Label Distribution Protocol (LDP) wurden bereits festgelegt. Allerdings fehlt es u.a. noch an einer gemeinsamen Signalisierung, da hier unterschiedliche Verfahren im Einsatz sind. Trotz fehlender Standards ist MPLS allerdings bereits von einigen Router-Herstellern implementiert worden und wird in einigen Netzen bereits eingesetzt. Die Vorteile von MPLS lassen sich so zusammenfassen:

- *Virtual Private Network (VPN)*: Durch das Label-Konzept ist MPLS sehr gut dazu geeignet, VPNs zu realisieren.
- *Multi-Protocol- und Multi-Link-Unterstützung*: Obwohl hauptsächlich für die IP-Versionen 4 und 6 genutzt, bietet MPLS auch die Möglichkeit, andere Protokolle wie IPX zu übertragen. Weiterhin ermöglicht es die Kommunikation und Ressourcenreservierung über verschiedene Techniken, zum Beispiel Ethernet, PoS, ATM und WDM.
- *Quality of Service (QoS)*: Mit MPLS ist es möglich, Ressourcenreservierung im Netz vorzunehmen. Hierfür gibt es speziell angepaßte Verbindungsaufbauprotokolle, die diese Reservierungen durchführen.
- *Verschiedene Service-Klassen*: MPLS ermöglicht es dem Service Provider, seinen Kunden verschiedene Dienstklassen anzubieten. Der Kunde profitiert von dieser Wahlmöglichkeit. Er wählt den Dienst nach seiner Qualität und den damit verbundenen Kosten.
- *Einsatz mit ATM*: Kombiniert mit ATM, bietet MPLS auch „harte“ Dienstgüte wie Bandbreitengarantien, Zusicherung von Verzögerungszeiten und Laufzeiten der Pakete. Vorhandene ATM-Switche

können in ein MPLS-Netz integriert und weitergenutzt werden, auch wenn sie selbst kein MPLS unterstützen.

MPLS funktioniert nach dem Prinzip des Label Swapping und verhält sich demnach sehr ähnlich wie Tag-Switching und ARIS (Aggregate Route-based IP-Switching). Allerdings spezifiziert MPLS deutlich mehr, als in beiden Verfahren beschrieben ist. Das heißt, MPLS skaliert kleine Netze und eine hohe Vergabezahl der Label bis hin zu großen Netzen, in denen die Label für große Routing-Gebiete vergeben werden. MPLS macht sich den Umstand zunutze, daß Layer-3-Protokolle Informationen besitzen, die nicht für die Bestimmung des nächsten Routers bzw. Hops nötig sind. Für die einfache Weiterleitung von Paketen reichen wesentlich weniger Informationen aus.

Die Vergabe eines Labels wird bei MPLS immer durch eine Netzkomponente durchgeführt. Nachdem dies geschehen ist, wird der ihr vorgelagerte Nachbar über diese Entscheidung informiert. Gleichzeitig wird mitgeteilt, welches Label für den Datenstrom vergeben wurde. Dabei besitzt MPLS eine hohe Skalierbarkeit,

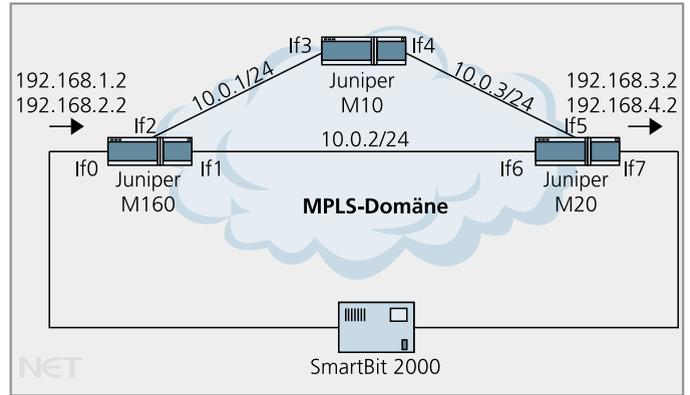
Das Thema in Kürze

Aufgrund seiner unbestreitbaren Vorteile gewinnt das Multiprotocol Label Switching weiter an Bedeutung. Der Autor schildert in dem Beitrag Tests, die auch eine Beurteilung der Praxistauglichkeit zulassen – etwa Messungen des Traffic Engineering oder Performance-Messungen. Hierbei ließen sich z.B. die Funktionalität des Constraint Based Routing für eine QoS-Unterstützung nachweisen oder die Eignung für Echtzeitanwendungen.

Kai-Oliver Detken ist Senior IT Consultant der Detken Consultancy & Internet Technologies sowie Dozent und freier Autor in Grasberg

da Label für ganze Routing-Gebiete sowie einzelne Datenströme vergeben werden können. Dabei wird den topologieorientierten Verfahren der Vorzug gegeben. Switchingpfade durch das Netz werden von einem Protokoll festgelegt, das die Verteilung der Label im Gesamtnetz steuert. Zusätzlich können aber auch einzelne Switching-Pfade zusammengefaßt werden. Dadurch entsteht ein vom Egress-LSR aus gesehener Switching-Baum.

Bild 1: MPLS-Testszenario



Messungen des Traffic Engineering (TE)

Die für die Darstellung des Constraint Based Routing (CBR) verwendete Testkonfiguration basiert auf den drei Juniper-Routern M160, M20 und M10. Als weiteres Element wird das Meßgerät SmartBit 2000 von Netcom Systems verwendet.

Es wird in diesem Test der Transport von zwei FECs über zwei CR-LSPs betrachtet. Die Aufgabe des SmartBit in diesem Test besteht in der Erzeugung von zwei kontinuierlichen Datenflüssen, welche die beiden FECs darstellen werden. Dazu werden über den SmartBit vier Hosts simuliert. Die Hosts 192.168.1.2 und 192.168.2.2 bilden die Quellen und die Hosts 192.168.3.2 und 192.168.4.2 die Senken. Das SmartBit wird über eine ATM-Schnittstelle an den M160 angeschlossen. An dieser Stelle erfolgt die Simulation der sendenden Hosts.

Über das zweite ATM-Interface wird das SmartBit mit dem M20 verbunden. Hierbei simuliert das SmartBit die empfangenden Hosts. Die drei Juniper-Router werden untereinander über Packet-over-SONET-Interfaces (PoS) verbunden (Bild 1).

Die Aufgabe der Juniper-Router besteht in dieser Testanordnung im Transport der vom SmartBit generierten Datenflüsse vom M160 zum M20. Dieser Transport soll über zwei Label Switching Paths (LSPs) erfolgen. Als Label-Distribution-Protokoll muß für CBR in der Juniper-Implementierung RSVP (Resource Reservation Protocol) verwendet werden.

Die Aufgabe des Routing-Protokolls übernimmt das Intermediate System to Intermediate System (IS-IS). Die beiden LSPs werden mit den Namen LSP_Direkt und LSP_Indirekt bezeichnet. Jedem LSP wurde der Transport eines Datenflusses fest zugewiesen. Auf diese Weise kann die Funktion ei-

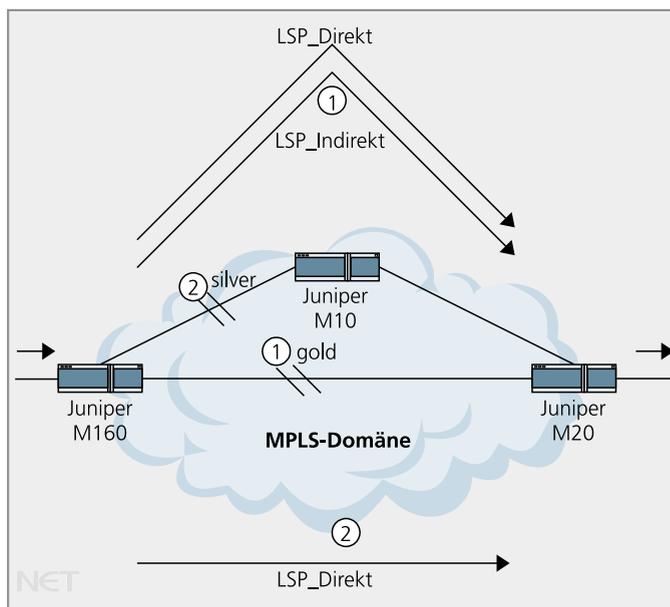
nes LSP über das Empfangen des entsprechenden Datenflows verifiziert werden.

In diesem Testszenario sollte das Traffic Engineering als Anwendung des Constraint Based Routing gezeigt werden. Das Traffic Engineering erlaubt das kontrollierte Einrichten einzelner LSPs mit dem Ziel, das Netz gleichmäßiger zu belasten. Zu diesem Zweck ist es im MPLS möglich, einzelnen Verbindungen unterschiedliche Übertragungsklassen mittels administrativer Attribute zuzuweisen. Über die Konfiguration eines LSP kann für dessen Aufbau die Verwendung bestimmter Links vorgeschrieben oder ausgeschlossen werden.

Die bisher beschriebene Testanordnung bildet die Basis für das Testszenario. Für die Anwendung des TE wird jedoch zusätzlich der Verbindung M160 – M20 die Übertragungsklasse gold und der Verbindung M160 – M10 – M20 die Übertragungsklasse silver zugewiesen. Dies wird über die Konfiguration der Schnittstellen des M160 realisiert. Auf diese Weise darf der Aufbau dieses LSP nicht über Verbindungen der Übertragungsklasse gold erfolgen. Dieser LSP realisiert das TE. Die Konfiguration des LSP_Direkt hingegen enthält außer der Zieladresse keine weiteren Argumente. Das Routing für diesen LSP basiert auf konventionellem Routing (Bild 2).

Die Bestimmung des Pfades, über den ein LSP aufgebaut wird, erfolgt durch eine Berechnung des Algorithmus Constraint Shortest Path First (CSPF). Der Algorithmus basiert auf dem Algorithmus Shortest Path First, der beispielsweise von Open Shortest Path First (OSPF) verwendet wird. Das CSPF ist aber zusätzlich in der Lage, weitere

Bild 2: LSP-Rerouting



Parameter für die Routing-Entscheidung zu berücksichtigen. Für den Aufbau des LSP_Direkt bei intakten Verbindungen sollte der CSPF-Algorithmus die Verbindung mit der geringsten Anzahl von Hops gewählt (M160 – M20) wählen. Für den LSP_Indirekt dürfte diese Verbindung nicht verwendet werden, da seine Konfiguration die Nutzung von Links der Übertragungsklasse gold ausschließt. Das sollte in den Berechnungen des CSPF-Algorithmus berücksichtigt werden. Der Aufbau des LSP_Indirekt sollte nur über den Pfad M160 – M10 – M20 möglich sein.

Die Trennung der Verbindung M160 – M20 ist automatisch mit dem Abbau des LSP_Direkt verbunden. Über konventionelles Routing sollte für diesen LSP der Alternativpfad über die Verbindung M160 – M10 – M20 gewählt werden. Zunächst wurde die Verbindung M160 – M20 wiederhergestellt. Die Trennung der Verbindung M160 – M10 ist automatisch mit dem Abbau der LSPs LSP_Direkt und LSP_Indirekt verbunden. Für den LSP_Direkt sollte über konventionelles Routing der Alternativpfad über die Verbindung M160 – M20 gewählt werden. Für den LSP_Indirekt sollte es über den CSPF-Algorithmus nicht möglich sein, einen Alternativpfad zu berechnen, da die Verwendung des Links M160 – M20 über die Konfiguration ausgeschlossen wurde. Der LSP_Indirekt sollte sich nicht mehr aufbauen können, was die Messungen auch bestätigten.

Das Basisszenario wird für die Tests weiterhin beibehalten. Für die QoS-Unterstützung wurde den beiden LSPs aber zusätzlich eine Bandbreitenanforderung von jeweils 100 Mbit/s zugewiesen. Um die Auswirkungen der Preemption auf das Rerouting darzustellen, wurden den LSPs unterschiedliche Prioritäten zugeordnet. Dem LSP_Direkt wurde eine Setup- und Holding-Priorität von jeweils 0 und dem LSP_Indirekt eine Setup- und Holding-Priorität von jeweils 7 zugewiesen. Auf diese Weise wurde der Aufbau des LSP_Direkt bevorzugt. Wegen der höheren Setup-Priorität sollte der CSPF-Algorithmus mit der Berechnung des Pfades für den

LSP_Direkt beginnen. Da auf allen Links eine Bandbreite von 155 Mbit/s zur Verfügung stand, wurde für diesen LSP die Verbindung mit der geringsten Anzahl von Hops gewählt (M160 – M20). Auf dem Link wurden 100 Mbit/s für den LSP_Direkt reserviert, wodurch sich die verfügbare Bandbreite auf 55 Mbit/s reduzierte. Für die anschließende Berechnung des Pfades für den LSP_Indirekt, der ebenfalls über eine Bandbreitenanforderung von 100 Mbit/s verfügte, bestand in der Verbindung M160 – M20 keine Option mehr. Deshalb mußte der CSPF-Algorithmus den alternativen Pfad M160 – M10 – M20, der auf allen Links eine verfügbare Bandbreite von 155 Mbit/s aufwies, für den Aufbau des LSP wählen. Auch an dieser Stelle sollte die Reservierung der geforderten 100 Mbit/s erfolgen.

Die Trennung der Verbindung M160 – M10 war automatisch mit dem Abbau des LSP_Indirekt verbunden. Für das Rerouting des LSP_Indirekt mußte der CSPF-Algorithmus einen alternativen Pfad zum M20 berechnen, der eine verfügbare Bandbreite von 100 Mbit/s aufwies. Die einzig mögliche Verbindung bestand dabei im Link M160 – M20. Wegen der Reservierungen des LSP_Direkt standen hier aber nur 55 Mbit/s zur Verfügung. Deshalb sowie wegen der geringen Setup-Priorität des LSP_Indirekt durfte hier ein Rerou-

dem Abbau des LSP_Direkt verbunden ist. Für das Rerouting des LSP_Direkt mußte der CSPF-Algorithmus einen alternativen Pfad zum M20 berechnen, der eine verfügbare Bandbreite von 100 Mbit/s aufwies. Die einzig mögliche Verbindung bestand im Link M160 – M10 – M20. Wegen der Reservierungen des LSP_Indirekt standen hier aber nur 55 Mbit/s zur Verfügung. Da die Setup-Priorität des LSP_Direkt jedoch höher als die Holding-Priorität des LSP_Indirekt ausgelegt war, mußten der LSP_Indirekt abgebaut und seine Ressourcen für den Aufbau des LSP_Direkt zur Verfügung gestellt werden. Das darauffolgende Rerouting des LSP_Indirekt war nun aufgrund der verfügbaren Ressourcen und der Preemption-Werte nicht mehr möglich, was die Messungen ebenfalls bestätigten.

Performance-Messungen

Ein weiterer Test wurde von der Network Test, Inc., mit den Core-Routern Cisco 12416 (12.0(14)SX) und Juniper Networks M160 (4.2R2.4/4.2E) durchgeführt. Bei einigen Tests waren auch Charlotte's Networks und Foundry dabei. Hierbei wurden ebenfalls ein SmartBit-6000-Analysator mit 3505-Terametrics-Karten als Meßgerät verwendet sowie als PoS-Schnittstellen OC-48c (2,5 Gbit/s) und OC-192c (10

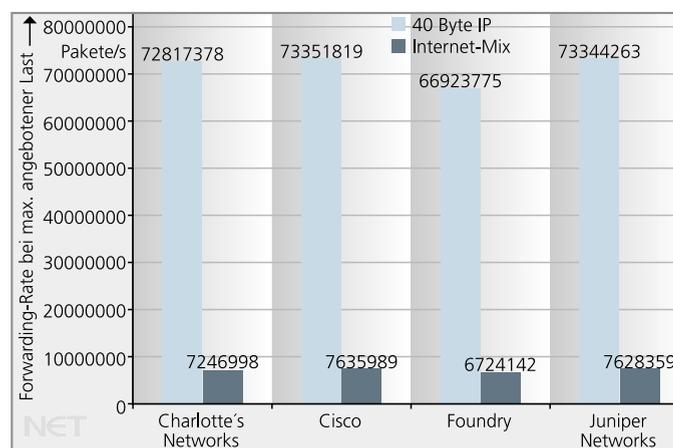


Bild 3: OC-48c-Forwarding-Rate bei IP-Routing: Charlotte's Networks wurde mit 95 % und Juniper Networks mit 99,9 % angebotener Last für den Internet-Mix getestet

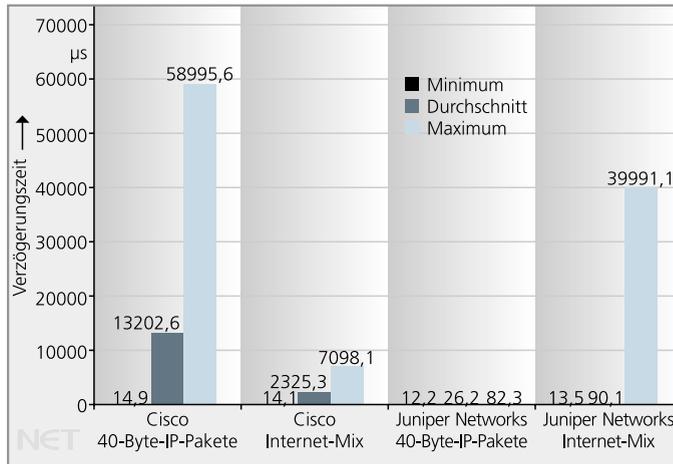
ting nicht möglich sein. Was auch tatsächlich der Fall war.

Anschließend wurde die Verbindung M160 – M10 wiederhergestellt. Als Ausgangssituation wurde dann die Trennung der Verbindung M160 – M20 initiiert, die automatisch mit

Gbit/s). Die Testplattformen bestanden aus drei Core- und drei Edge-Interfaces (OC-48c) sowie aus drei Core- und 12 Edge-Interfaces (OC-192c), so daß die Randschnittstellen beider Testplattformen die Kernkapazitäten ohne Überlast bewältigen konnten.

Um eine möglichst realistische Abbildung des Internetverkehrs zu bekommen, wurde ein Verkehrsmix von 40 Byte (55 %), 1.500 Byte (23 %), 576 Byte (17 %) und 52 Byte (5 %) erzeugt, der dem realen Verkehrsverhalten möglichst nah kommen sollte.

Bild 4: OC-192c-Verzögerung über MPLS (je niedriger die Werte, desto besser)



Zuerst wurden die maximale Forwarding-Datenrate sowie der zu erreichende Datendurchsatz der beiden eingesetzten Schnittstellen gemessen. Dabei konnten die Geräte von Cisco Systems und Juniper Networks den Datenverkehr mit Wirespeed bei der OC-48c-Schnittstelle handhaben, während die beiden anderen Hersteller Ausfälle aufwiesen. Dabei wurden 40-Byte-Pakete und der erwähnte Internet-Mix eingesetzt. Die OC-192c-Schnittstelle wies dann aber auch hier die ersten Einbrüche auf, obwohl Cisco den Internet-Mix ohne Verzögerung handhaben konnte und Juniper nur 10 % Einbrüche hatte. Allerdings machte Cisco die 40-Byte-Datenrate zu schaffen, so daß hier nur 52 % der Gesamt-Performance erreicht werden konnten, während Juniper immerhin auf 92,2 % kam. Beim Forwarding sahen die Ergebnisse der beiden unterschiedlichen Verkehrsströme schon anders aus. Hier konnte bei allen Geräten eine gute Performance der 40-Byte-Datenströme erreicht werden, während der Internet-Mix-Verkehr stark absackte, da dieser schwerer zu handhaben ist, wie Bild 3 verdeutlicht.

Im nächsten Schritt wurde MPLS auf den möglichen Datendurchsatz überprüft. Während vorher IP-Routing zum Einsatz kam, konnten jetzt direk-

te Pfade etabliert werden. Die Datenrate lag ungefähr bei 100 % bei allen Testkomponenten. Einzige Ausnahme war Juniper bei dem Internet-Mix-Verkehr, wo nur 88 % erreicht werden konnten. Anschließend wurden die Verzögerungen auf der OC-48c-

Schnittstelle gemessen. Hier verhielt sich Juniper ebenfalls konstant und konnte die gleichen Ergebnisse erzielen wie beim IP-Routing. Cisco hatte allerdings Probleme, da die Verzögerung von 2 ms auf 16 ms bei 40-Byte-Paketen anstieg, wenn die Pakete über MPLS transportiert wurden. Der Internet-Mix stieg im Vergleich dazu unwesentlich von 250 auf 650 µs an. Junipers Werte lagen dagegen im Bereich von 170 µs.

Bei der OC-192c-Schnittstelle verhält sich die Meßsituation anders. Die maximale Verzögerung von Juniper beläuft sich auf ca. 11 ms bei 40-Byte-Paketen. Dies entsprach einer ungefähren Verdopplung der Verzögerung im Vergleich zu Cisco. Man kann die Verzögerung zwar weiter senken. Dies würde allerdings bedeuten, daß man Paketverluste in Kauf nehmen müßte, was bei dieser Messung nicht getan wurde. Beim Einsatz von MPLS verhielt sich Juniper identisch zu den vorher ermittelten Werten. Ciscos Verzögerung stieg dagegen bei 40-Byte-Paketen von 26,4 µs auf bis zu 13 ms an. Beim Internet-Mix stieg der Wert ebenfalls von 500 µs auf 2,5 ms an. Die Tests machten hier deutlich, wie wichtig der korrekte Einsatz der Pufferspeicher ist.

Weiterhin spielt die Kapazität von MPLS bezüglich der zur Verfügung

stehenden Tunnel eine entscheidende Rolle. Da MPLS verbindungsorientiert arbeitet, muß ein Router zuerst einen Tunnel (Label Switched Path) aufsetzen, bevor Daten ausgetauscht werden können. Bei den Tests schaffte der Cisco-Router 5.000 LSPs, während Juniper bereits auf 10.000 kam. Beide Werte sind in Ordnung, können in einem Providernetz aber an ihre Leistungsgrenze stoßen (Bild 4).

Auswertung

Die Ergebnisse wurden in verschiedenen Testszenarien durchgemessen. Sowohl der Aufbau als auch das Routing waren fehlerfrei durchzuführen – sie arbeiten vergleichbar bereits in bestehenden Produkten. Auf diese Weise konnte die Funktionalität des Constraint Based Routing für die Anwendungen des Traffic Engineering und der QoS-Unterstützung vollständig nachgewiesen werden. Der zweite Abschnitt der Messungen zeigt anschließend auf, daß durch MPLS teilweise sehr geringe Verzögerungen umgesetzt werden können, aber auch Paketverluste teilweise in Kauf genommen werden mußten. Diese Verluste entstehen trotz Priorisierung durch Überbelastung, die sich auch durch größere Pufferspeicher nicht komplett unterdrücken lassen. Trotzdem lassen sich Echtzeitanwendungen dadurch wesentlich besser unterstützen, als es durch überzogene Bandbreiten möglich wäre.

Literatur

Detken, K.-O.: Echtzeitplattformen für das Internet – Grundlagen, Lösungsansätze der sicheren Kommunikation mit QoS und VoIP; ISBN 3-8273-1914-5; Addison-Wesley Verlag; München 2002.

Frommherz, D.: Multi Protocol Label Switching – Technische Betrachtung des MPLS und Bewertung seines Einsatzes im WAN-Bereich; Diplomarbeit an der Fachhochschule Darmstadt; Darmstadt 2001.

Newman, D.: Internet Core Router Test – Juniper Networks, Inc., Wins!; Test Report Network Test Inc.; Light Reading; March 2001.

(we)