INSTITUT FÜR INFORMATIK
DER LUDWIG–MAXIMILIANS–UNIVERSITÄT MÜNCHEN

Diplomarbeit

Einführung von Multi-Protocol Label Switching (MPLS) im Münchner Wissenschaftsnetz

Patrick Abeldt

INSTITUT FÜR INFORMATIK
DER LUDWIG–MAXIMILIANS–UNIVERSITÄT MÜNCHEN

Diplomarbeit

Einführung von Multi-Protocol Label Switching (MPLS) im Münchner Wissenschaftsnetz

Patrick Abeldt

Aufgabensteller:	Prof. Dr. Heinz-Gerd Hegering
Betreuer:	Dr. Helmut Reiser Dipl.-Ing. Helmut Tröbs
Abgabetermin:	30. September 2010

Hiermit versichere ich, dass ich die vorliegende Diplomarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

München, den 30. September 2010

. . . . . .
(Unterschrift des Kandidaten)

Abstract

Das Münchner Wissenschaftsnetz (MWN) ist ein regionales Computernetz, das in großer Zahl Standorte überwiegend wissenschaftlicher Einrichtungen in Bayern miteinander verbindet. Die Infrastruktur des Netzes hat sich über viele Jahre hinweg entwickelt und versorgt inzwischen eine sechsstellige Anzahl von Benutzern. Angesichts dieses Wachstums offenbaren einige technische Konzepte bzw. deren bisherige Umsetzung inzwischen Schwächen, u.a. hinsichtlich der Skalierbarkeit. Darüber hinaus sind bestimmte Dienste und Funktionen mit den derzeit eingesetzten Technologien nicht auf geeignete Weise realisierbar. Aus diesen Gründen erwägt der Netzbetreiber den zukünftigen Einsatz von MPLS im Kernnetz des MWN.

In der vorliegenden Diplomarbeit wird die Zweckmäßigkeit eines solchen Einsatzes erforscht und ein Konzept zur Einführung von MPLS erarbeitet. Dazu wird in drei Teilschritten vorgegangen. Nach einer einführenden Betrachtung von MPLS wird im ersten Schritt die Ausgangssituation im MWN analysiert, um einen Überblick für die weitere Planung zu schaffen. Wichtige Aufgaben dabei sind die Identifikation und Erörterung von Anwendungsfällen und die Erfassung relevanter Randbedingungen. Der zweite Schritt besteht in der Realisierungsplanung und umfasst die Exploration konkreter Lösungsansätze zu den Anwendungfällen und die Auswirkungen eines Einsatzes von MPLS auf die existierende Infrastruktur. Im dritten Schritt werden schließlich prototypische Implementierungen erstellt und Fragen zur praktischen Vorgehensweise bei der Migration diskutiert.

Im Verlauf der Arbeit werden sechs Anwendungsfälle gefunden und nach praxisorientierten Kriterien hinsichtlich Machbarkeit und Nutzen untersucht. Sie beinhalten sowohl spezielle Problemstellungen aus dem MWN als auch klassische Einsatzszenarien von MPLS wie z.B. die Umsetzung von Virtual Private Networks (VPN), Verkehrsklassen und Traffic Engineering. Für geeignete Anwendungsfälle werden Lösungen entworfen und Musterkonfigurationen für Hardware der Firma Cisco konstruiert.

Inhaltsverzeichnis

1. Einleitung

Der Betrieb eines regionalen Computernetzes ist eine fortwährende Herausforderung. Dies gilt insbesondere dann, wenn ein solches Netz von mehreren Parteien genutzt wird, die nicht einer gemeinsamen Administration unterstehen. Die Anforderungen der verschiedenen Benutzergruppen hinsichtlich gewünschter Leistungsmerkmale und benötigter Dienste können in einem solchen Szenario sehr heterogen sein. Neben den klassischen Aufgaben des Netzmanagements ist im Betrieb deshalb eine hohe Flexibilität in Bezug auf die Bedürfnisse der zu bedienenden Benutzer von zentraler Bedeutung.

Eine unmittelbare Voraussetzung, um diese Flexibilität effizient erbringen zu können, ist die Unterstützung des Managements durch geeignete technische Mechanismen und Werkzeuge. Das Spektrum an Möglichkeiten, die den eingesetzten Netztechnologien zugrundeliegen, bildet hierfür die Basis.

1.1. Motivation

Das im Rahmen dieser Diplomarbeit betrachtete Münchner Wissenschaftsnetz (MWN) ist ein als Metropolitan Area Network (MAN) zu klassifizierendes Computernetz, das die Standorte verschiedener – in erster Linie wissenschaftlicher – Institutionen miteinander verbindet. Der Betrieb des Netzes obliegt dem Leibniz-Rechenzentrum (LRZ) mit Sitz in Garching bei München. Zu den Benutzern des Netzes gehören vornehmlich öffentliche Einrichtungen (u.a. die beiden Münchner Universitäten). Die einzelnen Standorte befinden sich hauptsächlich im Großraum von München, mit einigen Ausnahmen weiter entfernter Orte in Bayern.

Das MWN kann als ein über viele Jahre organisch gewachsenes Netz charakterisiert werden. Bei der Planung und dem Ausbau des Netzes wurden in der Vergangenheit aus den verfügbaren Technologien jeweils solche ausgewählt, die zu dem betreffenden Zeitpunkt technisch angemessen erschienen, kostengünstig waren und den damaligen Designkriterien genügten. Mit wachsenden Kapazitäten und Teilnehmerzahlen sind inzwischen neue Anforderungen entstanden, die aufgrund von Beschränkungen dieser Technologien teilweise nicht mehr oder nur noch mit verhältnismäßig hohem Aufwand zu erfüllen sind.

Ein mögliches Instrument zur Bewältigung einiger dieser Anforderungen stellt Multi-Protocol Label Switching (MPLS) dar. MPLS ist eine relativ junge Technologie, die aus mehreren proprietären Initiativen der Industrie hervorgegangen ist und mittlerweile weitgehend standardisiert wurde. Einige der im MWN vorhandenen Hardwarekomponenten unterstützen bereits MPLS, so dass ein zeitnaher Einsatz möglich wäre. Die Erforschung der Zweckmäßigkeit eines solchen Einsatzes und die Entwicklung eines Konzepts zur Einführung von MPLS sind die Kernthemen dieser Arbeit.

1.2. Übersicht

Der Inhalt des nächsten Kapitels ist eine Einführung in die technischen Grundlagen, die im Verlauf der Arbeit benötigt werden. Im Wesentlichen geht es dabei um die Funktionsweisen und Eigenschaften von MPLS. In diesem Kontext werden auch die mit MPLS verbundenen Möglichkeiten und Grenzen ergründet.

Der Hauptteil der Diplomarbeit gliedert sich in drei Phasen. Zunächst wird eine Anforderungsanalyse durchgeführt, welche den ersten Schwerpunkt der Arbeit bildet. In einer detaillierten Betrachtung des Szenarios werden hierbei die thematisch relevanten technischen und betrieblichen Aspekte des MWN erfasst. Auf dieser Grundlage werden sinnvolle Einsatzbereiche für MPLS und die sich daraus ergebenden Anforderungen ermittelt.

Im darauffolgenden Abschnitt der Arbeit wird auf der Analyse aufbauend ein Konzept zur Einführung von MPLS entwickelt. Die zuvor identifizierten Einsatzbereiche werden konkretisiert, praxisorientiert erörtert und potentielle Lösungsansätze dafür erarbeitet. Weiterhin werden allgemeine Bedingungen der Einführung von MPLS sowie die Auswirkungen auf die existierende Infrastruktur untersucht. Die gewonnenen Einsichten und Informationen werden schließlich in einem Realisierungsplan zusammengeführt.

Der praktische Teil umfasst den Entwurf und Test von prototypischen Implementierungen. Darüber hinaus werden verschiedene Fragen der Migration erörtert. Eine Evaluation der Ergebnisse im letzten Kapitel schließt die Arbeit ab.

Da die gestalterischen Möglichkeiten in der Netzplanung umfangreich sind und MPLS ein vielseitiges und kreativ einsetzbares Werkzeug darstellt, würde allein der Versuch, alle denkbaren Anwendungen von MPLS im MWN erschöpfend zu betrachten, den Rahmen dieser Arbeit bei weitem sprengen. Darüber hinaus befindet sich das MWN, ebenso wie MPLS als Technologie, in einem Zustand fortwährender Evolution, so dass sich alle Analysen zwangsläufig auf eine zeitliche Momentaufnahme beschränken müssen. Vielmehr ist das Ziel der Arbeit daher, die grundsätzlichen Einsatzmöglichkeiten für MPLS im MWN zu erkunden, ihre Implikationen aufzuzeigen und die elementaren Anwendungstechniken von MPLS im Umfeld des MWN zu demonstrieren.

2. Technische Grundlagen

In diesem Kapitel werden fachliche Begriffe eingeführt und technische Verfahren informell erläutert, die für diese Arbeit relevant sind. Das Kernthema bildet dabei MPLS als standardisierte Technologie. Neben der grundlegenden Funktionsweise von MPLS werden spezielle Konzepte und häufig eingesetzte Verfahren dargestellt, die auf MPLS aufbauen. Einige weitere Betrachtungen in diesem Kapitel dienen darüber hinaus der Beurteilung des Potentials von MPLS und der Einordnung von MPLS in sein technologisches Umfeld.

Am Ende des Kapitels wird kurz auf die Eigenschaften der verfügbaren Hardware der Firma Cisco Systems eingegangen und eine Übersicht über das zum Einsatz kommende Betriebssystem gegeben.

2.1. Multi-Protocol Label Switching (MPLS)

Die Grundidee von MPLS ist die Einbindung von verbindungsorientierten Mechanismen in ein verbindungsloses Netz. MPLS vereint die Vorteile von virtuellen Leitungen mit der Robustheit von Hop-by-Hop-Routing. Bei einer Betrachtung des ISO/OSI-Referenzmodells kann MPLS nicht eindeutig einer Schicht zugeordnet werden. Der Grund hierfür liegt in der Tatsache, dass MPLS sich nicht an das Abstraktionsparadigma des Referenzmodells hält, demzufolge eine Schicht zur Erfüllung seiner Aufgabe nur Dienste und Informationen nutzen darf, die von der unmittelbar darunterliegenden Schicht bereitgestellt werden. Am ehesten könnte MPLS zwischen dem Data Link Layer und dem Network Layer angesiedelt werden. Entsprechend wird MPLS in der Literatur gelegentlich als eigene Schicht angesehen ("Shim Layer" bzw. "Layer 2.5").

Grundsätzlich ist MPLS eine reine Carrier-Technologie, die hauptsächlich in Backbones von Internet Service Providern (ISP) und Großunternehmen eingesetzt wird und für Hostrechner vollständig transparent ist. Zwar steht "Multi-Protocol" für die Fähigkeit, mit jedem beliebigen Protokoll zusammenarbeiten zu können; dennoch hat sich MPLS weniger als ein eigenständiges Verfahren, sondern überwiegend vor dem Hintergrund einer zweckmäßigen Erweiterung von IP entwickelt. Auch im Rahmen der Diplomarbeit wird die Betrachtung von MPLS maßgeblich auf die Anwendung in IP-Netzen beschränkt, da es sich bei dem MWN um ein solches handelt.

Der ursprüngliche Gedanke hinter MPLS war es, die Weiterleitung von Paketen in IP-Routern zu beschleunigen. Dazu sollte die vergleichsweise rechenintensive Analyse des Headers bzw. der Zieladresse eines Pakets, die aufgrund des verbindungslosen Charakters von IP in jedem Router entlang des Übertragungspfades erneut durchgeführt werden muss, überwiegend durch einfache Tabellenzugriffe ersetzt werden. Das Verfahren des Longest-Prefix-Matching von IP-Adressen besitzt eine höhere Komplexität als der Zugriff auf Labels in einer Tabelle und ist daher wesentlich rechenintensiver und zeitaufwendiger. Durch die fortlaufende Weiterentwicklung von ASICs und die weitgehende Implementierung von Hochleistungsroutern in Hardware spielt dieser Geschwindigkeitsvorteil jedoch heute nur noch eine untergeordnete Rolle.

Inzwischen wird MPLS in erster Linie mit dem Ziel der Erbringung von Diensten eingesetzt, die in einem konventionellen IP-Netzwerk nicht verfügbar sind. Der Einsatz von MPLS wirkt sich dabei v.a. auf die für den Netzbetreiber relevanten strukturellen Aspekte des Netzes aus und weniger auf die spezifischen Eigenschaften der Kommunikationsbeziehungen zwischen Hostrechnern. Unter diesem Gesichtspunkt ist MPLS auch klar von verbindungsorientierten Ende-zu-Ende-Protokollen wie z.B. TCP abzugrenzen, da es auf einer wesentlich niedrigeren Ebene in die Datenübertragung eingreift. Da auf dieser Ebene Informationen über die Topologie des Netzes zur Verfügung stehen, kann auch das Routing beeinflusst werden.

Wichtige Einsatzzwecke von MPLS sind die Bildung von virtuellen privaten Netzen (VPN) und die Umsetzung von Traffic Engineering (TE). MPLS ist aus proprietären Verfahren von Cisco (Tag Switching), IBM (Aggregate Route IP Switching, ARIS), Toshiba (Cell Switching), u.a. [OH09] [Bor02] hervorgegangen und wurde seit 2001 von der IETF standardisiert. Die wichtigsten Dokumente dazu sind RFC 3031 [RVC01] und 3032 [RTF+01], in denen die Architektur von MPLS bzw. die Rahmenbildung und -kodierung definiert werden.

2.1.1. Terminologie und Funktionsweise

Ein zusammenhängender Verbund von Routern, die Datenpakete über MPLS weiterleiten, wird als MPLS-Domäne bezeichnet. Die einzelnen Router in einer MPLS-Domäne werden als Label Switching Router (LSR) bezeichnet. Ein LSR, der sich am Rand einer MPLS-Domäne befindet (d.h. beispielsweise am Übergang zu einem konventionellen IP-Netz), heißt Label Edge Router (LER).

Die Paketweiterleitung mit MPLS basiert auf der Markierung von Paketen mit sog. Labels, anhand derer auf den LSR die zur Weiterleitung erforderlichen Ausgangsports bestimmt werden können. Labels sind unstrukturierte/flache numerische Identifikatoren mit einer festen Länge.

Ein Label kann mit einem Paket transportiert werden, indem ein bereits vorhandenes Feld des Rahmens dafür verwendet wird, der das Paket enkapsuliert (z.B. VPI/VCI bei ATM). Falls kein geeignetes Feld zur Verfügung steht (z.B. bei Ethernet), ist die Bildung eines zusätzlichen Rahmens über den sog. MPLS-Header (“Shim Header”) erforderlich (Abbildung 2.1).

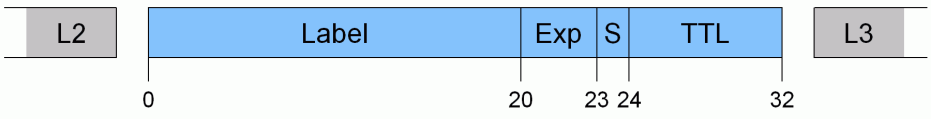

Abbildung 2.1.: Aufbau des MPLS-Headers

Der MPLS-Header wird zwischen die Header der Sicherungsschicht und der Vermittlungsschicht eingefügt. Neben dem Label selbst enthält der Header drei weitere Felder:

- Experimental Bits (Exp): dieses Feld wurde ursprünglich nicht näher spezifiziert, in der Praxis jedoch von Anfang an nahezu ausschließlich zur Unterstützung einer differenzierten Quality of Service (QoS) genutzt. Vor diesem Hintergrund wurde es in jüngerer Vergangenheit von der IETF in RFC 5462 offiziell in “Traffic Class” (TC)

umbenannt. Da in der Literatur aber bisher noch überwiegend die etablierte Bezeichnung vorzufinden ist, wird diese im Rahmen der Arbeit beibehalten.

- S-Bit: dieses Feld kennzeichnet das unterste Label bei Verwendung von Label Stacking (s.u.).
- Time to live (TTL): dieses Feld ist äquivalent zu dem korrespondierenden Feld im IP-Header und hat denselben Zweck. Während ein Paket über MPLS weitergeleitet wird, wird das TTL-Feld des enkapsulierten IP-Headers nicht weiter dekrementiert. Der IETF-Standard schlägt vor, dass das TTL-Feld des MPLS-Headers die Anzahl der Hops innerhalb der MPLS-Domäne widerspiegeln soll, d.h. bei Eintritt in die Domäne sollte der TTL-Wert des IP-Headers vom MPLS-Header übernommen werden, und bei Verlassen der Domäne der aktualisierte Wert zurück in den IP-Header kopiert werden. Dies ist aber im Standard nicht zwingend vorgeschrieben, da auch Policies denkbar sind, die damit in Konflikt stehen würden. Dieser Sachverhalt kann bei der Fehlersuche in einem Netz von Bedeutung sein.

Das Label wird beim Forwarding verwendet, um den Ausgangsport für das betreffende Paket zu bestimmen. Vereinfacht betrachtet speichert ein LSR in einer Tabelle zu jedem Label für eingehende Pakete (Eingangslabel) ein neues Label (Ausgangslabel) und die Adresse des Next-Hop-LSR. Bei Ankunft eines Pakets im LSR werden die Label entsprechend dieser Tabelle getauscht (Label Swapping) und das Paket dann auf dem zugeordneten Ausgangsport weitergeleitet. Die Label müssen getauscht werden, da sie jeweils nur lokal zwischen zwei LSRs gültig sind. Die Vereinbarung der Labels in einem Netz wird dadurch dezentralisiert und ihre Verteilung vereinfacht.

Zur Übertragung von Paketen innerhalb einer MPLS-Domäne sind zunächst Übertragungspfade einzurichten, sog. Label Switched Paths (LSP). Konzeptionell entspricht dieser Vorgang einem Verbindungsaufbau; dabei werden die Einträge für die o.g. Tabelle erstellt. Ein LSP der Länge n kann definiert werden durch eine Sequenz von $n+1$ LSRs, geschrieben als

$$< R_0, R_1, ..., R_n >$$

für die gilt ($\forall i \in \{0, ..., n-1\}$):

- R_i und R_{i+1} sind benachbart, d.h. zwischen ihnen existiert eine direkte Übertragungsstrecke (physisch oder logisch/transparent), und
- R_i und R_{i+1} haben das zwischen ihnen zu verwendende Label für diesen LSP miteinander ausgehandelt.

LSPs sind unidirektional, wobei der Router R_0 als LSP-Ingress und der Router R_n als LSP-Egress bezeichnet wird. Diese Rollen beziehen sich jeweils auf einen bestimmten LSP, d.h. beispielsweise kann ein Router aus Sicht eines oder mehrerer LSPs der LSP-Ingress sein, und aus Sicht eines anderen LSP lediglich ein zwischenliegender LSR. Der LSP-Ingress hat die Aufgabe, an ein in den LSP einzuschleusendes Paket das entsprechende Label anzuhängen (Label Push). Entlang des LSP wird vor jeder Weiterleitung ein Label Swap durchgeführt. Der LSP-Egress muss das Label schließlich wieder entfernen (Label Pop). Der Label Pop kann auch bereits in R_{n-1} durchgeführt werden, da der LSP-Egress es zur Weiterleitung

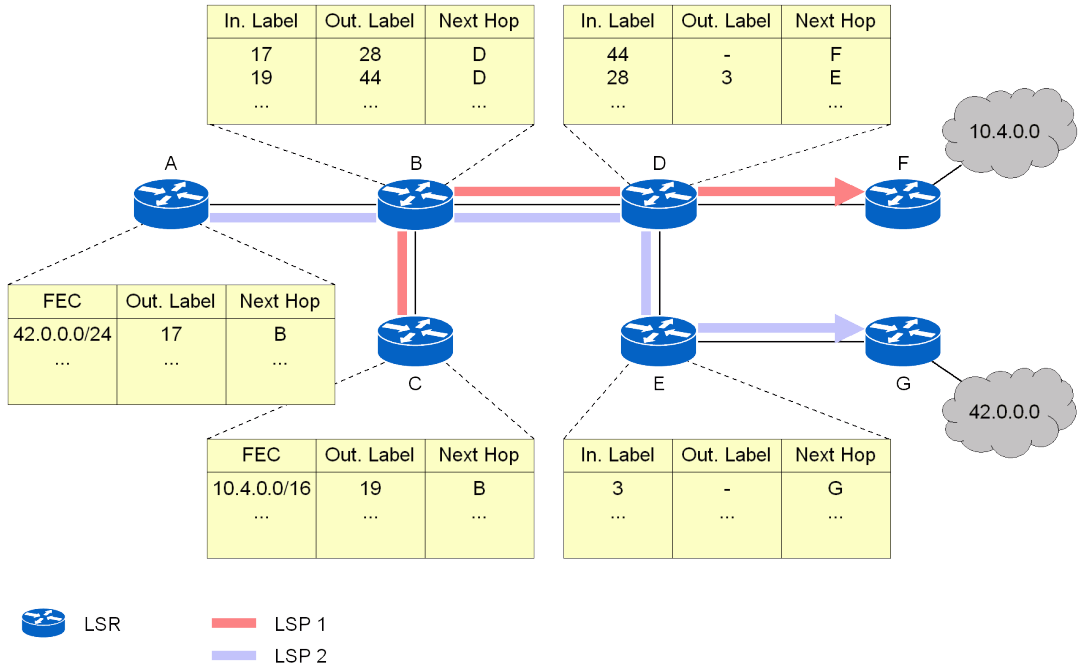

Abbildung 2.2.: Paketweiterleitung mit MPLS

nicht mehr benötigt (Penultimate Hop Popping, PHP); dies ist i.A. die Standardmethode (Abbildung 2.2).

Das von einem LSP-Ingress zugewiesene Label legt den LSP und damit den zukünftigen Pfad eines Pakets fest. Damit ein LSP-Ingress das korrekte Label für ein IP-Paket bestimmen kann, muss das Paket einer sog. Forwarding Equivalence Class (FEC) zugeordnet werden. FECs sind ein abstraktes Konzept zur Beschreibung von Mengen von Paketen. Eine FEC repräsentiert nach Definition [RVC01] auf einem LSR eine Menge von Paketen, die bzgl. der Weiterleitung identisch zu behandeln sind. Um bestimmte Pakete über einen LSP weiterzuleiten, wird er mit der diesen Paketen entsprechenden FEC assoziiert. Das einem Paket zugewiesene Label repräsentiert somit seine FEC.

Die Klassifikation eines Pakets wird einmalig im LSP-Ingress durchgeführt, und beim weiteren Durchlaufen des LSP wird nur noch sein jeweils aktuelles Label betrachtet. Pakete, die an einem gegebenen Punkt des LSP das gleiche Label besitzen, werden folglich identisch behandelt und sind daher bis zum Verlassen des LSP nicht mehr unterscheidbar (FEC-Gültigkeitsbereich).

Eine FEC kann nach vielfältigen Kriterien konstruiert werden, z.B. nach IP-Prefixen von Quell- bzw. Zielnetz, nach dem TOS-Feld des IP-Headers, nach der Größe des Pakets, etc. Denkbar ist auch die Nutzung von Protokollinformationen höherer Schichten, wie z.B. TCP-/UDP-Portnummern. Die Einflussfaktoren, die Komplexität und die Granularität einer FEC können theoretisch beliebig gewählt werden; entscheidend für die Zuordnung ist, dass die Weiterleitung der entsprechenden Pakete innerhalb des FEC-Gültigkeitsbereichs auf dieselbe Art geschehen soll (d.h. über dieselben Ausgangsports, mit derselben Priorität, etc.). Dies impliziert, dass Pakete, die an einem LSP-Ingress derselben FEC zugeordnet werden, an völlig unterschiedliche Zielnetze adressiert sein können.

LSPs können ineinander verschachtelt werden; man erhält dann einen Stapel von Labels (Label Stack). Bei Verwendung des MPLS-Headers sieht das derart aus, dass bei Eintreten in einen LSP ein weiterer Header vor den/die ggf. bereits vorhandenen MPLS-Header eingefügt wird. Die Pakete werden dann am LSP-Ingress des verschachtelten LSP anhand ihres bisherigen Labels in eine FEC eingeordnet. Für die Weiterleitung wird jeweils nur das oberste Label (Top Label) betrachtet. Nach Verlassen eines LSP wird der äußerste Header entfernt und das Paket auf Basis des Label der nächstniedrigeren Ebene weitergeleitet bzw. falls das unterste Label des Stapels (Bottom Label) entfernt wurde, auf konventionelle Weise anhand der IP-Zieladresse. Auf diese Weise können Hierarchien von LSPs gebildet werden.

Das Stapeln mehrerer Labels kann auch auf ein und demselben Netzknoten geschehen, d.h. es wird dann gleich ein kompletter Label Stack an ein Paket angehängt. Eine Reihe interessanter Funktionalitäten von MPLS wird erst durch diese Technik realisierbar. Das für die Weiterleitung relevante Top Label wird in diesem Zusammenhang auch als Transport-Label bezeichnet, während darunterliegende Labels beispielsweise zur Kennzeichnung von verschiedenen Diensten verwendet werden können und dann als Service-Label bezeichnet werden.

Es ist zulässig, dass ein LSR mehreren unterschiedlichen Eingangslabels dasselbe Ausgangslabel zuordnet (Label Merging). Bei diesem Vorgang geht Information verloren, da alle nachfolgenden LSRs die Pakete identisch behandeln. Jedoch kann die Anzahl der benötigten Labels so reduziert werden. Bei Verwendung von Label Merging bilden die betroffenen LSPs eine Baumstruktur mit dem LSP-Egress als Wurzel.

Es ist ebenfalls zulässig, für ein Eingangslabel verschiedene Ausgangsports zu definieren [RVC01]. Das Ausgangslabel hängt dann von dem letztendlich selektierten Port ab. Dies kann theoretisch für folgende Zwecke genutzt werden:

- ECMP-Routing (dazu müssen die LSPs an einem späteren Punkt wieder zusammengeführt werden, z.B. über Label Merging)
- Punkt-zu-Mehrpunkt-Kommunikation (z.B. zur Abbildung von IP-Multicasting; die Pakete müssen dann an den Verzweigungspunkten repliziert werden)

Eine MPLS-Domäne kann sich theoretisch über mehrere Autonome Systeme (AS) erstrecken. Dies ist aber mit einigen Komplikationen verbunden [Bor02] und im Rahmen der Arbeit auch nicht erforderlich, weshalb dieser Aspekt hier nicht weiter betrachtet wird.

2.1.2. Verwendung von LSPs

Prinzipiell kann Datenverkehr über zwei Techniken in einen LSP geleitet werden. Die einfachste Variante ist statisches Routing, d.h. die explizite Auswahl der Daten, die über den LSP geführt werden sollen. Der LSP hat dann über dieses statische Routing hinaus keine Auswirkungen auf das Netz.

LSPs können jedoch auch in das IGP-Routing integriert werden. Ein solcher LSP erscheint dann in der Routing-Tabelle des LSP-Ingress als eine direkte Verbindung zu dem LSP-Egress und wird wie eine physische Übertragungsstrecke behandelt. In analoger Weise erhält der LSP dabei einen Kostenwert, der maßgeblich für seine Berücksichtigung beim Routing ist. Dieser kann aus den Kosten der zugrundeliegenden Übertragungsstrecken abgeleitet werden.

Beim Traffic Engineering wird dieser Kostenwert aber auch oft modifiziert, um die Bevorzugung eines LSP bei der Wegewahl zu bewirken, d.h. beispielsweise wird er relativ zu

den IGP-Kosten oder auch willkürlich festgelegt. Er wird dann als TE-Metrik bezeichnet.[1] Existieren zu einem Ziel sowohl eine konventionelle Route als auch ein LSP, so wählt das IGP den LSP genau dann zur Weiterleitung aus, wenn dessen Kosten kleiner oder gleich denen der konventionellen Route sind.

2.1.3. Virtual Private Networking (VPN)

Eine der häufigsten Anwendungen von MPLS ist die Bereitstellung von VPNs. Generell besteht der Zweck eines VPN darin, seinen Datenverkehr und seine Adressräume von anderen VPNs, vom Netz des VPN-Providers und evtl. auch vom öffentlichen Internet zu trennen. MPLS wird hierbei verwendet, um den Datenverkehr zwischen den einzelnen VPN-Standorten durch das Backbone des ISP zu transportieren.

Bei der Betrachtung von VPNs – aber auch generell von MPLS-Anwendungsszenarien in einem ISP-Umfeld – sind i.A. die Geräte an den Übergängen zwischen den Netzen von Nutzern und dem Netz des ISP von besonderem Interesse. Diese werden als Customer Edge Device (CE), respektive als Provider Edge Device (PE), bezeichnet. Bei PE-Geräten handelt es sich üblicherweise um Router (im Kontext von MPLS genauer um LERs). Geräte innerhalb des ISP-Netzes, die keine Nutzer direkt bedienen, werden als Provider Core Device (P) bezeichnet (Abbildung 2.3).

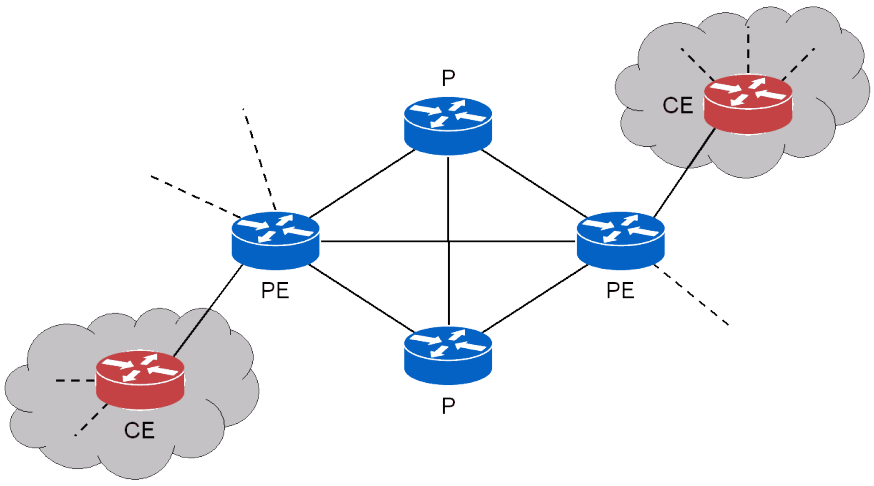

Abbildung 2.3.: P-, PE- und CE-Geräte

Diese Begriffe sind dabei nicht als absolut zu verstehen, sondern beziehen sich auf verschiedene Rollen, die ein Gerät je nach gegebener Situation einnehmen kann. Dies wird z.B. in Carrier-to-Carrier-Verkehrsbeziehungen deutlich.

Zur Bildung eines VPN werden die Teilnetze des VPN-Nutzers jeweils bidirektional über

[1] In der Literatur werden gelegentlich auch andere Typen von Kostenwerten als TE-Metrik bezeichnet, die im Zusammenhang mit Traffic Engineering verwendet werden. Insbesondere in Dokumentationen von Geräteherstellern bestehen oft diffizile Unterschiede.

LSPs miteinander verbunden, so dass die Infrastruktur des ISP vollständig transparent erscheint. Wie dies im Detail aussieht, hängt von der Art des VPN ab. Im Folgenden werden die Charakteristika der wichtigsten Typen von VPNs dargestellt, die mit MPLS realisiert werden können.

Layer 3-VPN

In einem Layer 3-VPN werden die Teilnetze auf der Vermittlungsschicht miteinander verbunden. Aus Sicht des Nutzers erscheint das Netz des ISP als ein einzelner virtueller Router, an dem die Teilnetze direkt angeschlossen sind. Um die IP-Adressräume verschiedener VPNs voneinander zu isolieren, werden auf den involvierten PE-Routern getrennte Routing-Domänen (Virtual Routing and Forwarding, VRF) eingerichtet. PE- und CE-Router tauschen über BGP untereinander Routing-Informationen für die Ziele in den Teilnetzen aus. Dieser Typ von VPN wird daher auch oft als BGP/MPLS-VPN bezeichnet.

Ausgehend davon kann ein PE-Router für jede VRF eine separate Routing-Tabelle erstellen. Im Normalfall leiten die PE-Router nur Pakete innerhalb derselben VRF weiter. Dadurch sind Überlappungen der Adressräume verschiedener VPNs zulässig. Sollen Daten zwischen unterschiedlichen VPNs ausgetauscht werden, sind i.A. zusätzliche Mechanismen erforderlich [OH09]. Es ist grundsätzlich auch möglich, ein Teilnetz mehreren verschiedenen VPNs zuzuordnen.

Die Weiterleitung von Paketen zwischen den PE-Routern erfolgt über LSPs zusammen mit VRF-spezifischen Service-Labels. Für das Kernnetz des ISP sind die BGP-Routen der Nutzer daher unsichtbar.

Layer 3-VPNs via BGP sind der älteste standardisierte Typ von MPLS-basierten VPNs und wurden in RFC 2547 [RR99] und 4364 [RR06] spezifiziert.

Any Transport over MPLS (AToM)

Wie zu Beginn des Kapitels erwähnt können über MPLS nicht nur IP-Pakete, sondern prinzipiell beliebige Daten transportiert werden (theoretisch auch MPLS-Rahmen selbst). Angewandt auf Protokolle der Bitübertragungs- und Sicherungsschicht wird dieses Konzept allgemein als AToM bezeichnet.

Die spezielle Ausprägung von AToM für Ethernet wird als Ethernet over MPLS (EoMPLS) bezeichnet. Hierüber können Layer 2-VPNs zur transparenten Punkt-zu-Punkt-Verbindung zweier Ethernet-LANs realisiert werden. Aus Sicht des Nutzers erscheint dabei das Netz des ISP als ein virtueller “Draht”. Zur Weiterleitung von Ethernet-Rahmen werden diese in MPLS-Rahmen enkapsuliert; die ihrerseits von den Ethernet-Rahmen enkapsulierten Protokolle höherer Schichten sind dabei irrelevant. In den Teilnetzen ggf. vorhandene CE-Router des Nutzers sind für den ISP unsichtbar.

AToM basiert auf Pseudo Wire Emulation Edge-to-Edge (PWE3), einem standardisierten Verfahren zur transparenten Übertragung verschiedener Layer 1- und Layer 2-Protokolle über IP- oder MPLS-Netze. Die Weiterleitung von Rahmen mittels MPLS erfolgt über mit PWE3-Verbindungen assoziierte LSPs. PWE3 ist u.a. in RFC 3916 und 3985 spezifiziert und unterstützt neben Ethernet z.B. Frame Relay, PPP, ATM und SDH [OH09].

Virtual Private LAN Service (VPLS)

VPLS stellt die Verallgemeinerung von EoMPLS auf Mehrpunkt-Konnektivität dar, so dass damit Ethernet-LAN-Dienste realisiert werden können. Analog zu einem Layer 3-VPN erscheint hier das Netz des ISP aus Sicht des Nutzers wie ein einzelner virtueller Switch, an dem die Teilnetze direkt angeschlossen sind.

Die PE-Router sind über PWE3-Verbindungen bzw. deren LSPs vollvermascht verbunden (dient zusammen mit dem Split-Horizon-Algorithmus zur Schleifenvermeidung). Der virtuelle Switch wird als Virtual Forwarding Instance (VFI) bezeichnet und verhält sich ähnlich wie ein physischer Switch, d.h. MAC-Adressen werden von den PE-Routern verteilt gelernt und Rahmen an unbekannte Ziele, vergleichbar mit einem Broadcast, an alle Teilnetze versendet.

Analog wie bei einem Layer 3-VPN ist die VFI für die P-Router unsichtbar. Signalisierungsverfahren für VPLS sind in RFC 4761 und 4762 standardisiert.

Die Vollvermaschung der PE-Router mit PWE3-Verbindungen kann einen begrenzenden Faktor für die Skalierbarkeit von VPLS darstellen. In dieser Richtung wurden daher in jüngerer Vergangenheit neue Ansätze entwickelt. Eine nennenswerte Erweiterung stellt H-VPLS dar, bei der durch eine hierarchische Unterteilung der Vermaschung die Anzahl der benötigten LSPs deutlich verringert werden kann.

Vergleich Layer 3- vs. Layer 2-VPN

Ein maßgebliches Argument für Layer 3-VPNs ergibt sich aus deren Fähigkeit, Teilnetze unabhängig von den dort auf der Sicherungsschicht eingesetzten Technologien miteinander verbinden zu können. Weiterhin kann sich für den Nutzer der Betrieb seiner Netze vereinfachen, da er die Verantwortung für das WAN-Routing teilweise an den ISP abgeben kann.

Demgegenüber sind Layer 2-VPNs unabhängig von den innerhalb des VPN verwendeten Protokollen der Vermittlungsschicht, was zum Betrieb von Altsystemen vonnöten sein kann. Ein interessanter Vorteil für den ISP ist, dass aus seiner Sicht die Inbetriebnahme und Pflege eines Layer 2-VPN mit einem wesentlich geringeren Aufwand als bei einem Layer 3-VPN verbunden ist, da hier kein BGP-Peering stattfindet und somit eine klare funktionale Abgrenzung zwischen seinem Netz und denen des Nutzers besteht. Die Fehlersuche gestaltet sich daher zumeist einfacher, erfordert keinen so hohen Grad an Expertenwissen zu Routing-Protokollen und kann weitgehend ohne Einwirkung des Nutzers durchgeführt werden. Die Hardware-Anforderungen an die PE-Router sind zudem i.A. niedriger.

Beide Konzepte skalieren verhältnismäßig gut, da die VPN-relevanten Strukturen jeweils auf den PE-Routern und somit dezentral verwaltet werden. Bei einer Gegenüberstellung der Skalierbarkeit sind zwei Dimensionen zu betrachten, zum einen die mögliche Größe eines einzelnen VPN, und zum anderen die mögliche Anzahl an VPNs im Netz eines ISP:

- Der limitierende Faktor von Layer 3-VPNs in beide Dimensionen ist die Größe der VRF-spezifischen Routing-Tabellen. Hier kann es beispielsweise problematisch werden, wenn ein Nutzer sehr viele Routen in seinem Netz vorhalten möchte. VRFs sind inhärent speicher- und rechenintensiv.

- Die Skalierbarkeit eines einzelnen Layer 2-VPN ist, sofern der Nutzer in den Teilnetzen keine eigenen Router betreibt oder eine anderweitige Strukturierung vornimmt, durch das praktikable Ausmaß einer einzelnen Broadcast-Domäne beschränkt. Für die Gesamtanzahl an möglichen VPNs ist dabei die Anzahl von Hosts je VPN als kritische

Größe einzustufen, da alle MAC-Adressen in der Broadcast-Domäne von den beteiligten PE-Routern gelernt werden müssen. Diese Beschränkungen entfallen jedoch, falls Router in den Teilnetzen vorhanden sind – dann skalieren Layer 2-VPNs in beide Dimensionen deutlich besser als Layer 3-VPNs.

Grundsätzlich sind beide Ansätze als gleichberechtigt anzusehen; welcher von ihnen in einem konkreten Fall besser geeignet ist, hängt sehr stark von dem gegebenen Einsatzszenario ab. Hinsichtlich der unterstützten Topologien existieren keine signifikanten Unterschiede. Eine ausführliche Übersicht findet sich z.B. in [Abd02] und [mc006]. In der Industrie ist aufgrund der betrieblichen Vorteile für ISPs ein wachsender Trend zu Layer 2-VPNs, insbesondere zu VPLS, zu beobachten [nor03] [alc09].

2.1.4. Traffic Engineering (TE)

MPLS kann in einem IP-Netz verwendet werden, ohne dass sich die dabei benutzten Datenübertragungswege von der Routing-Tabelle des IGP unterscheiden. In einem solchen Fall werden die Routen für LSPs auf konventionelle Weise bestimmt, d.h. für gewöhnlich über einen SPF-Algorithmus. Dies kann realisiert werden, indem z.B. jeder Eintrag in der Routing-Tabelle als eine FEC betrachtet wird [OH09]; es ändert sich dann lediglich die Methode des Forwardings. Eine solche Vorgehensweise ist besonders dann geeignet, wenn man lediglich an MPLS-Diensten interessiert ist, die keine speziellen Routen erfordern. Zur Signalisierung von LSPs wird dann das Label Distribution Protocol (LDP) eingesetzt, das in RFC 3036 [ADF$^+$01] u.a. standardisiert wurde.

MPLS kann jedoch auch mit explizitem Routing eingesetzt werden, d.h. der Pfad für einen LSP kann auch abweichend von der kürzesten Route angelegt werden. Dieses Konzept kann zur Verkehrssteuerung genutzt werden, um bestimmte Ziele der Netzplanung zu erreichen. Da ein LSP konzeptionell eine virtuelle Leitung (VC) darstellt, ist TE eine natürliche Anwendung von MPLS.

Derart erstellte LSPs werden auch als TE-LSP oder TE-Tunnel bezeichnet.[1] Die primitivste Möglichkeit zur Einrichtung eines solchen Pfades ist die statische Konfiguration, d.h. die Angabe aller LSR, über die der LSP geführt werden soll. Andere Varianten sind halbautomatische Verfahren wie Constrained Based Routing (CBR) bzw. Constrained SPF (CSPF), die bei der Berechnung des besten Pfades vorgegebene Randbedingungen einbeziehen. Solche Bedingungen können z.B. eine minimal verfügbare Bandbreite sein oder bestimmte Wegpunkte, die zwingend durchlaufen bzw. vermieden werden müssen.

Damit die Bedingungen mittels CSPF überprüft werden können, müssen zusätzliche TE-relevante Informationen im Netz verteilt werden. Zu diesem Zweck wurden Erweiterungen gängiger IGPs entwickelt, z.B. TE Extensions for OSPF (OSPF-TE) oder IS-IS Extensions for Traffic Engineering (IS-IS-TE).

Verfügbare Protokolle zur Signalisierung der LSPs selbst sind Constraint Based LDP (CR-LDP), eine Erweiterung von LDP, sowie TE Extensions to RSVP for LSP Tunnels (RSVP-

[1] Die Begriffe TE-Tunnel und TE-LSP werden – auch im fachlichen Sprachgebrauch – oft synonym verwendet, sind aber genau genommen zu unterscheiden. Zwar besteht eine 1 : 1-Beziehung zwischen beiden Entitäten, allerdings handelt es sich bei einem TE-Tunnel um ein virtuelles Konstrukt, während ein TE-LSP dessen konkrete Ausprägung repräsentiert. Dies zeigt sich z.B. darin, dass einem TE-Tunnel während seiner Lebensdauer verschiedene TE-LSPs bzw. verschiedene Pfade zugewiesen sein können.

TE), eine Erweiterung von RSVP. CR-LDP wurde in RFC 3212 u.a. und RSVP-TE in RFC 3209 [ABG+01] u.a. standardisiert.

Hybride Szenarien zwischen MPLS mit IGP-Routing und MPLS mit explizitem Routing sind möglich. In vielen Fällen wird es ausreichend sein, TE-Mechanismen nur an bestimmten kritischen Orten eines Netzes einzusetzen. Da die Konfiguration von TE-LSPs einen gewissen Aufwand mit sich bringt, kann so die Komplexität gering werden.

Bei der Integration von TE-LSPs in das IGP-Routing bestehen zwei Möglichkeiten. Sie können entweder lediglich lokal (d.h. nur sichtbar für den LSP-Ingress) verwendet oder auch an benachbarte Router propagiert werden. Für letztere ist dann nicht unterscheidbar, ob es sich bei einer mitgeteilten Übertragungsstrecke um eine direkte physische Verbindung handelt oder um einen LSP. Zusammen mit der Variation von TE-Metriken kann das Gesamtverhalten eines Netzes auf diese Weise erheblich beeinflusst werden.

2.1.5. MPLS Class of Service (CoS)

Seit ein rasant steigender Anteil von über das öffentliche Internet übertragenen Daten multimedialer Natur ist, treten spezifische Dienstgütemerkmale wie minimale Datentransferrate, Latenz und Jitter im Kontext der Netzplanung zunehmend in der Vordergrund. Als verbindungslose Technologie können reine IP-Netze grundsätzlich keine festen Zusicherungen für solche Parameter abgeben, so dass dafür zusätzliche Mechanismen notwendig sind. In der Vergangenheit hat es verschiedene Initiativen in diese Richtung gegeben, z.B. IP Integrated Services (IntServ) oder IP Differentiated Services (DiffServ).

MPLS CoS definiert keine eigene QoS-Architektur, sondern baut auf dem IP DiffServ-Modell auf [Alv06]. Das zugrundeliegende Prinzip hierbei ist die Unterteilung des Datenverkehrs in verschiedene Verkehrsklassen, die jeweils unterschiedliche Anforderungen an die Dienstgüte repräsentieren. Verkehrsklassen bilden diskrete Prioritätsstufen ab.

Bevor der Datenverkehr einer differenzierten Verarbeitung zugeführt werden kann, ist eine Zuordnungsfunktion erforderlich. In einem verbindungslosen Netz wird dies i.A. durch eine Markierung erreicht, d.h. das Versehen des Pakets mit einem klassenspezifischen Identifikator. Die korrekte Markierung wird durch eine Klassifikation ermittelt, d.h. die Auswertung des Pakets nach administrativen oder technischen Kriterien.

Aufgrund dieser zweiteiligen Vorgehensweise muss die Klassifikation – ein potentiell rechenintensiver Prozess – nur einmalig durchgeführt werden und kann daher am Rand des Netzes erfolgen, wo sich oft aggregierende Netzknoten mit geringeren Datentransferraten als im Kernnetz befinden. Den Hochgeschwindigkeitsknoten im Kernnetz fällt lediglich die vergleichsweise einfache Aufgabe zu, die Klassenidentifikatoren zu analysieren und entsprechende Regeln auf die Pakete anzuwenden (Per-Hop-Behaviour, PHB).

Die Zugehörigkeit zu einer Verkehrsklasse (PHB Scheduling Class, PSC) kann mit MPLS über zwei Methoden abgebildet werden. Bei der ersten Methode werden Verkehrsklassen über das Exp-Feld des MPLS-Headers voneinander unterschieden und können gemeinsam auf demselben LSP übertragen werden ($n : 1$). Ein solcher LSP wird als Exp-Inferred-PSC LSP (E-LSP) bezeichnet. Das Exp-Feld entspricht hierbei der Markierung. Die Anzahl der Klassen ist dabei entsprechend des mit 3 Bit darstellbaren Wertebereichs auf 8 beschränkt. Neben einer freien Gestaltung von Prioritätsstufen ist auch eine Abbildung der Precedence-Bits aus dem TOS-Feld des enkapsulierten IP-Headers bzw. des DSCP bei IP DiffServ möglich.

Bei der zweiten Methode wird jeweils eine Verkehrsklasse auf einen LSP abgebildet ($1 : 1$), d.h. die Klassen werden über das Label selbst unterschieden. Ein solcher LSP wird als Label-

Inferred-PSC LSP (L-LSP) bezeichnet. Mischformen der beiden Methoden sind ebenfalls möglich.

Es ist an dieser Stelle wichtig festzuhalten, dass eine MPLS-Infrastruktur für sich genommen noch keine QoS-Funktionen bereitstellt [OH09] [Pep07a]. Beispielsweise können mit MPLS-TE alleine keine Garantien bzgl. Bandbreite oder Latenz für einen LSP gegeben werden. Hierzu sind weitere verkehrskonditionierende Mechanismen wie z.B. Traffic Shaping notwendig.

2.1.6. MPLS Fast Reroute (FRR)

MPLS-FRR ist ein Verfahren zur Verringerung von Ausfallzeiten, wie sie bei der Konvergenz des Netzes nach einer Topologieänderung auftreten. Kommt es zu einer Störung in einem Netzelement, so kann ein über dieses Element laufender LSP ggf. keine Daten mehr weiterleiten. Um seine Funktionstüchtigkeit wiederherzustellen, ist ein Ab- und anschließender Neuaufbau des gesamten LSP auf einem alternativen Pfad notwendig. Sinnvollerweise geschieht dies erst nach der Konvergenz des Netzes; währenddessen kommt es vorübergehend zu einer Unterbrechung des betroffenen Datenverkehrs.

Diese Phase kann mit MPLS-FRR durch Umleitung des Datenverkehrs auf vorbereitete Backup-LSPs überbrückt werden. Dabei kann zwischen zwei Konzepten unterschieden werden:

- FRR Link Protection überbrückt den Ausfall einer Übertragungsstrecke zwischen zwei LSR. Dazu wird vorab für jede einzelne Übertragungsstrecke, die ein zu schützender LSP durchläuft, ein Backup-LSP zwischen den zwei an der Strecke angrenzenden LSRs eingerichtet.

- FRR Node Protection überbrückt den Ausfall eines LSR. Ähnlich wie bei Link Protection wird vorab für jeden LSR, den ein zu schützender LSP durchläuft, ein Backup-LSP zwischen dem vorangehenden und dem nachfolgenden LSR eingerichtet. Vereinfacht betrachtet überbrückt Node Protection zwei konsekutive Übertragungsstrecken im Gegensatz zu einer bei Link Protection.

Bei einem Ausfall wird ein betroffener LSP mittels Label Stacking auf den Backup-LSP umgeleitet (Abbildung 2.4). Nach Erreichen des LSP-Egress des Backup-LSP wird die weitere Übertragung auf dem ursprünglichen LSP fortgesetzt. Die Pakete erscheinen dabei identisch, unabhängig davon, welchen Weg sie genommen haben.[1] Der Backup-LSP wird solange genutzt, bis das IGP konvergiert ist und ein neuer Pfad für den ursprünglichen LSP etabliert wurde. (Eine dauerhafte Nutzung ist keine Option, da er möglicherweise zur Absicherung noch weiterer LSPs dient und zudem der resultierende Gesamtpfad meist suboptimal ist.)

MPLS-FRR skaliert sehr gut (insbesondere Link Protection), da mit einem einzelnen Backup-LSP beliebig viele LSPs geschützt werden können. Da die Umleitung ausschließlich lokal geregelt werden kann, sind sehr schnelle Umschaltzeiten möglich. Nach Angaben von Geräteherstellern [Alv08] sind Größenordnungen von ca. 50 ms erreichbar. Zu bedenken

[1] Zwingende Voraussetzung hierfür ist eine entsprechende Konfiguration der Label-Allokation, so dass alle Labels auf einem LSR global gültig sind und nicht per Schnittstelle. In letzterem Fall kann MPLS-FRR aufgrund von Uneindeutigkeiten bei der Label-Zuordnung nicht genutzt werden.

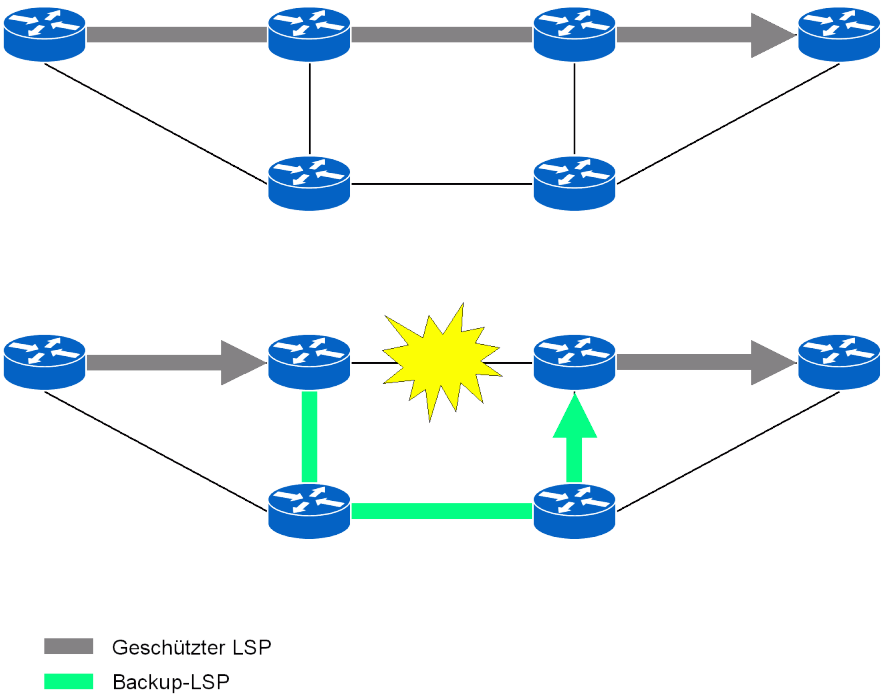

Abbildung 2.4.: Lokale Umleitung bei MPLS-FRR Link Protection

ist, dass neben der Zeit zur Anpassung der Forwarding-Tabellen im LSR die Zeit zur Erkennung eines Ausfalls einen begrenzenden Faktor darstellt. Technologien der Bitübertragungs- bzw. Sicherungsschicht, die spezielle Mechanismen dafür unterstützen (z.B. SONET/SDH), bzw. gesonderte Protokolle zu diesem Zweck (z.B. BFD) sind hier vorteilhaft.

Eine offensichtliche Restriktion von MPLS-FRR ist, dass nur LSP-Datenverkehr geschützt werden kann, aber nicht z.B. parallel dazu existierender konventioneller IP-Datenverkehr. MPLS-FRR kann außerdem nur zusammen mit MPLS-TE verwendet werden, d.h. ein Schutz von über LDP signalisierten LSPs ist nicht vorgesehen.

Diese Einschränkungen können aber umgangen werden, indem für jede betroffene Übertragungsstrecke ein TE-LSP der Länge 1 errichtet wird. Diese Tunnel werden bei Wahl einer geeigneten TE-Metrik vom IGP zur Weiterleitung bevorzugt und können schließlich wie andere TE-LSPs auch geschützt werden. Der Pfad der Weiterleitung ändert sich im Normalbetrieb nicht, es findet lediglich über jeden Hop eine Ver-/Entkapselung statt. Ein signifikanter Leistungsverlust bzgl. des Durchsatzes hierdurch ist aufgrund der Ausführung in Hardware eher nicht zu erwarten; in der Tat ist dies die von Cisco vorgeschlagene Vorgehensweise [Alv08] [MW09].

Ein drittes Überbrückungskonzept, das jedoch eher selten zur Anwendung kommt, wird als MPLS Path Protection bezeichnet. Es ist MPLS-FRR insofern ähnlich, als dass auch hier vorab ein Backup-LSP eingerichtet wird, auf den bei einem Ausfall schnell umgeschaltet werden kann. Während es sich aber bei MPLS-FRR um ein lokales Verfahren handelt, bei dem ein einzelner Backup-LSP beliebig viele LSPs schützen kann $(1 : n)$, stellt der bei MPLS Path Protection erzeugte Backup-LSP eine vollständige, disjunkte und exklusive Alternativroute für den ursprünglichen LSP dar $(1 : 1)$. Die Skalierbarkeit ist daher eingeschränkt; darüber

hinaus sind die Umschaltzeiten nicht so gering wie bei MPLS-FRR, da für den Wechsel auf den Backup-LSP eine Signalisierung zum LSP-Ingress erforderlich ist. In bestimmten Topologien kann MPLS Path Protection aber besser geeignet sein als MPLS-FRR.

2.1.7. Vergleich mit anderen Technologien

Die nachfolgenden Abschnitte stellen MPLS zum Vergleich spezifischer Aspekte einigen ausgewählten anderen Technologien gegenüber.

ATM

Auf den ersten Blick besitzt MPLS einige Ähnlichkeiten mit ATM. Bei beiden handelt es sich um verbindungsorientierte Transportverfahren, und zwischen Labels und VPI/VCI sind Analogien ersichtlich.

Dennoch bestehen gravierende Differenzen. MPLS und ATM liegen jeweils unterschiedliche Konzepte und Ziele zugrunde. Am deutlichsten spiegelt sich dies darin wider, dass ATM ein eigenes Kommunikationsmodell besitzt, dass bis auf die Bitübertragungsschicht hinunter definiert ist, während MPLS mit beliebigen Protokollen niedrigerer Schichten eingesetzt werden kann.

Ein markanter technischer Unterschied besteht in der Rahmenbildung. Während bei MPLS ein kleiner Header an ein theoretisch beliebig großes Paket angehängt wird, werden bei ATM kleine Zellen einer festen Länge gebildet. Hierdurch kann ATM bestimmte QoS-Garantien erbringen wie z.B. einen konstante Latenz, die für interaktive Echtzeitanwendungen von Bedeutung sind. Der Preis hierfür ist ein hohe Komplexität in der ATM-Anpassungsschicht, da Aufgaben wie Segmentierung und Reassemblierung anfallen. Die für ATM erforderliche Hardware ist entsprechend auch mit deutlich höheren Anschaffungskosten verbunden.

MPLS ist einfacher konzipiert, kann aber solche absoluten Garantien nicht leisten. Aus heutiger Sicht liegt der größte Vorteil von MPLS gegenüber ATM darin, dass es ideal mit IP zusammenarbeiten kann, während dies bei ATM wesentlich komplizierter ist. Durch die Möglichkeit von Label Stacking ist die Adressierung mit MPLS außerdem enorm flexibel, während ATM an die VPI/VCI-Struktur gebunden ist. Weiterhin ist Traffic Engineering bei ATM mit gewissen Einschränkungen verbunden, die bei MPLS nicht existieren [Lak05].

IP Source Routing

Das Label eines Pakets legt seinen Weg durch eine MPLS-Domäne fest. Dies kann genutzt werden, um für bestimmten Datenverkehr feste Pfade vorzugeben.

Theoretisch könnte das gleiche auch durch IP Source Routing erreicht werden. Im Unterschied zu MPLS stellt dies jedoch einen hostgesteuerten Ansatz dar, während diese bei MPLS nicht direkt involviert sind. Die notwendigen technischen und administrativen Rahmenbedingungen sind somit vollkommen unterschiedlich.

IP Source Routing verursacht im Vergleich zu MPLS einen erheblichen Overhead, da die Route in dem IP-Header jedes einzelnen Pakets transportiert werden muss. Weiterhin ist dieser Mechanismus im Vergleich zu MPLS in hohem Maße sicherheitsbedenklich und daher in den meisten Szenarien ungeeignet.

IP Differentiated Services (DiffServ)

Sowohl über IP DiffServ als auch über MPLS CoS können Verkehrsklassen realisiert werden. Beide Verfahren sind wie bereits erwähnt miteinander verwandt und besitzen daher einige Gemeinsamkeiten [SQ]; beispielsweise wird jeweils die Klassifikation und damit die Komplexität in den Rand des Netzes verlagert, und in beiden Fällen wird der Datenverkehr in diskrete Klassen unterteilt.

Ein signifikanter Unterschied besteht darin, dass IP DiffServ ein auf die Vermittlungsschicht beschränktes Konzept darstellt. Demgegenüber sind für MPLS Möglichkeiten spezifiziert, wie QoS-Funktionalitäten niedrigerer Schichten einbezogen werden können [Ste98].

Bei IP DiffServ agiert grundsätzlich jeder Router innerhalb der DiffServ-Domäne unabhängig von den anderen, d.h. das PHB ist jeweils von der lokalen Konfiguration des Routers abhängig. Im Zusammenhang mit MPLS-TE sind jedoch auch QoS-Konzepte denkbar, die nicht auf PHB-Methoden (wie z.B. WFQ oder WRED), sondern auf einer klassenspezifischen Wegewahl basieren; die Klassen wären dann für alle LSRs außer denen am Rand der MPLS-Domäne unsichtbar. Vergleichbare Konzepte wäre mit IP DiffServ aufgrund der fehlenden TE-Fähigkeiten i.A. wesentlich umständlicher zu realisieren, so dass MPLS CoS als mächtigeres Verfahren eingestuft werden kann.

MPLS kann leicht zusammen mit IP DiffServ eingesetzt werden, indem die Verkehrklassen an dem Übergang zwischen einer DiffServ- und einer MPLS-Domäne aufeinander abgebildet werden. Sinnvoll ist dies z.B. dann, wenn ein DSCP-markiertes IP-Paket durch eine MPLS-Domäne transportiert werden soll und die Charakteristiken der DiffServ-Verkehrsklassen erhalten bleiben sollen.

2.1.8. Abgrenzung zu Generalized MPLS (GMPLS)

GMPLS steht als Oberbegriff für eine Reihe von Verfahren, welche die Grundideen von MPLS auf niedrigere Schichten des OSI-Modells ausweiten. GMPLS verallgemeinert dabei das Konzept der Labels von numerischen Werten hin zu anderen Informationsarten, wie z.B. Wellenlängen bei WDM, Zeitschlitzen bei TDM oder Schnittstellen bei SDM. Die zu GMPLS gehörenden Protokolle konzentrieren sich auf Aufgaben der Steuerungsebene wie z.B. die Signalisierung von LSPs auf diesen Schichten. GMPLS ist u.a. in RFC 3945 standardisiert.

GMPLS ist nicht nur eine wesentlich umfassendere Technologie als MPLS, die entsprechende Hardware benötigt, sondern wird auch aus fundamental anderen Gründen eingesetzt. Im Kontext dieser Arbeit wird GMPLS nicht von Bedeutung sein, weshalb hier keine genauere Betrachtung erfolgt.

2.2. Cisco Internetworking Operating System (IOS)

Für seine Hardware hat der Hersteller Cisco ein eigenes einheitliches Betriebssystem entwickelt, das Cisco IOS. Nahezu alle Geräte von Cisco können über IOS betrieben werden; derzeit verwendet nur eine relativ kleine Anzahl von Produktserien andere Betriebssysteme wie das Cisco Catalyst OS (CatOS), Cisco IOS XR oder Cisco NX-OS.

Es existieren Hunderte verschiedener Versionen von IOS. Ursächlich dafür ist zum einen, dass Cisco das Betriebssystem bisher in zahlreichen verzweigten Versionslinien fortgeführt hat, die jeweils an eine bestimmte Kundengruppe angepasst sind; zum anderen basiert das Lizenzierungskonzept von Cisco teilweise auf Modifikationen im Funktionsumfang von IOS. Die

Versionsbezeichnungen von IOS sind entsprechend oft sehr länglich, und die Eigenschaften verschiedener Abkömmlinge derselben Hauptversionsnummer können sich stark unterscheiden.

Das IOS eines Gerätes ist i.A. nicht fest installiert, sondern befindet sich als Datei auf einem auswechselbaren Speichermedium (z.B. einer Flash-Speicherkarte) und kann damit bei Bedarf durch eine andere bzw. neuere Version ersetzt werden.

2.2.1. IOS Command Line Interface (CLI)

Alle Versionen von IOS haben eine einheitliche Benutzerschnittstelle gemeinsam. Beim IOS CLI handelt es sich um eine textuelle Kommandozeile, über das die Steuerung und die Konfiguration der Hardware vorgenommen werden können. Der Zugriff auf das CLI kann über mehrere Wege erfolgen, z.B. über

- eine serielle Schnittstelle (RS-232) mittels eines Terminals bzw. Terminalemulators (physischer Zugang zum Gerät erforderlich)
- telefonische Einwahl mittels eines Modems (out-of-band)
- einen SSH-Zugang (in-band)
- eine HTTP-basierte Web-Schnittstelle (muss zuvor explizit auf dem Gerät aktiviert werden)

Das CLI wird durch eine Eingabeaufforderung (Prompt) dargestellt. Ein Befehl in IOS ist eine Folge von alphanumerischen Wörtern, die ggf. durch Leerzeichen voneinander getrennt werden.

Grundsätzlich kann zwischen zwei Betriebsmodi des CLI unterschieden werden: dem unprivilegierten, und dem privilegierten Modus. Nur in Letzterem sind Befehle zulässig, die die Konfiguration eines Gerätes verändern. Der privilegierte Modus ist daher üblicherweise passwortgeschützt. Der unprivilegierte Modus erlaubt lediglich einfache Aktionen wie z.B. das Anzeigen bestimmter Statusinformationen oder das Anpingen eines Hosts.

Die Konfiguration eines Gerätes wird intern in seinem nichtflüchtigen Speicher (NVM) abgelegt. Die Konfigurationsdatei ist als Klartext lesbar und kann über TFTP exportiert werden (z.B. zu Backup-Zwecken). Da sie neben Details der Netzkonfiguration auch Passwörter im Klartext enthalten kann, ist sie als sicherheitskritisch einzustufen.

2.2.2. Befehlshierarchie

Charakteristisch für das CLI ist die baumartige Struktur der verschiedenen Systembereiche. Zur Ausführung eines bestimmten Befehls muss zunächst in den entsprechenden Bereich navigiert werden. Die Bedeutung und der Effekt eines Befehls sind von diesem Kontext abhängig.

Um beispielsweise eine Schnittstelle zu konfigurieren, muss zunächst mit dem Befehl `configure` in den Konfigurationsbereich gewechselt werden. Anschließend kann mittels des Befehls `interface` eine Schnittstelle ausgewählt werden. Erst dann können die entsprechenden Aktionen auf der Schnittstelle ausgeführt werden. Durch den Befehl `exit` kann in die

übergeordnete Ebene zurückgegangen werden. Bei der Navigation durch die einzelnen Bereiche zeigt das Prompt, vergleichbar mit der Navigation durch eine Dateiverzeichnisstruktur, jeweils die aktuelle Position in der Hierarchie an.

Neben dem CLI sind auch die Befehle bzw. ihre Parameter selbst i.A. hierarchisch strukturiert. Über ein Hilfesystem können die jeweils gerade zulässigen Befehle bzw. Parameter angezeigt werden.

3. Anforderungsanalyse

In Zukunft soll MPLS im MWN eingesetzt werden. Um ein umfassendes Verständnis der Ausgangslage zu entwickeln und potentielle Anwendungsfälle für MPLS zu identifizieren, ist zunächst eine Anforderungsanalyse erforderlich. Das Ziel dieser Analyse ist es festzustellen, in welchen Bereichen des MWN der Einsatz von MPLS sinnvoll erscheint, welche Aufgaben damit erfüllt werden sollen und welche Rahmenbedingungen dabei jeweils gelten. Formal soll eine Anforderungsspezifikation entstehen.

3.1. Vorgehensweise

Zur Erhebung und Erfassung der benötigten Informationen werden drei Quellen herangezogen:

- Schriftliche Dokumente, die Auskunft über die bestehende Netzstruktur geben (soweit vorhanden)
- Interviews von Mitarbeitern des Netzbetreibers
- Konfigurationsdateien der an den Netzknoten eingesetzten Geräte

Im Folgenden wird zunächst die Motivation des Netzbetreibers für die Einführung von MPLS dargestellt. Anschließend wird die bestehende Infrastruktur des MWN untersucht. Dabei werden die im Kontext der Analyse relevanten Informationen herausgearbeitet. Um die Anwendungsfälle zu ermitteln, werden danach die im Betrieb ersichtlich gewordenen Schwierigkeiten erörtert, bei denen sich die momentan eingesetzten Technologien als unzureichend erweisen. Darüber hinaus werden einige weitere typische Einsatzmöglichkeiten von MPLS betrachtet, die im MWN nützlich sein könnten. Anhand der gesammelten Fakten werden schließlich die Anforderungen formuliert.

Hinsichtlich der Einsatzgebiete für MPLS wurde vom Netzbetreiber vorgegeben, dass "minimalinvasiv" vorgegangen werden soll, d.h. zum einen soll MPLS vorerst lediglich in solchen Situationen eingesetzt werden, in denen sich dadurch unmittelbare Vorteile ergeben, und zum anderen soll dies auf eine Weise geschehen, bei der möglichst wenige Seiteneffekte auf andere Netzbereiche auftreten. Dort, wo MPLS einsetzbar wäre, aber der Umstellungaufwand den Nutzen derzeit nicht rechtfertigen würde, sind die bisherigen Strukturen zu belassen.

3.2. Gründe für die Einführung von MPLS

Das MWN besitzt eine über Jahre gewachsene Infrastruktur. Angesichts enorm gestiegener Nutzerzahlen haben sich die Anforderungen im Bereich von Betrieb und Wartung des Netzes verändert. Einige Aufgaben sind aus heutiger Sicht nicht mehr optimal mit den bisherigen technischen Verfahren erfüllbar; es ist anzunehmen, dass sich ohne geeignete Anpassungsmaßnahmen diese Problematik in Zukunft mit dem fortschreitenden Netzausbau verschärfen

würde. Des weiteren können die vorhandenen Technologien bestimmte Funktionen nicht auf angemessene Weise leisten. Im Wesentlichen wird die Einführung von MPLS durch folgende Themen motiviert:

- An das Netz sind zahlreiche Benutzergruppen angeschlossen, die aus sehr unterschiedlichen betrieblichen Umfeldern stammen können. Entsprechend weichen die Anforderungen dieser Gruppen bzgl. der Netznutzung teilweise stark voneinander ab. Diese Abweichungen spiegeln sich in Gestalt von speziellen Konstrukten innerhalb der Konfiguration des Netzes wider, die bisher weitgehend manuell eingerichtet und gewartet werden müssen. Deren Anzahl hat im Laufe der Zeit stetig zugenommen, so dass ihre Pflege sich inzwischen aufwendig gestaltet. Dies wird durch den Umstand verstärkt, dass immer mehr Netzelemente betroffen sind. Der Netzbetreiber vermutet, dass durch den Einsatz von MPLS dieser Aufwand verringert werden kann.
- Die derzeit im MWN eingesetzten Mechanismen zur Verkehrssteuerung sind abhängig von konventionellen dynamischen Routing-Protokollen, deren Möglichkeiten in dieser Hinsicht sehr eingeschränkt sind. MPLS kann hier voraussichtlich die vorhandenen Instrumente erweitern.
- MPLS bietet die Möglichkeit, mittels einer einheitlichen und standardisierten Technologie eine Reihe verschiedener interessanter Dienste und Funktionen umzusetzen, die im MWN aktuell nicht in dieser Form verfügbar sind.

Von dem Einsatz von MPLS erwartet der Netzbetreiber ein einfacheres und effektiveres Netzmanagement und insgesamt eine Verbesserung der Robustheit des Netzes. Bevor konkret auf die einzelnen Sachverhalte eingegangen werden kann, ist eine genauere Betrachtung des MWN und seiner Umgebung erforderlich.

3.3. Szenario: Das Münchner Wissenschaftsnetz

Das MWN verbindet vor allem die Standorte verschiedener wissenschaftlicher Lehr- und Forschungseinrichtungen im Raum München miteinander. Die wesentliche Aufgabe des MWN ist die Bereitstellung des Zugangs zu verschiedenen internen Netzdiensten, zu überregionalen Wissenschaftsnetzen und zum öffentlichen Internet. Auf nationaler Ebene ist das MWN in das Deutsches Forschungsnetz (DFN/X-WiN) und auf internationaler Ebene an das europäische Forschungsnetz (GEANT) eingebunden. Das LRZ ist neben der Verwaltung, dem Betrieb und der Wartung des MWN auch für die Netzplanung verantwortlich (Operation, Administration & Maintenance, OAM).

Die überwiegende Mehrzahl der angebundenen Standorte sind Gebäude der Technischen Universität München (TUM), der Ludwig-Maximilians-Universität (LMU) und weiterer örtlicher Hochschulen sowie solche des Studentenwerks München bzw. die dazugehörigen Studentenwohnheime. Daneben sind in kleinerer Zahl etliche weitere öffentliche Einrichtungen angeschlossen wie z.B. die Bayerische Staatsbibliothek, das Deutsche Herzzentrum und verschiedene Museen. Das MWN wird darüber hinaus von einigen nichtstaatlichen Einrichtung mitgenutzt, u.a. von einigen Instituten der Max-Planck-Gesellschaft (MPG) und der Fraunhofer-Gesellschaft.

Im aktuellen Bericht des LRZ zum Netzkonzept [LR10] wird eine Versorgung von mehr als 50 Arealen mit über 510 Gebäudegruppen (Unterbezirke) durch das MWN ausgewiesen.

Die Anzahl der Arbeitsplatzrechner und Server liegt bei etwa 80000 und die geschätzte Nutzerzahl bei über 117000.

Die Gewährleistung einer hohen Ausfallsicherheit des Netzes besitzt für den Netzbetreiber oberste Priorität. Dabei ist die redundante Auslegung kritischer Netzbestandteile von zentraler Bedeutung. Hinsichtlich der Wartbarkeit ist aus der Sicht des Netzbetreibers ein hoher Grad an Automatisierung bzw. ein niedriger Konfigurationsaufwand und eine gute Übersichtlichkeit der Netzstrukturen wünschenswert. Dies ist langfristig insbesondere angesichts des fortlaufenden Ausbaus des Netzes wichtig. Um auf wirtschaftliche Weise eine optimale Verfügbarkeit zu erreichen, wird zudem ein möglichst gleichmäßiger Auslastungsgrad des Netzes angestrebt.

3.3.1. Topologie

Die bestehende interne Infrastruktur des MWN baut über seine gesamte Ausdehnung auf dem traditionellen TCP/IP-Modell auf. Auf der Netzzugangsschicht wird flächendeckend Ethernet eingesetzt. Die physische Struktur gliedert sich in zwei Ebenen: das Kernnetz/Backbone, und die Zugangsnetze. Das Backbone wird in unregelmäßigen Zeitabständen nach Bedarf erweitert und ausgebaut. Zur Zeit besteht es aus 9 Netzknoten, an denen Layer 3-Geräte arbeiten. Die Knoten sind untereinander teilvermascht verbunden und versorgen lokale Zugangsnetze. Diese decken jeweils einen Standort bzw. einen Gebäudekomplex ab und werden über Layer 2-Geräte hierarchisch weiter gegliedert.

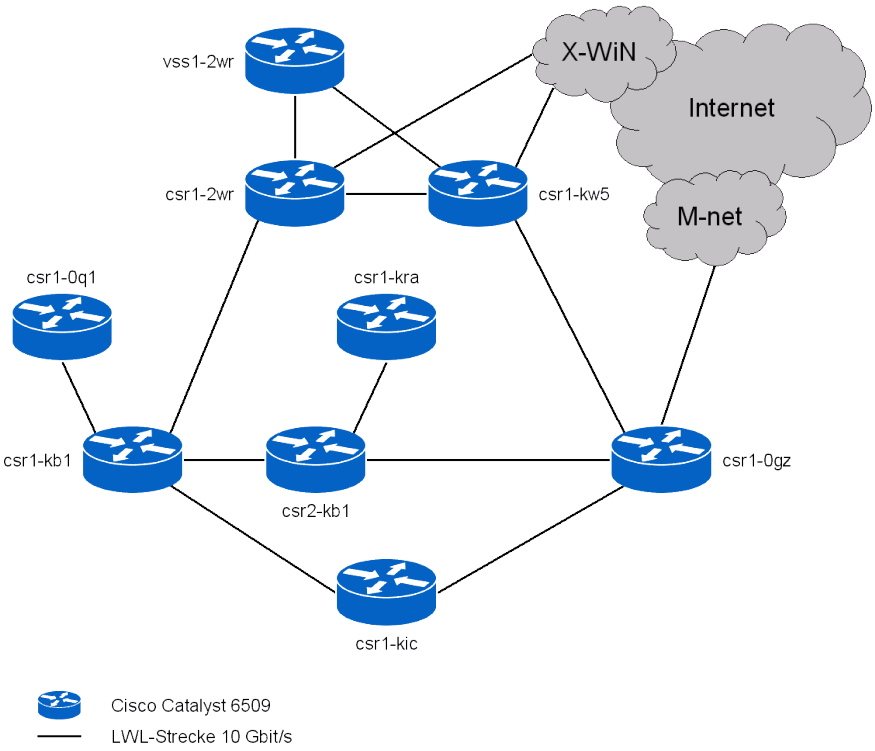

Abbildung 3.1.: Schematische Darstellung des Kernnetzes des MWN

Die Topologie des Backbone (Abbildung 3.1) kann nicht eindeutig einem bestimmten Typ

Router	Standort
csr1-2wr, vss1-2wr	Leibniz-Rechenzentrum
csr1-kw5	Campusgelände der TUM in Garching
csr1-0gz	Stammgelände der LMU
csr1-kb1, csr2-kb1	Stammgelände der TUM
csr1-kic	Campusgelände der LMU in Großhadern/Martinsried
csr1-0q1	Campusgelände in Weihenstephan
csr1-kra	Stammgelände der Fachhochschule München

Tabelle 3.1.: Standorte der wichtigsten Netzknoten

zugeordnet werden. Ein dominantes Merkmal ist der Ring, der durch fünf Router gebildet wird. Aus Gründen der Redundanz bzw. Ausfallsicherheit enthält die Topologie noch weitere Zyklen. Die Knoten sind an strategisch wichtigen Standorten positioniert.

Das MWN besitzt drei Pfade in das öffentliche Internet:

- zwei unabhängige Verbindungen zum X-WiN, die als primärer Internetzugang genutzt werden, und
- eine Backup-Verbindung über den lokalen Netzbetreiber M-net.

Die meisten versorgten Standorte sind direkt an das Kernnetz angeschlossen; einige wenige werden aufgrund ihrer geographischen Entfernung indirekt durch Tunnelkonfigurationen über das X-WiN angebunden.

3.3.2. Hardware

Die Datenübertragungsrate beträgt im Kernnetz durchgängig 10 Gbit/s. An den Knoten des Kernnetzes werden Hochleistungsrouter der Firma Cisco vom Typ Catalyst 6509, sog. Multi-Layer-Switches (MLS), eingesetzt. Sie befinden sich an zentral gelegenen Universitätsgeländen und sind miteinander über gemietete Lichtwellenleiter (LWL), sog. “dark fibre”, verbunden.

Die Geräte sind mit modularen Routing-Prozessoren ausgerüstet (Cisco Supervisor 720), wodurch sich ihre technischen Eigenschaften um zusätzliche Fähigkeiten – vergleichbar mit denen von Cisco-Routern der 7200-Serie – erweitern [Men03]. Alle Router im Kernnetz sind MPLS-fähig. Wesentliche austauschbare Komponenten der Routersysteme (Netzteile, Supervisor-Engines, etc.) sind redundant ausgeführt, wobei i.A. jeweils eine Instanz aktiv ist und die andere sich im Hot-Standby befindet. Bei Versagen einer Komponente findet umgehend ein Stateful-Failover statt, d.h. es gehen im Idealfall keine Pakete verloren und der Betrieb wird praktisch nicht beeinträchtigt. (Eine der wichtigsten Komponenten, die Backplane, ist jedoch nur in einfacher Ausführung vorhanden.) Die aktuell eingesetzte Version von IOS ist *12.2(33)SXI3 Advanced IP Services SSH*.

In den Zugangsnetzen werden Hochleistungsswitches verschiedener Typen der Firma Hewlett Packard eingesetzt (HP ProCurve 4000/5000 Serien). Die Switches befinden sich an den von ihnen versorgten Standorten. Die Gebäude sind zumeist strukturiert mit Twisted-Pair-Kupferleitern verkabelt. Die Anbindung der Zugangsnetze an das Backbone ist sehr unterschiedlich und variiert – abhängig von geographischer Lage und Bedarf des entsprechenden

Standorts – sowohl in der Datenübertragungsrate (von 64 kbit/s bis 10 Gbit/s) als auch in der Übertragungstechnik (ISDN, DSL, Funk, LWL, u.a.).

Die Switches arbeiten als reine Layer 2-Geräte, da darüber hinausgehende Funktionalitäten höherer Schichten deaktiviert wurden. Die Hardware in den Zugangsnetzen wird daher von MPLS nicht betroffen sein.

3.3.3. Netzstruktur und Konfiguration

Das MWN bildet ein eigenes Autonomes System (AS 12816). Als IGP wird im Kernnetz OSPF eingesetzt (in einigen Bereichen auch RIP). Es existiert lediglich eine OSPF-Area (Area 0). Das Routing in das X-WiN erfolgt über BGP. Auf der Vermittlungsschicht wird neben IPv4 auch flächendeckend IPv6 parallel eingesetzt (Dual Stack). Neben dem LRZ selbst besitzen auch einige der versorgten Benutzer eigene öffentliche IP-Adressräume, die entsprechend geroutet werden.

Als Hauptinstrument zur Bildung von logischen Strukturen dienen derzeit Virtuelle Lokale Netze (VLAN) nach IEEE 802.1Q. Generell dienen diese der Aggregation einzelner Hosts zu Benutzergruppen bzw. der Isolation dieser Benutzergruppen und ihres Datenverkehrs voneinander. Die Kommunikation zwischen zwei Hosts in unterschiedlichen VLANs kann nur über eine geroutete Infrastruktur erfolgen.

Im MWN werden VLANs derzeit obligatorisch angewandt, d.h. die an das Backbone angeschlossenen physischen Netze werden normalerweise grundsätzlich als VLANs konfiguriert. Dies gilt auch für die Punkt-zu-Punkt-Verbindungsnetze zwischen einem Paar von benachbarten Routern. Infolgedessen arbeiten praktisch alle physischen Schnittstellen der Router – mit Ausnahme einzelner nach außerhalb des MWN gerichteter Schnittstellen – als reine VLAN-Trunk-Links. Dabei sind jeder physischen Schnittstelle mehrere virtuelle Schnittstellen für die einzelnen VLANs zugewiesen. Jede dieser VLAN-Schnittstellen besitzt eine eigene IP-Adresse, die in dem entsprechenden VLAN eine Gateway-Adresse darstellt.

Einige VLANs sind mittels VLAN-Trunking über das Backbone hinweg zwischen mehreren entfernten Standorten aufgespannt. Sind Rahmen zwischen solchen Standorten zu übertragen, so arbeitet das Backbone in diesem Kontext als ein rein geswitchtes Netz. Erst wenn ein Paket an ein Ziel außerhalb des ursprünglichen VLAN bzw. an ein fremdes IP-Subnetz adressiert ist, wird es geroutet. Die funktionale Rolle der Netzknoten im Backbone ist daher verkehrsabhängig.

Wegen der Berücksichtigung von redundanten Pfaden durch das Backbone können Zyklen von VLAN-Trunks entstehen. Zur Schleifenvermeidung ist daher ein Spanning-Tree-Protokoll eingerichtet (Per-VLAN-Spanning-Tree (PVST) bzw. Rapid-PVST+, Cisco-proprietär). Dieses verwaltet aus logischer Sicht einen eigenen Spanning-Tree je VLAN, wodurch ein einfacher Lastverteilungsmechanismus realisiert wird.

Üblicherweise ist je ein IP-Subnetz mit einem VLAN assoziiert. In einigen Fällen werden im MWN auch mehrere Subnetze innerhalb desselben VLAN betrieben (z.B. im Zshg. mit VoIP-Telefonanlagen). Dann ist der betreffenden VLAN-Schnittstelle des Routers die entsprechende Anzahl von IP-Adressen zugeordnet.

Alle VLANs werden statisch konfiguriert, d.h. es existiert kein VLAN Membership Policy Server (VMPS). Im Regelfall wird eine solche Konfiguration über die Zuweisung von Ports bzw. Portgruppen an den standortlokalen Layer 2-Switches erreicht. Die VLAN-Trunks sind ebenfalls statisch konfiguriert, d.h. die Cisco-Protokolle DTP und VTP sind deaktiviert.

LWL-Strecke	Auslastung →	Auslastung ←
csr1-kb1 ↔ csr1-0q1	104	300
csr1-2wr ↔ csr1-kb1	544	1043
csr1-2wr ↔ csr1-kw5	450	215
csr1-0gz ↔ csr1-kic	17	47
csr1-kb1 ↔ csr1-kic	148	135
csr1-kb1 ↔ csr2-kb1	273	336
csr2-kb1 ↔ csr1-kra	80	53
csr2-kb1 ↔ csr1-0gz	184	69
csr1-2wr ↔ vss1-2wr	2154	1555
csr1-kw5 ↔ vss1-2wr	1299	244
csr1-2wr ↔ X-WiN	832	1003
csr1-kw5 ↔ X-WiN	0	307
csr1-0gz ↔ M-Net	0	0

Tabelle 3.2.: Mittlere Auslastung der Hauptverbindungsstrecken (in Mbit/s, Mittelwerte für den Monat Februar 2010)

3.3.4. Quality of Service

Bisher wird im MWN aus Gründen der Effizienz auf die Anwendung expliziter Mechanismen zur Anforderung und Bereitstellung spezieller Dienstgüte-Merkmale für Datenübertragungen verzichtet. Stattdessen wird der Grundsatz angewendet, möglichst jederzeit ausreichende Reserven an Netzkapazität vorzuhalten (Over-Provisioning) [LR10]. Um hierbei unter wirtschaftlichen Gesichtspunkten sinnvoll planen zu können, wird die Auslastung der Standortanbindungen überwacht und die Messungen ausgewertet. Überschreitet der Datendurchsatz (Mittelwert über 1 Stunde) mehrfach den Schwellenwert von 30% der Gesamtkapazität, so werden Maßnahmen zum Ausbau der Strecke geplant.

Prinzipiell erbringt das Netz den Best-Effort-Dienst von IP. Durch die prozentual niedrige Auslastung der Netzkapazität können dennoch verhältnismäßig gute Werte für QoS-Kriterien wie Latenz, Paketverlustrate, Jitter etc. erzielt werden.

3.3.5. Auslastung und Datenverkehrsflüsse

Alle Router exportieren NetFlow-Daten über die Verkehrsvolumina an einen entsprechenden Server, so dass bis auf die Granularitätsebene von VLAN-Schnittstellen Zeitreihen zur Auslastung zur Verfügung stehen. Tabelle 3.2 zeigt eine Übersicht der Auslastung für die Hauptverbindungsstrecken im Backbone [lrz].

Bei der Betrachtung dieser Werte ist zu beachten, dass es sich um Mittelwerte handelt und die Spitzenwerte um ein Vielfaches höher liegen können. Zum einen ist dies darin begründet, dass die Auslastung tageszeitbedingten Schwankungen unterliegt, und zum anderen darin, dass das typische Kommunikationsverhalten vieler Anwendungen eine diskontinuierliche Charakteristik besitzt, d.h. deren Daten häufig in kurzen “Bursts” übertragen werden, die sich erst ab einer bestimmten Nutzeranzahl statistisch ausgleichen. Zur Beurteilung der relativen Verteilung des Verkehrsaufkommens können die Werte jedoch als Anhaltspunkte dienen.

Eine weitergehende pauschale Untersuchung der Messwerte erscheint an dieser Stelle nicht

zweckmäßig, da die Auswirkungen eines Einsatzes von MPLS stark von dem jeweiligen Anwendungsfall abhängen. In Fällen, in denen MPLS lediglich aus strukturellen Gründen eingesetzt werden sollte (z.B. zur Minderung des Konfigurationsaufwands), würde die Auslastung möglicherweise gar nicht beeinflusst werden. Genauere Betrachtungen dieser Werte könnten ggf. später vor dem Hintergrund einer konkreten Umsetzungsplanung bei der Routenauswahl sinnvoll sein.

3.3.6. Sicherheit, externer Netzzugang

Die vom Netzbetreiber eingesetzten Sicherheitsmaßnahmen konzentrieren sich hauptsächlich darauf, die Funktionsfähigkeit der Netzinfrastruktur sicherzustellen und den Missbrauch des Netzes zu unterbinden. Für den Schutz von Hostsystemen wie Servern und Arbeitsplatzrechnern sind die Benutzer selbst verantwortlich, d.h. besondere Abschirmungsmaßnahmen sind bei Bedarf dezentral umzusetzen.

Der Zugang von außen in das MWN kann über Modemeinwahl, einen bestehenden Internetzugang, öffentliche Ethernet-Netzanschlussdosen oder WLAN hergestellt werden. Mit Ausnahme des ersten Falls ist zusätzlich die Verwendung eines IPsec-VPN-Clients erforderlich, um interne Netzdienste nutzen zu können.

Der Netzbetreiber setzt diverse Monitoring- und Intrusion-Detection-Werkzeuge ein, um Angriffe zu erkennen und zu verhindern. Darüber hinaus sind die Router im Backbone derart konfiguriert, dass IP-Spoofing und damit verbundene Angriffstechniken (v.a. Denial-of-Service (DOS) bzw. Distributed DOS (DDOS)) erschwert werden. Dies wird zum einen durch spezielle Verfahren (z.B. Unicast RPF und CoPP) erreicht, und zum anderen durch eine möglichst lückenfreie und konsistente Konfiguration mit geeigneter Anwendung von Zugriffskontrollisten.

Aus Gründen der Organisation und Effektivität existiert im MWN kein globales Firewall-System. Der Netzbetreiber bietet seinen Kunden virtualisierte Firewall-Dienste als Alternative zu eigenen Systemen an. Für Benutzer, die keine öffentlichen IP-Adressen benötigen, existiert ein NAT-Dienst ("NAT-o-MAT"), so dass an diese Benutzer private IP-Adressen vergeben werden können. Dadurch wird eine Minimalbarriere für externe Angreifer hergestellt.

Der NAT-o-MAT und sein im Aufbau befindlicher Nachfolger ("Secomat") erfüllen auch eine Kontrollfunktion, indem sie den über sie geführten Datenverkehr auf Auffälligkeiten hin untersuchen und verdächtige IP-Adressen ggf. temporär sperren. Aus diesem Grund wird auch der Datenverkehr von Hosts, die über einen der o.g. externen Netzzugänge an das MWN angeschlossen sind, zwangsweise über diese Systeme geleitet.

3.3.7. Accounting

Die bereits genannte Erfassung von Nutzungsstatistiken findet derzeit nur zur Informations- und Planungszwecken statt [LR10]. Eine nutzerbezogene Abrechnung für das Dienstangebot des MWN findet derzeit nicht statt. Dieser Aspekt wird daher im Rahmen der Analyse nicht gesondert berücksichtigt. Es ist jedoch bei der Einführung von MPLS darauf zu achten, dass eine solche Abrechnung technisch realisierbar bleibt, sofern dies zu einem zukünftigen Zeitpunkt gewünscht werden sollte.

3.4. Anwendungsfälle für MPLS

Einige Anwendungsfälle ergeben sich aus den zu Beginn dieses Kapitels angesprochenen Problemstellungen, durch die die Einführung von MPLS motiviert wurde. In den folgenden Unterabschnitten werden die dazugehörigen Sachverhalte präzisiert. Darüber hinaus sind – mit Hinblick auf die in Kapitel 2 dargestellten Fähigkeiten von MPLS – einige weitere mögliche Einsatzbereiche im MWN vorstellbar. Diese werden im Anschluss daran erörtert.

Zur Vermeidung von unnötigem Aufwand ist es sinnvoll, möglichst frühzeitig die Machbarkeit der Anwendungsfälle zu validieren. Obwohl es sich bei MPLS um eine standardisierte Technologie handelt, ist dies eine nichttriviale Aufgabe. Der Grund dafür ist, dass sich Standards i.A. überwiegend auf solche Aspekte konzentrieren, die für die Interoperabilität von Systemen unterschiedlicher Hersteller von Bedeutung sind. Entsprechend liegt der Fokus der IETF-Dokumente vorwiegend auf der Spezifikation grundlegender Funktionsweisen von MPLS wie z.B. der Signalisierung und den damit verbundenen Protokollen.

Höhere Funktionen auf Basis von MPLS und deren technische Realisierung sind aber oft in beträchtlichem Umfang herstellerspezifisch. Darüber hinaus existieren gelegentliche Abweichungen von den Standards und/oder zusätzliche proprietäre Restriktionen. Der Einfluss dieser Faktoren auf die Machbarkeit der Anwendungsfälle hängt unmittelbar von den konkreten Lösungsansätzen ab. In dieser Phase der Arbeit können daher keine abschließenden Urteile dazu gefällt werden. Die jeweils angefügten Bewertungen zur Machbarkeit sind als vorläufige Einschätzung zu betrachten, um Anhaltspunkte für das weitere Vorgehen zu gewinnen.

3.4.1. Standortübergreifende VLANs

Derzeit werden im MWN in hoher Anzahl VLANs eingesetzt (> 1300) [LR10]. Angesichts dieser Größenordnung erscheint ihre Konfiguration recht unübersichtlich, was insbesondere daran liegt, dass die zugrundeliegenden Strukturen historisch gewachsen sind und die Granularitäten der VLANs daher sehr unterschiedlich sind. Die meisten von ihnen dienen der Isolation standortlokaler Netze voneinander, während andere zu speziellen Einsatzzwecken (z.B. für das Netzmanagement) und mit entsprechend abweichenden Konfigurationen eingerichtet wurden.

Zu den Letzteren gehören momentan insgesamt 12 VLANs, die sich über mehrere Standorte hinweg erstrecken. Ihre Aufgabe ist die Bereitstellung einer isolierten, transparenten Kommunikation zwischen diesen Standorten, d.h. die Erbringung eines VPN-Dienstes auf der Sicherungsschicht. Problematisch hierbei ist, dass diese VLANs auf den zwischen den Standorten liegenden VLAN-Trunks im Backbone statisch konfiguriert sind. Dieses Verfahren ist unflexibel, da es impliziert, dass bei Veränderungen der Topologie oder anderer struktureller Eigenschaften des Netzes die Konfiguration ggf. an mehreren Netzknoten im Backbone manuell angepasst (oder zumindest überprüft) werden muss.

Der Aufwand und die Komplexität werden weiter dadurch erhöht, dass zur Gewährleistung eines Mindestmaßes an Ausfallsicherheit zwischen jedem Paar von Standorten wenigstens zwei redundante Pfade berücksichtigt werden müssen. Dabei können Zyklen in der Layer 2-Topologie entstehen, die über Rapid-PVST+ aufgelöst werden. Dieses Protokoll arbeitet vollständig automatisch und kann nicht auf einfache Weise kontrolliert bzw. restringiert werden, was aus Sicht der Netzplanung ungünstig ist. Eine ungleichmäßige Auslastung kann die Folge sein; z.B. könnten sich zwei Spanning-Trees auf nachteilige Weise überlagern

und punktuell ein hohes Verkehrsaufkommen verursachen, während gleichzeitig ein niedrig ausgelasteter Alternativpfad mit den gleichen Kosten existiert. Darüber hinaus können grundsätzlich die entstehenden Kommunikationswege deutlich von dem kürzesten Weg, wie er beim SPF genommen werden würde, abweichen.[1]

Die Problematik kann eingegrenzt werden, indem nicht alle möglichen redundanten Pfade zwischen jedem Paar von entfernten Standorten als VLAN-Trunks konfiguriert werden, sondern nur eine Teilmenge von ihnen. Dies ist im MWN auch der Fall, hat jedoch den entscheidenden Nachteil, dass somit nicht mehr das volle Redundanzpotential der Backbone-Topologie ausgeschöpft wird.

Unabhängig von den bisherigen Betrachtungen ist außerdem anzunehmen, dass die über das Backbone hinweg aufgespannten VLANs eine erhöhte Ressourcenbelastung des Netzes darstellen, u.a. da je VLAN ein separater Spanning-Tree berechnet und verwaltet wird, und weil alle beteiligten Netzknoten jeweils in ihrer Funktion als Layer 2-Switches ständig die MAC-Adressen sämtlicher im betreffenden VLAN befindlichen Netzschnittstellen lernen.

Insgesamt ist die OAM der o.g. VLANs nicht nur verhältnismäßig umständlich und aufwendig, sondern auch fehleranfällig. Die Skalierbarkeit des bisherigen Konzepts ist infolgedessen eingeschränkt (nicht zuletzt auch dadurch, dass der Adressraum für VLANs relativ klein ist). Mit dem Einsatz von MPLS soll eine Verbesserung der Situation erreicht werden.

Die auf den ersten Blick naheliegende Idee, VTP zur Konfiguration der VLAN-Trunks einzusetzen, erscheint aus mehreren Gründen nicht sinnvoll:

- Jeder von VTP verwaltete VLAN-Trunk transportiert normalerweise alle innerhalb einer VTP-Domäne konfigurierten VLANs. Im Fall des MWN-Backbone wäre dies aus administrativer und ressourcenorientierter Sicht gänzlich ungeeignet, da sich nur ein kleiner Anteil aller VLANs über mehrere Standorte erstreckt. Mittels VTP Pruning kann nur teilweise Abhilfe geschaffen werden; die Effizienz von VTP Pruning hängt außerdem von der konkreten Form des Spanning-Tree ab. Bei Umsetzung einer Mischvariante in dem Sinne, nur einige gesonderte VLANs von VTP konfigurieren zu lassen, wäre das Resultat nahezu identisch mit der bestehenden Konfiguration, so dass damit kein Gewinn zu erwarten ist.

- Die bestehende VLAN-Struktur im Backbone kann mit VTP nicht auf einfache Weise nachgebildet werden. Bei einer komplexen Konstruktion mehrerer VTP-Domänen würde effektiv der Nutzen von VTP zunichte gemacht.

- Die Verwendung von VTP birgt darüber hinaus einige grundsätzliche Risiken. VTP kann zu großen Spanning-Trees und den damit verbundenen Schwachpunkten führen. Darüber hinaus kann die Integrität der VLAN-Datenbank des VTP-Servers beschädigt oder zerstört werden, wenn Fehler bei einer Erweiterung der Netztopologie gemacht werden. Die Folge wäre im worst-case ein Netzausfall der gesamten VTP-Domäne, und eine Neukonfiguration würde notwendig.

[1] Durch die Vergabe von geeigneten Werten für die Bridge-Prioritäten könnten diese Schwierigkeiten zwar theoretisch behoben werden, dies käme jedoch einer manuellen Konfiguration der Spanning-Trees gleich.

Beispiel

Ein Beispiel für ein über das Backbone aufgespanntes VLAN ist das Verwaltungsnetz der TUM. Hier dient das MWN als reines Transportnetz zwischen einer eigenen physischen abgeschirmten Infrastruktur und entfernten Standorten der TUM-Verwaltung. Datenpakete aus diesem VLAN werden im MWN nicht geroutet.

Machbarkeit

MPLS-basierte VPNs sind ein natürlicher Lösungsansatz zur Ersetzung standortübergreifender VLANs. Grundsätzlich könnte dabei das dynamische Routing einbezogen werden, so dass der Konfigurationsaufwand vsl. beträchtlich reduzierbar wäre. Das Redundanzpotential der Topologie könnte auf diese Weise vollständig ausgenutzt werden. Als erfreulicher Nebeneffekt entfiele hierbei der Einsatz eines Spanning-Tree-Protokolls. Alle Cisco-Geräte im Backbone unterstützen sowohl Layer 2- als auch Layer 3-VPNs.

3.4.2. Policy-basiertes Routing

In einigen Situationen im MWN ist der Einsatz von Policy-basiertem Routing erforderlich. Die Konfiguration der dafür notwendigen Regeln muss manuell erstellt und gepflegt werden und gestaltet sich daher aufwendig.

Die Beweggründe für Policies können sehr unterschiedlich sein. Auf das Abweichen von den Vorgaben dynamischer Routing-Protokolle lässt sich bei dem Vorhandensein von externen Rahmenbedingungen daher i.A. nicht verzichten. Grundsätzlich ist diese Form des Routings aber inhärent mit einem gewissen Grad an manueller Konfiguration verbunden, da Policies oft nichttechnischer Natur sind, individuelle Ursachen haben und sich daher meist nicht sinnvoll in automatisierter Form abbilden lassen. Eine Reduzierung des Aufwands, soweit das möglich sein sollte, wäre wünschenswert.

Beispiel

Ein Beispiel für Policy-basiertes Routing besteht im Zusammenhang mit der Weiterleitung von Datenverkehr aus den am MWN angeschlossenen Instituten der MPG. Die MPG besitzt einen eigenen Zugang in das öffentliche Internet am Max-Planck-Institut für Plasmaphysik (IPP) in Garching. Aus administrativen Gründen soll der Datenverkehr von allen MPG-Instituten in das Internet über diesen Zugang geführt werden, und nicht über die MWN-eigene Verbindung zum X-WiN. Daher ist das Backbone derart konfiguriert, dass die Default-Route für Daten aus den MPG-Instituten über das IPP läuft, wo das weitere Routing in das Internet vorgenommen wird.

Eine zentrale Rolle bei Policy-basiertem Routing nimmt der Router `csr1-2wr` ein, der sich direkt am LRZ befindet. Über diesen werden etliche solche Routen geführt, unter anderem im Zusammenhang mit dem NAT-o-MAT-Dienst, dem VPN-Dienst für den Netzzugang von außerhalb des MWN und die soeben beschriebene Route für den Datenverkehr aus den MPG-Instituten.

Machbarkeit

Pauschal kann keine Aussage darüber getroffen werden, ob MPLS-basierte Verfahren hier Betriebsvorteile bringen können. Man kann aber zumindest festhalten, dass Policy-basiertes

Routing innerhalb eines AS meist schlicht dadurch umgesetzt wird, dass feste Pfade für den betroffenen Datenverkehr vorgegeben werden, die sich von dem Shortest-Path unterscheiden. Dieses Ziel könnte theoretisch ebenso mit TE-Tunneln erreicht werden. Ein Unterschied ergäbe sich daraus, dass statische Routen Hop-by-Hop zu konfigurieren sind (d.h. potentiell auf mehreren Routern entlang der gewünschten Route), während ein TE-Tunnel bei Nutzung von dynamischem Routing nur die entsprechende Konfiguration von Start- und Endpunkt erfordert. Möglicherweise könnten sogar mehrere Policies mit einem einzelnen TE-Tunnel aggregiert werden.

Daneben sind auch noch andere Lösungsansätze für die Problematik denkbar. Inwieweit diese praktikabel sind und ob sich die Umstellung lohnt, muss aufgrund der Verschiedenartigkeit von Policies jedoch von Fall zu Fall einzeln untersucht werden.

3.4.3. Bandbreitenbegrenzung

Viele der an das MWN angeschlossenen Zugangsnetze werden direkt vom LRZ oder in enger Kooperation mit diesem betrieben. Einige Nutzer verwalten ihre Netze jedoch weitgehend selbständig, so dass diese außerhalb der administrativen Hoheit des LRZ liegen. Entsprechend kann kaum Einfluss auf die dortigen Betriebsrichtlinien, Netzstrukturen und Konfigurationen genommen werden (evtl. sind diese Informationen noch nicht einmal bekannt). Für den Netzbetreiber ist es daher von vitalem Interesse, durch gesonderte Maßnahmen sicherzustellen, dass potentielle Störungen in diesen Zugangsnetzen bzw. deren Auswirkungen die Betriebssicherheit des MWN nicht gefährden können.

In diesem Zusammenhang soll untersucht werden, ob MPLS die technischen Möglichkeiten zur Begrenzung der von einem Benutzer verwendeten Bandbreite verbessern oder erweitern kann. Eine solche Begrenzung kann verhindern, dass das MWN-Backbone durch fehlerbedingtes oder Malware-verursachtes Datenaufkommen überlastet wird (z.B. bei Routingschleifen, Spanning-Tree-Fehlern oder Problemen mit Broadcast-Traffic).

Als weitere mögliche Gründe für die Begrenzung der Bandbreite eines Übertragungskanals im MWN sind drei Motive ersichtlich:

- Es soll allgemein verhindert werden, dass der Datenverkehr einzelner Benutzer zuviele Ressourcen in Anspruch nimmt und der übrige Verkehr dadurch verhungert (Fairnessprinzip).
- Transitverkehr kann unmittelbar zusätzliche finanzielle Kosten verursachen (z.B. in das öffentliche Internet). Für bestimmte Nutzergruppen soll daher das übertragbare Datenvolumen beschränkt werden.
- Es wurde ein SLA vereinbart, das ungeachtet sonstiger Umstände durchgesetzt werden soll.

Beispiel

Derzeit wird in Einzelfällen von Traffic Policing Gebrauch gemacht. Ein Beispiel hierfür ist das Zugangsnetz des Max-Planck-Instituts für psychologische Forschung, dessen mittlerer Datendurchsatz in das MWN hinein auf 100 Mbit/s beschränkt ist.

Machbarkeit

MPLS verfügt über keinerlei Eigenschaften, die explizit zur Bandbreitenbegrenzung genutzt werden könnten. Zwischen MPLS und den gebräuchlichen Verfahren zur Bandbreitenbegrenzung, Traffic Policing und Traffic Shaping, sind keine unmittelbaren funktionalen Abhängigkeiten erkennbar. Im weiteren Verlauf der Arbeit ist daher hauptsächlich zu klären, ob Bandbreitenbegrenzungen indirekt durch MPLS unterstützt werden können und auf welche Weise sie im Kontext von MPLS im MWN einsetzbar sind.

Traffic Shaping ist als Mittel zur Bandbreitenbegrenzung für diesen Anwendungsfall nicht relevant, da es aus anderen Intentionen als den o.g. Motiven eingesetzt wird; das wesentliche Merkmal des Verfahrens, das Bewirken einer konstanten Senderate bei der Paketweiterleitung, spielt hier keine Rolle. Darüber hinaus ist unklar, ob Traffic Shaping von der Hardware unterstützt wird. Der Fokus weiterer Betrachtungen richtet sich daher auf Traffic Policing.

3.4.4. Verkehrsklassen

Bisher ist im MWN kein einheitliches und flächendeckendes Verfahren zur Priorisierung von Datenverkehr etabliert. Es können jedoch durchaus – sowohl aktuell existierende als auch potentielle – Verkehrsbeziehungen identifiziert werden, deren Anforderungen höher (oder auch niedriger) als die der Mehrheit des Datenverkehrs liegen.

An den Statistiken zur Netzauslastung [lrz] kann man erkennen, dass viele Übertragungsstrecken im Backbone auch zu den Hauptzeiten nur eine schwache bis mittlere Auslastung aufweisen. Dies spiegelt das im MWN praktizierte Konzept des Over-Provisioning wider. Dieses Vorgehen ist zwar einfach und bis zu einem gewissen Grad ökonomisch günstig, aber unspezifisch und daher vergleichsweise ineffizient. Darüber hinaus ist die zukünftige Skalierbarkeit fragwürdig.

Die gezielte Unterstützung von QoS-sensiblen Anwendungen durch Priorisierung wird momentan vom Netzbetreiber nicht angestrebt, könnte aber in Zukunft interessant sein. Hierdurch entstünden zwei wesentliche Vorteile:

- Zum einen ließe sich die Netzkapazität selektiver und damit ökonomischer ausnutzen. Eine zukünftige Erweiterung der Kapazität könnte dadurch evtl. verzögert und die gewünschten QoS-Ziele mit einem geringeren wirtschaftlichen Einsatz erreicht werden.
- Zum anderen könnten die QoS-Ziele durch dafür ausgelegte technische Mechanismen effektiver (z.B. bei erhöhter Netzauslastung) durchgesetzt werden.

Aus diesem Grund soll untersucht werden, wie MPLS hierfür sinnvoll genutzt werden kann.

Machbarkeit

Die QoS-Architektur von MPLS CoS basiert, wie in Kapitel 2 erläutert wurde, auf IP DiffServ. Sie definiert dabei keine zusätzlichen oder neuartigen Mechanismen. Das technische Potential von MPLS zur Umsetzung von Verkehrsklassen ist daher größtenteils identisch zu dem, was im MWN auch ohne MPLS schon jetzt machbar ist.

Die Frage, ob MPLS CoS bestimmte Vorteile gegenüber den bestehenden Möglichkeiten aufweist, ist allerdings irreführend. MPLS CoS ist nicht als Argument für den Einsatz von MPLS geeignet. Vielmehr ergeben sich Nutzen und Zweck von MPLS CoS aus der Motivation, Verkehrsklassen in einem Netz realisieren zu wollen, in dem bereits – aus anderen

Gründen – MPLS eingesetzt wird. Im weiteren Verlauf der Arbeit werden daher vorwiegend die Eigenschaften von MPLS CoS mit Hinblick auf einen eventuellen zukünftigen Einsatz betrachtet.

3.4.5. Lastverteilung

Aufgrund der Tatsache, dass sich damit der Verlauf von Datenübertragungspfaden auf sehr einfache Weise gezielt steuern lässt, erscheint MPLS geradezu prädestiniert als Werkzeug für das Traffic Engineering. Eine wesentliche Motivation für TE ergibt sich aus der Überlegung, dass nicht nur jedes einzelne Datenpaket für sich betrachtet effizient durch das Netz transportiert werden soll, sondern auch die Netzressourcen global möglichst effektiv genutzt werden sollen.

Eine transportbezogene Lastverteilung, d.h. das Weiterleiten von Datenverkehr über potentiell suboptimale Pfade, um die Auslastungsverhältnisse im Netz zu verbessern, wird auf den Hauptverbindungsstrecken des Backbone derzeit nicht explizit angewandt. Bisher besteht dafür kein dringender Bedarf, da sich aufgrund der bereits erwähnten Praxis des Over-Provisioning die mittlere Auslastung i.A. auf einem relativ geringen Niveau bewegt.

Es gilt jedoch als nahezu sicher, dass das Verkehrsaufkommen zukünftig weiter zunimmt. Da die Nutzerschaft des MWN örtlich sehr ungleichmäßig verteilt ist und man angesichts ihrer Diversität von einem inhomogenen Nutzungsverhalten ausgehen kann, ist zu erwarten, dass sich dabei Hot-Spots entwickeln werden, an denen die Auslastung punktuell besonders hoch ist; dies zeichnet sich bereits heute ab, z.B. an der Übertragungsstrecke zwischen `csr1-kb1` und `csr1-2wr`.

Der gezielte Einsatz von Maßnahmen zur Lastverteilung kann die Kosteneffektivität der Infrastruktur verbessern und präventiv gegen überlastbedingte Netzinstabilitäten wirken. Durch eine effizientere Nutzung der vorhandener Ressourcen könnten somit neue Ausgaben für den Netzausbau evtl. verzögert oder sogar vermieden werden. Da – mit Ausnahme von `csr1-0q1` und `csr1-kra` – zwischen jedem Paar von Routern im Backbone mindestens zwei disjunkte Wege existieren, bietet sich eine Lastverteilung im MWN auch intuitiv an.

Statisches Routing ist aufgrund seiner mangelnder Flexibilität zur Lastverteilung eher ungeeignet. Die technischen Instrumente zur Lastverteilung beschränken sich im MWN daher derzeit auf ECMP innerhalb des IGP. Grundlage für die Auswahl gleichwertiger Routen ist dabei eine konventionelle Routing-Metrik, bei der die Kosten eines Pfades sich additiv aus denen der einzelnen Übertragungsstrecken ergeben. Die Möglichkeiten zur Verkehrssteuerung sind hier recht limitiert:

- Es werden nur Wege mit identischen Kosten berücksichtigt. Andere denkbare Optionen wie z.B. eine Aufteilung zwischen Wegen mit unterschiedlichen Kosten oder eine gewichtete Aufteilung sind auf diese Weise nicht möglich. Zwar könnte über eine manuelle Anpassung der Kosten einzelner Übertragungsstrecken Einfluss genommen werden, jedoch wäre ein derartiges Konzept schwer zu balancieren, empfindlich für topologische Veränderungen, extrem wartungsintensiv und daher praktisch kaum brauchbar.

- In die Berechnung fließen keine dynamischen Parameter ein. Eine Anpassung von Routing-Tabellen findet nur bei diskreten Veränderungen der Topologie statt (z.B. dem Ausfall einer Übertragungsstrecke), nicht jedoch bei graduellen Veränderungen (z.B. stetig zunehmende Auslastung einer Übertragungsstrecke).

- Da IP ein verbindungs- und zustandsloses Protokoll ist, können die Ressourcenanforderungen von Datenströmen bei der Wegewahl nicht ohne weiteres vorab berücksichtigt werden, um eine gleichmäßige Auslastung zu erzielen.

Aufgrund dieser Einschränkungen soll untersucht werden, welche Möglichkeiten zur Lastverteilung sich mit MPLS ergeben.

Machbarkeit

MPLS-TE kann das IGP um einige zusätzliche Optionen zur Lastverteilung erweitern. Alle diese Optionen haben gemeinsam, dass sie jeweils für eine gegebene Menge von TE-LSPs deren Auslastung untereinander balancieren, d.h. der Einsatz von LSPs ist insofern obligatorisch, als dass konventioneller IP-Datenverkehr außerhalb dieser LSPs nicht in die Lastverteilung eingezogen werden kann. Diese Bedingung ist aber unproblematisch, da seitens der Netzplanung im MWN ohne weiteres LSPs errichtet werden könnten, die ausschließlich dem Zweck der Lastverteilung dienen.

Eine Lastverteilung könnte mit MPLS auf bedeutend flexiblere Weise realisiert werden als mit der o.g. konventionellen Methode. Der Hauptvorteil bestünde darin, dass die dafür zu verwendenden Pfade unabhängig von deren Länge, Lage und IGP-Kosten gewählt werden können. Diese Faktoren haben keinen zwingenden Einfluss darauf, ob Datenverkehr über den entsprechenden LSP geführt wird oder nicht, da hierfür die TE-Metrik maßgeblich ist, die sich unabhängig davon variieren lässt. Die Planung und Gestaltung der zu verwendenden Pfade würde sich dadurch erheblich erleichtern.

3.4.6. Überbrückung von Ausfällen

Um ein möglichst robustes Netz zu erhalten, wurden im MWN auf unterschiedlichen technischen Ebenen Redundanzen gebildet. Kommt es zu einer Störung, so kann durch Ausnutzung dieser Redundanzen die Funktionsfähigkeit des Netzes erhalten werden. Bei Totalausfällen einzelner Übertragungsstrecken oder Netzknoten im Backbone greifen im Wesentlichen die im IGP dafür vorgesehenen Mechanismen, d.h. es wird eine Neuberechnung der Routing-Tabellen initiiert. Die damit erreichbaren Konvergenzzeiten liegen im Sekundenbereich, was eine spürbare Beeinträchtigung des Netzbetriebs darstellen kann. Die zeitliche Größenordnung ist dadurch bedingt, dass Topologieinformationen zwischen Netzknoten ausgetauscht werden müssen und das Verfahren verteilt arbeitet.

Eine weitere kritische Größe ist die notwendige Zeitdauer bis zur Erkennung eines Ausfalls. Diese hängt zumeist von der Art der Störung ab und kann erheblich über der reinen Konvergenzzeit des IGP liegen, da die dazugehörigen Timeout-Werte i.A. mit einer hohen Toleranz gewählt werden, um die Netzstabilität nicht zu gefährden. Die Netzknoten des Backbone benutzen für die Timeout-Werte hauptsächlich Standardeinstellungen des Hardware-Produzenten, die recht konservativ bemessen sind.

Derzeit sind im MWN keine technischen Mittel im Einsatz, durch die diese Zeiträume verringert oder überbrückt werden könnten. Viele Echtzeitanwendungen reagieren jedoch empfindlich auf Unterbrechungen; zu dieser Klasse von Anwendungen zählt auch z.B. VoIP-Telefonie, die im MWN extensiv eingesetzt wird.

MPLS bietet einige Möglichkeiten an, um die Unterbrechungszeiten bei einem Ausfall erheblich zu verringern. Es ist daher zu untersuchen, ob durch diese im MWN eine Verbesserung der Ausfallsicherheit erreicht werden kann.

Machbarkeit

Im Umfeld von MPLS existieren zwei Verfahren, die der Verkürzung von ausfallbedingten Unterbrechungen dienen: MPLS-FRR, und MPLS Path Protection. Damit diese Verfahren effektiv eingesetzt werden können, ist eine schnelle Ausfallerkennung notwendig. Da im MWN auf der Sicherungsschicht Ethernet eingesetzt wird und diese Technologie keine eigenen Mechanismen dafür bereitstellt, bietet sich der Einsatz des in Kapitel 2 genannten BFD an. Alle diese Verfahren werden von der vorhandenen Hardware unterstützt.

3.5. Weitere Stakeholder im Szenario

Bis zu dieser Stelle der Analyse wurden alle Themen aus der Sicht des Netzbetreibers betrachtet. Da sich die Einführung bzw. der Einsatz von MPLS in erster Linie auf die Arbeit und das Umfeld des Netzbetreibers auswirken, ist dies ein logischer Ansatz. Es können jedoch noch weitere Interessengruppen identifiziert werden, die indirekt betroffen sein werden.

3.5.1. Nutzer des MWN

Für die Endgeräte und lokalen Zugangsnetze im MWN wäre MPLS transparent, d.h. die Nutzer des Netzes würden durch dessen Einsatz vsl. keine unmittelbaren Veränderungen bemerken. Von der Realisierung der o.g. Anwendungsfälle würden sie auf eher subtile Weise profitieren:

- Bei durch Priorisierung unterstützten Anwendungen könnte für deren Nutzer eine positive Veränderung verschiedener QoS-Parameter spürbar werden.
- Bei einer hohen Netzauslastung könnte sich der Einsatz von Lastverteilung günstig auf die Verfügbarkeit des MWN auswirken. Der Effekt einer verbesserten Ausfallsicherheit ginge ebenfalls in diese Richtung.

Die bisherige Umsetzung von VPNs mittels standortübergreifender VLANs schottet diese Netze vollständig ab, d.h. deren Datenpakete werden nicht geroutet. Für manche Benutzer im MWN könnten jedoch auch VPNs mit Internetzugang interessant sein. In der Planungsphase sollte dieser Aspekt daher berücksichtigt werden.

3.5.2. Betriebliches Management des Netzbetreibers

Aus Sicht der Geschäftsleitung sind vor allem die wirtschaftlichen Aspekte einer Einführung von MPLS von Bedeutung. In Bezug auf die technische Ausstattung wäre sie aus jetziger Sicht kostenneutral, da die erforderliche Hardware bereits vorhanden ist und momentan keine weiteren Anschaffungen notwendig sind.

Ein zusätzlicher finanzieller Aufwand entstünde einmalig durch den Personaleinsatz zur Vorbereitung und Durchführung der Migration. Es ist nicht anzunehmen, dass der laufende Betrieb von MPLS langfristig erhöhte Kosten zur Folge hat. In Zukunft könnten durch gesteigerte Effizienz des Netzmanagements und der Ressourcenausnutzung vsl. eher Kosten eingespart werden.

3.6. Anforderungsspezifikation

Im folgenden Abschnitt werden die Anforderungen an die Einführung von MPLS im MWN zusammengefasst.

3.6.1. Übersicht und Ziele

Der wesentliche Einsatzzweck von MPLS ist die Ergänzung der bereits vorhandenen Technologien, um verschiedene mit diesen verbundene Beschränkungen zu überwinden und neue Funktionalitäten im Netz zu integrieren. Folgende Ziele werden dabei angestrebt:

- Unterstützung und Erleichterung des Netzmanagements
- Verbesserung der Robustheit und Effizienz des Netzes
- Erweiterung des Dienstangebots und der technischen Möglichkeiten zur Verkehrssteuerung und -kontrolle

3.6.2. Voraussetzungen, Restriktionen und Abhängigkeiten

- Die Anwendungsmöglichkeiten von MPLS sind im Rahmen dieser Arbeit auf den technischen Funktionsumfang der vorhandenen Geräte beschränkt. Eine unmittelbare Anschaffung weiterer Hardware ist nicht vorgesehen.
- MPLS wird im MWN vorerst isoliert eingesetzt werden, d.h. eine Interaktion mit benachbarten Netzen bzw. anderen Autonomen Systemen ist derzeit nicht geplant. Der Netzbetreiber besitzt daher bzgl. MPLS die vollständige administrative Entscheidungsfreiheit.
- Inwieweit die im weiteren Verlauf der Arbeit betrachteten Anwendungsfälle tatsächlich realisierbar sind, hängt sowohl von der technischen Machbarkeit auf den gegebenen Plattformen als auch von der Komplexität möglicher Implementierungen ab. Diese Fragen sind sehr lösungsspezifisch und werden daher erst in Kapitel 4 abschließend geklärt.

3.6.3. Operatives Umfeld

Allgemeiner Einsatzbereich für MPLS ist das gesamte Backbone des MWN. Als mögliche aktive MPLS-Komponenten sind 9 Routersysteme verfügbar, die bereits mit der erforderlichen Hardware ausgerüstet sind. Die Konfiguration und Überwachung von MPLS erfolgt auf diesen Routern über das Cisco IOS CLI. In direkter Interaktion mit MPLS werden vsl. nur Mitarbeiter des Netzbetreibers stehen.

3.6.4. Funktionale Anforderungen

Folgende Anwendungsfälle für MPLS wurden im MWN gefunden und werden in die Planungsphase einbezogen:

1. Ersetzung standortübergreifender VLANs durch auf MPLS-basierende VPNs

2. Vereinfachung oder Reduzierung von Policy-basiertem Routing
3. Einsatz von Traffic Policing im Kontext von MPLS
4. Bildung von Verkehrsklassen über MPLS CoS
5. Lastverteilung zwischen verschiedenen Übertragungsstrecken im Backbone mit MPLS-TE
6. Verbesserung der Ausfallsicherheit durch MPLS-FRR oder MPLS Path Protection

3.6.5. Nichtfunktionale Anforderungen

Folgende Rahmenbedingungen wurden ermittelt:

1. Die Erstellung und Pflege der Konfiguration von MPLS soll möglichst einfach sein, ein hoher Automatisierungsgrad ist wünschenswert.
2. Das Redundanzpotential der Backbone-Topologie soll durch MPLS weitestgehend ausgenutzt werden.
3. MPLS soll auf effiziente und skalierbare Weise eingesetzt werden.
4. Die Seiteneffekte des Einsatzes von MPLS auf bestehende Netzstrukturen sollen minimal sein. Insbesondere soll die Einführung von MPLS die Stabilität und Zuverlässigkeit des Netzes nicht gefährden.
5. Durch den Einsatz von MPLS soll keine zusätzliche Angriffsfläche entstehen, durch die die Sicherheit des Netzes kompromittiert werden könnte.
6. Die Migration bzw. Implementierung von MPLS soll möglichst zügig und reibungslos realisierbar sein.
7. Die Option einer zukünftigen nutzerbezogenen Abrechnung im MWN soll auch nach der Einführung von MPLS erhalten bleiben.
8. Auf proprietäre Funktionen und Technologien soll weitestgehend verzichtet werden. Im Bedarfsfall ist sorgfältig gegenüber dem Nutzen abzuwägen.

4. Planung

Anhand der in der Anforderungsanalyse ermittelten Informationen soll nun ein Konzept zur Einführung von MPLS entwickelt werden. Das Ziel der Planungsphase ist es, ein Grundgerüst dafür zu entwerfen und die verfügbaren Optionen zur praktischen Umsetzung der Anwendungsfälle zu untersuchen.

4.1. Vorgehensweise

In einigen Anwendungsfällen ist davon auszugehen, dass die zur Umsetzung erforderlichen Eingriffe in die Netzkonfiguration temporär einen Konnektivitätsverlust zur Folge haben, so dass die Umstellung vsl. nicht während des Live-Betriebs stattfinden kann. Im MWN ist jedoch regulär nur einmal wöchentlich ein verhältnismäßig kurzes Wartungsfenster von ca. 2 Stunden vorgesehen. Zur Erleichterung der Migration erscheint es daher sinnvoll, das Konzept derart zu gestalten, dass es in kleinen separaten Schritten umgesetzt werden kann. Um dies zu erreichen, wird ein möglichst geringer Grad an Querabhängigkeiten zwischen den Lösungsansätzen für die Anwendungsfälle angestrebt, damit sie weitestgehend unabhängig voneinander implementiert werden können.

Um zu geeigneten Lösungsansätzen zu gelangen, werden für jeden Anwendungsfall verschiedene Optionen, die aussichtsreich erscheinen, erforscht und deren Eigenschaften analysiert. Anschließend werden die Interoperabilität der Lösungsansätze untereinander betrachtet und ihre Integration in das bestehende Umfeld des MWN erörtert. Die Ergebnisse werden am Ende zum Konzeptentwurf zusammengeführt.

4.2. Exploration von Lösungsansätzen

Die wesentlichen Kriterien zur Bewertung verschiedener Lösungsansätze sind ihr potentieller Nutzen gegenüber der bisherigen Situation, ihre technische Machbarkeit und ihre Komplexität bzw. der mit ihnen verbundene OAM-Aufwand. Der jeweils am besten geeignetste Ansatz wird anschließend anhand seiner funktionalen Aspekte konkretisiert. Dabei werden auch betriebsrelevante Implikationen wie z.B. das Verhalten des Netzes im Normalbetriebs- und Fehlerfall untersucht.

4.2.1. Standortübegreifende VLANs

Die bisherige Umsetzung von VPNs im MWN basiert auf VLANs und damit auf der Sicherungsschicht. Da dort Ethernet eingesetzt wird, stellt VPLS eine naheliegende Alternative dazu dar.

Der wesentliche Vorteil von VPLS gegenüber standortübergreifenden VLANs besteht in der einfacheren Handhabung. Die Einrichtung und Pflege von VPNs ist deutlich flexibler und weniger aufwendig. Während der Konfigurationsaufwand bei dem bisherigen Konzept

auf VLAN-Basis linear mit der Anzahl der zwischen den Teilnetzen liegenden Netzknoten zunimmt, ist der Aufwand mit VPLS diesbezüglich konstant (Abbildungen 4.1 und 4.2).

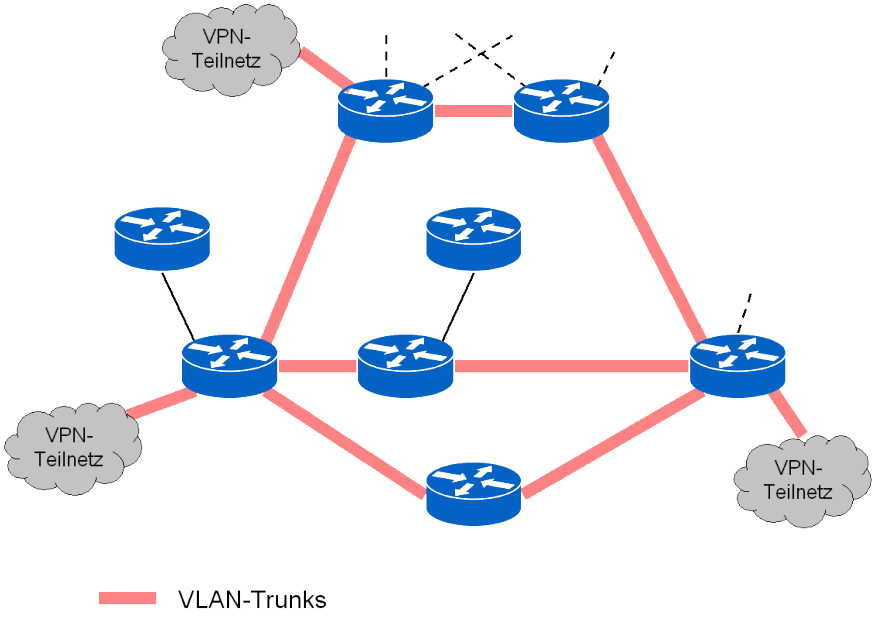

Abbildung 4.1.: Zu konfigurierende Schnittstellen, L2VPN mittels VLANs

VPLS erscheint als Lösungsansatz insbesondere deshalb sinnvoll, weil aus Sicht der VPN-Nutzer keine Veränderung durch die Umstellung stattfinden würde. Die Migration könnte somit unabhängig von ihnen und daher mit vsl. relativ geringem Aufwand durchgeführt werden. Dass Layer 2-VPNs deutlich einfacher einzurichten und zu betreiben sind als Layer 3-VPNs, ist ein zusätzlicher Bonus. Für zukünftige VPNs ist diese Designentscheidung nicht bindend, d.h. zu einem späteren Zeitpunkt neu ins MWN hinzukommende VPNs könnten auch auf der Vermittlungsschicht umgesetzt werden, sofern das opportun erscheint.

Modell

Anstatt wie bisher für alle Teilnetze eines VPN gemeinsam ein einzelnes VLAN mittels Trunking über das Backbone aufzuspannen und die Teilnetze diesem VLAN zuzuordnen, wird nun jedes Teilnetz jeweils in ein separates VLAN eingeteilt. Für das VPN wird eine VFI definiert und die neuen VLANs dieser VFI zugeordnet. Zwischen den beteiligten PE-Routern werden paarweise PWE3-Verbindungen untereinander aufgebaut, so dass eine logische Vollvermaschung resultiert.

Um die Konfiguration übersichtlich zu halten, kann im MWN für die VLANs der VPN-Teilnetze eine identische Nummerierung verwendet werden, da sie nunmehr über Router voneinander getrennt sind und sich daher in unterschiedlichen Adressräumen befinden. Idealerweise wird die Nummer des bisherigen standortübergreifenden VLAN übernommen.

Der Datenverkehr zwischen den Teilnetzen wird dann, sofern das Zielnetz nicht lokal an demselben PE-Router angeschlossen ist, über die PWE3-Verbindungen geleitet. Der Label Stack der übertragenen Rahmen beinhaltet dabei neben dem Transport-Label ein Service-

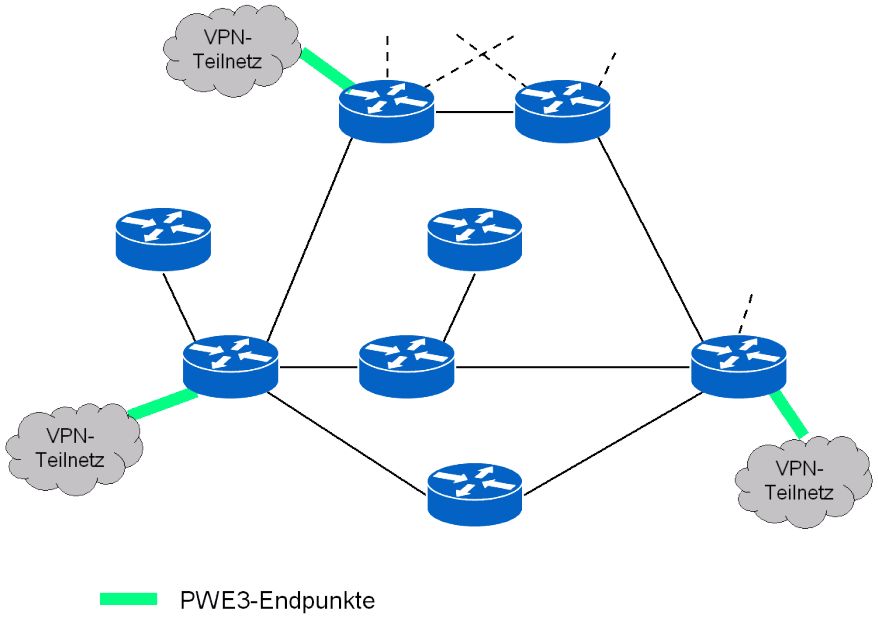

Abbildung 4.2.: Zu konfigurierende Schnittstellen, L2VPN mittels VPLS

Label zur Kennzeichnung der jeweiligen PWE3-Verbindung (bezeichnet als Interworking- bzw. VC-Label).

Auf welche Weise die LSPs aufgebaut werden und über welche Pfade innerhalb des Backbone sie genau führen, ist für die Funktion des VPN unerheblich. Im einfachsten Fall werden sie daher automatisch mittels LDP signalisiert. Falls TE-Restriktionen bestehen, können auch gesondert eingerichtete TE-LSPs für die PWE3-Verbindungen verwendet werden.

Unmittelbar für die Funktion des VPN verantwortlich sind die PE-Router, die Start- bzw. Endpunkte der PWE3-Verbindungen darstellen und Label Push bzw. Label Pop des VC-Label durchführen; ggf. dazwischenliegende P-Router betrachten lediglich das Transport-Label des Label Stack.

Kommunikationsablauf

Aus Sicht der VPN-Nutzer verhält sich das Backbone wie ein gewöhnlicher Ethernet-Switch. Das Lernen der MAC-Adressen läuft analog ab, d.h. die Absenderadresse eines übertragenen Rahmens wird entsprechend ihres Ursprungs mit einer PWE3-Verbindung bzw. einer lokalen VLAN-Schnittstelle assoziiert. Ein zu übertragender Rahmen wird, falls die Empfängeradresse bereits gelernt wurde, somit nur über die korrekte PWE3-Verbindung in das Zielnetz geleitet. Lediglich bei unbekannten MAC-Adressen wird der Rahmen vom PE-Router repliziert und an alle übrigen PE-Router (und damit in alle anderen Teilnetze) übertragen.

Inbetriebnahme und Wartung

Zur Erstellung eines VPN ist eine initiale Konfiguration an den PE-Routern durchzuführen. Bei nachfolgenden Topologieänderungen bzw. -erweiterungen im Backbone sind i.A. keine

zusätzlichen Arbeiten erforderlich, da die mit den PWE3-Verbindungen assoziierten LSPs bei Bedarf automatisch neu aufgebaut werden.

Eine genaue Prüfung und ggf. Überarbeitung der Konfiguration wird notwendig bei Änderungen am VPN selbst (z.B. das Hinzufügen/Entfernen von Teilnetzen oder eine Änderung der VFI) und bei Änderungen an den PE-Routern, die den Aufbau der PWE3-Verbindungen betreffen (z.B. Änderungen von Schnittstellen-Adressen). Besondere Aufmerksamkeit ist geboten bei der Verwendung von TE-LSPs, da sich hier u.U. auch Änderungen der Topologie auf die Konnektivität des VPN auswirken können (z.B. wenn die zulässigen Pfade durch das Backbone über CBR-Constraints restringiert wurden).

Fehlerverhalten

Bei Ausfall einer PWE3-Verbindung können sich die Teilnetze, die durch sie miteinander verbunden werden, bis zur Umleitung des zugehörigen LSP nicht mehr gegenseitig erreichen. Die übrigen Teilnetze des VPN sind davon nicht betroffen. Eine zwischenzeitliche Überbrückung mittels MPLS-FRR o.ä. ist möglich.

Bei Ausfall eines PE-Routers werden alle von ihm versorgten Teilnetze des VPN unerreichbar. Eine Überbrückung ist nur möglich, falls die betroffenen Teilnetze redundant an weitere PE-Router angebunden sind.

Internet-Zugang

Bis zum Abschluss der Arbeit konnte nicht geklärt werden, ob es mit der vorhandenen Hardware möglich ist, als VPLS-Provider einen Zugang in das öffentliche Internet einzurichten. Während dies für Layer 3-VPNs definitiv möglich ist, konnte für VPLS weder durch Recherchen in der Dokumentation von Cisco noch aus anderen Quellen eine klare Antwort ermittelt werden. Aufgrund dessen wird dieser Aspekt für die weitere Planung verworfen.

Theoretisch wäre eine solche Konstellation ohne weiteres denkbar. Dazu wäre es erforderlich, auf einem der PE-Router eine virtuelle Schnittstelle einzurichten und diese der entsprechenden VFI zuzuordnen. Der Schnittstelle müsste dann eine IP-Adresse zugewiesen werden, die von den Hosts innerhalb des VPN als Standard-Gateway verwendet werden könnte.

Gegen die Machbarkeit eines solchen Modells spräche allerdings das Argument, dass diese Vorgehensweise bei einem Layer 2-VPN der Unabhängigkeit vom Vermittlungsschichtprotokoll zuwiderliefe. Auch wäre die Problematik möglicher Überschneidungen von Adressbereichen zu lösen.

4.2.2. Policy-basiertes Routing

Policies werden i.A. umgesetzt, indem gewisse Daten explizit entlang bestimmter Pfade weitergeleitet werden bzw. bei der Weiterleitung bestimmte Pfade vermieden werden. Im MWN wird dies derzeit praktisch erreicht durch die Konfiguration von Filterregeln in einer Zugriffskontrollliste (ACL) zur Auswahl des für die Policy relevanten Datenverkehrs und die statische Vorgabe eines oder mehrerer Transitpunkte (d.h. einer Folge von Netzelementen), die auf dem Weg zum Ziel durchlaufen werden müssen. An zwei wichtigen Fällen sollen exemplarisch mögliche Vereinfachungen diskutiert werden.

- Fall 1: Umleitung von in das öffentliche Internet gerichteten Datenverkehr aus Zugangsnetzen der MPG über das IPP. Dieser Fall wurde bereits in Kapitel 3 als Anwen-

dungsfallbeispiel beschrieben.

- Fall 2: Umleitung von Datenverkehr aus Zugangsnetzen, denen private Adressbereiche zugewiesen wurden, über den NAT-o-MAT. Bevor ausgehender Datenverkehr von Hosts mit privaten IP-Adressen in das öffentliche Internet weitergeleitet werden kann, muss eine Adressübersetzung stattfinden. Für den aus dem Internet zurückkommenden Verkehr ist der Vorgang umgekehrt durchzuführen. Datenverkehr von externen Systemen, die via IPSec-VPN-Client o.ä. mit dem MWN verbunden sind, wird ebenfalls über den NAT-o-MAT geleitet.

Intuitiv erscheinen zwei technische Ansätze mittels MPLS eine genauere Untersuchung wert. (Im Folgenden werden die Zugangsnetze, von denen aus der umzuleitende Datenverkehr im Backbone ankommt, als Quellnetze bezeichnet.) Die erste Idee besteht darin, TE-Tunnel zwischen den Quellnetzen und den Transitpunkten einzusetzen und zu versuchen, dadurch eine Vereinfachung gegenüber der bisherigen Konfiguration zu erreichen. Die zweite Idee geht das Problem indirekt an und besteht darin, die Quellnetze zusammen mit den Transitpunkten in einem VPN zusammenzufassen.

TE-Tunnel

Normalerweise sind statische Routen an jedem Router entlang des gewünschten Pfades zu konfigurieren, da für gewöhnlich jeweils nur der Next-Hop angegeben werden kann. Der betroffene Datenverkehr muss dabei auf jedem Router bis zum Ziel durch eine Filterregel selektiert und auf die statische Route gelenkt werden. Ein TE-Tunnel böte den Vorteil, dass die Konfiguration von Filterregeln auf Start- und Endpunkt des Tunnels beschränkt werden könnte. Es wäre dann zwischen jedem Quellnetz und dem letzten Transitpunkt (oder so nahe wie möglich davor, abhängig von der Anwensenheit MPLS-fähiger Netzknoten) ein Tunnel zu erstellen; zwischenliegende Transitpunkte könnten mittels CBR berücksichtigt werden.

Im MWN wird ein vergleichbarer Effekt aber bereits durch Verwendung eines speziellen, eigens für Policies errichteten Backbone-weiten VLAN erreicht (VLAN 998). Der umzuleitende Datenverkehr wird beim Übergang von den jeweiligen Quellnetzen zum Backbone in dieses VLAN eingespeist und am Router `csr1-2wr` wieder entnommen. Von dort aus werden die Pakete direkt an den nächsten Transitpunkt weitergeleitet. Das VLAN isoliert die Policy-basierten Routen ähnlich wie ein TE-Tunnel.

Eine geringfügige Vereinfachung ergäbe sich höchstens dadurch, dass bei einer Verwendung von TE-Tunneln das VLAN 998 nicht mehr auf den VLAN-Trunks im Backbone konfiguriert werden müsste. Da es sich jedoch nur um ein einzelnes VLAN handelt, das für verschiedene Policies gemeinsam genutzt wird, erscheint der Aufwand vernachlässigbar klein und wird mehr als kompensiert durch die Tatsache, dass die Verwaltung zahlreicher TE-Tunnel aufwendiger als die eines einzelnen VLAN zu bewerten ist. Eine signifikante Aufwandseinsparung bzgl. OAM oder Verringerung der Komplexität sind unter dem Strich folglich nicht zu erwarten. Dieser Lösungsansatz erscheint somit eher ungeeignet.

Layer 3-VPN

Der Gedanke hinter der Verwendung von VPNs für Policy-basiertes Routing ist, dass die Konfiguration von statischen Routen überflüssig wird, wenn man alle übrigen Routen für den umzuleitenden Datenverkehr unsichtbar macht. Hierbei wird davon ausgegangen, dass nur ein

einzelner Transitpunkt vorliegt; dies trifft in den beiden o.g. Fällen zu. Die Quellnetze und der Transitpunkt wären jeweils als VPN-Teilnetze zu konfigurieren. Die vormals statische Route könnte dann als Default-Route deklariert und automatisch mittels eines Routing-Protokolls innerhalb des VPN verteilt werden.

Voraussetzung hierfür wäre, dass das Netzelement am Transitpunkt in der Lage ist, seine IP-Adresse über BGP als Next-Hop für die Default-Route zu propagieren. Für Fall 1 sollte dies unproblematisch sein, da das Netzelement im IPP ein Router ist; für Fall 2 sieht der Netzbetreiber ebenfalls die technische Machbarkeit gegeben, da es sich beim NAT-o-MAT bzw. beim Secomat jeweils um Verbunde von Linux-Rechnern handelt, die entsprechend konfiguriert werden könnten. Alternativ könnte auch der unmittelbar vor dem NAT-o-MAT befindliche Router verwendet und für den letzten Hop eine statische Route eingerichtet werden.

Für die Rückrichtung des Verkehrs müsste der Transitpunkt seine IP-Adresse außerhalb des VPN im IGP als Next-Hop für die Adressbereiche der Quellnetze propagieren. Auf diese Weise kann der Datenverkehr zurück in das VPN gelangen. Durch die Kapselung der Quellnetze in einer VRF sind diese für das IGP nicht sichtbar, so dass es nicht zu Konflikten kommen dürfte.

Im Idealfall fiele außer der Errichtung des VPN und den genannten Modifikationen am letzten Transitpunkt kein weiterer Konfigurationsaufwand an. In den beiden Fallbeispielen könnten theoretisch jeweils alle von der Policy betroffenen Standorte mittels eines einzelnen VPN abgedeckt werden. Als Vorteil ergäbe sich ein etwas höherer Automatisierungsgrad, da vsl. an weniger Stellen explizit IP-Adressen angegeben werden müssten und die Konfiguration der VPNs grundsätzlich auf die PE-Router beschränkt wäre.

Allerdings unterschiede sich dieses Konzept erheblich von dem bisherigen Verfahren, so dass es in den Fallbeispielen nicht ohne einige unerwünschte Seiteneffekte bliebe:

- Fall 1: Isolierte man die Quellnetze der MPG in einem VPN, so würde nicht nur der Datenverkehr in das öffentliche Internet über das IPP geleitet, sondern zwingenderweise der gesamte Datenverkehr an Ziele außerhalb der Quellnetze. Insbesondere könnten Ziele innerhalb des MWN nicht mehr direkt erreicht werden, sondern nur noch über einen Umweg durch das X-WiN. Es konnte keine Methode ermittelt werden, wie diese Restriktion auf einfache Weise umgangen werden könnte.

- Fall 2: Der NAT-o-MAT wird aus Ausfallsicherheits- und Lastverteilungsgründen über mehrere verschiedene IP-Adressen erreicht. Anders als viele IGP unterstützt BGP jedoch kein ECMP; d.h. um eine identische Funktionalität zu erhalten, wäre ein zusätzlicher Aufwand notwendig (z.B. in Form eines anderweitigen SLB-Mechanismus oder über eine Variante mit statischen Routen).

Darüber hinaus ist unklar, ob bzw. wie ein solches Modell auf der vorhandenen Hardware realisiert werden könnte. Das in der Literatur und von Herstellern verwendete Standardmodell zu MPLS Layer 3-VPNs nimmt implizit an, dass in den vom VPN-Nutzer betriebenen Teilnetzen CE-Router existieren, über die diese Teilnetze an die PE-Router des ISP angebunden werden. Diese Annahme kann für die Quellnetze in den beiden o.g. Fällen nicht getroffen werden, was für das Routing des Datenverkehrs in Rückrichtung kritisch ist. Bei genauerem Studium der IETF-Standards [RR99] [RR06] ergibt sich zwar, dass sie keine zwingende Voraussetzung ist; in den Dokumenten dazu sind für “CE devices” ausdrücklich

	Statisches Routing über Backbone-weites VLAN	**MPLS Layer 3-VPN**
Quellnetz → Backbone	Festlegung einer Route-Map in der Interface-Konfiguration des Quellnetz-VLAN, Angabe des Next-Hop in der Route-Map	Festlegung eines VRF-Identifikators in der Interface-Konfiguration des Quellnetz-VLAN
Backbone → Transitpunkt	Festlegung einer Route-Map in der Interface-Konfiguration von VLAN 998, Angabe des Next-Hop in der Route-Map	Festlegung eines VRF-Identifikators in der Konfiguration des Interface, an dem der Transitpunkt angeschlossen ist
Sonstiges	Konfiguration von VLAN 998 auf den VLAN-Trunks im Backbone	Erstellung des VPN und zusätzliche BGP-Konfiguration auf allen PE-Routern, BGP-Konfiguration des CE-Systems am Transitpunkt

Tabelle 4.1.: Aufwandsvergleich des derzeitigen Konzepts Policy-basierten Routings mit dem VPN-Konzept

auch Netzelemente wie Switches und einzelne Hosts als zulässig angegeben. In den Konfigurationsrichtlinien von Cisco wurden jedoch keine Hinweise auf eine derartige Konstellation gefunden.

Um zu beurteilen, ob diese Schwierigkeiten gerechtfertigt sind, erscheint eine genauere Klärung des konkreten praktischen Nutzens des Modells angebracht. In einer Gegenüberstellung von jetzigem Verfahren und VPN-Konzept (Tabelle 4.1) wird ersichtlich, dass der Umfang der Konfigurationsaufgaben sich kaum verringert, sondern sich im Wesentlichen nur verlagert.

Angesichts dieses Ergebnisses und der genannten Seiteneffekte erscheint das Argument eines etwas höheren Automatisierungsgrades recht schwach. Kalkuliert man außerdem ein, dass die Verwendung einer zusätzlichen Technologie wie VPNs die Komplexität zunächst einmal erhöht und der Aufwand einer Umstellung sich in absehbarer Zeit amortisieren sollte, so macht es keinen Sinn, diesen Lösungsansatz weiter zu verfolgen.

Layer 2-VPN

Diese Option scheidet als Alternative zu einem Layer 3-VPN aus, da dadurch alle Quellnetze der MPG bzw. alle Quellnetze mit privaten IP-Adressbereichen jeweils in eine gemeinsame Broadcast-Domäne verlegt würden. Ohne massive Änderungen an der Netzstruktur und -konfiguration würde dies nicht funktionieren; desweiteren sprechen auch eine Vielzahl von technischen und administrativen Gründen dagegen.

Fazit

Keiner der untersuchten Lösungsansätze stellt im Vergleich zur gegenwärtigen Umsetzung von Policy-basiertem Routing eine klare und überzeugende Verbesserung in Aussicht. Weitere

Alternativen auf Grundlage von MPLS sind nach dem jetzigen Stand der Technik nicht bekannt.

4.2.3. Bandbreitenbegrenzung

Traffic Policing kann auf der gegebenen Hardware auf verschiedene Arten konfiguriert werden. Bei allen Varianten wird die Bandbreitenbegrenzung an eine physische oder virtuelle Schnittstelle sowie an eine bestimmte Übertragungsrichtung gekoppelt. Neben diesem als Aggregate Policing bezeichneten Vorgehen unterstützt die Hardware auch das sog. Microflow Policing, das eine Anwendung von Bandbreitenbegrenzung auf der Granularitätsebene von einzelnen Datenströmen erlaubt. Darüber hinaus sind zusätzlich feingranularere Selektionsmechanismen wie z.B. IP-Adresslisten oder Protokolltypen möglich.

Auf LSPs ist Traffic Policing nicht direkt anwendbar:

- Über LDP signalisierte LSPs werden automatisch aufgebaut und sind daher unter IOS nicht explizit als Schnittstellen ansprechbar.
- TE-LSPs werden unter IOS zwar als Tunnel-Schnittstellen abstrahiert, jedoch wird auf diesem Typ von Schnittstellen kein Traffic Policing unterstützt.[1]

Im MWN stellen diese Einschränkungen aber kein Hindernis dar, da Bandbreitenbegrenzungen sinnvollerweise möglichst am Rand der Netzes angewendet werden und dort normalerweise VLAN-Schnittstellen vorliegen.

Traffic Policing kann effektiv in Zusammenwirkung mit VPNs eingesetzt werden. Ein konkreter Einsatzzweck, der in diesem Abschnitt genauer betrachtet wird, ergibt sich im MWN mit VPLS, das standortübergreifende VLANs zukünftig ersetzen soll. Da die Teilnetze dieser VPNs in den meisten Fällen einer eigenen administrativen Instanz unterstehen, entspricht dies einer Konstellation wie in Kapitel 3 erläutert, in der eine Bandbreitenbegrenzung als vorbeugende Sicherheitsmaßnahme angebracht ist.

Modell

Zur Kontrolle der von einem VPN genutzten Netzkapazitäten ist es erforderlich, auf den PE-Routern für jedes VPN-Teilnetz einen Datenvolumenzähler zu konfigurieren. Es ist nicht ohne weiteres möglich, den Datenverkehr innerhalb eines VPN in seiner Gesamtheit als einzelne Größe zu begrenzen, da hierzu eine Koordination der einzelnen Zähler notwendig wäre, d.h. der Datendurchsatz der Teilnetze kann nur individuell beschränkt werden. Ausnahme: sind mehrere Teilnetze eines VPN an demselben PE-Router angebunden, so kann für diese über Aggregate Policing ein gemeinsamer Zähler eingerichtet werden. Bei der Umsetzung von Traffic Policing als Sicherheitsvorkehrung sind diese Beschränkungen jedoch unproblematisch.

Die Bandbreitenbegrenzung wird an den VLAN-Schnittstellen der VPN-Teilnetze konfiguriert. Um nichtkonformen Datenverkehr ggf. so nahe wie möglich am Entstehungsort einzudämmen, muss das Traffic Policing für die Übertragungsrichtung vom Teilnetz aus in das Backbone hinein angelegt werden.

[1] Eine Ausnahme sind GRE-Tunnel [Cisb].

Traffic Policing wird mittels eines Token-Bucket-Algorithmus gesteuert. Bei der Konfiguration des Zählers ist daher neben dem zulässigen mittleren Datendurchsatz auch das maximale Burst-Datenvolumen zu definieren. Diese Angabe ist erforderlich, da kurzzeitige Überschreitungen des mittleren Datendurchsatzes in einem vorgegebenen Rahmen toleriert werden sollen.

Kommunikationsablauf

Beim Übergang von Paketen aus einem Teilnetz in das Backbone wird das übertragene Datenvolumen am jeweiligen PE-Router laufend aktualisiert und überprüft. Wird das gesetzte Limit verletzt, so wird für die nichtkonformen Daten eine vorher definierte Aktion ausgeführt (i.A. das Verwerfen der überschüssigen Pakete).

Inbetriebnahme und Wartung

Die Konfiguration eines Zählers erfolgt zusammen mit dem VPN an den PE-Routern. Es gelten daher dieselben Bedingungen wie für den Anwendungsfall der standortübergreifenden VLANs.

Zusammenspiel mit anderen MPLS-Funktionen

Es sind zwei Zusammenhänge zwischen Traffic Policing und QoS-relevanten Mechanismen von MPLS erkennbar:

- Üblicherweise resultiert eine Verletzung von durch Traffic Policing vorgegebenen Grenzen in Paketverlust. Bei Anwendungen, die vorübergehend starke Lastspitzen aufweisen, kann es dadurch häufig zu Paketverlust kommen, wenn das maximale Burst-Volumen zu restriktiv konfiguriert ist. Eine weniger restriktive Konfigurationen kann andererseits dazu führen, dass das Traffic Policing an Effektivität verliert. Bei schwankenden Verkehrscharakteristika kann das Finden einer guten Balance schwierig sein. Eine Alternative zum Verwerfen bestünde darin, die Priorität der nichtkonformen Pakete im Vergleich zum übrigen Datenverkehr herabzustufen, so dass es erst im Fall einer Überlastung des Netzes zu Paketverlust käme.

- Um in einem Netz QoS-Garantien erbringen zu können, muss nicht nur die Zuteilung von Kapazitäten geplant werden, sondern auch durchgesetzt werden, dass diese Planung eingehalten wird. Im Kontext von MPLS-TE kann die erste Aufgabe beispielsweise von RSVP-TE und CBR übernommen werden, während Traffic Policing die zweite Aufgabe erfüllen kann.

4.2.4. Verkehrsklassen

Derzeit ist vom Netzbetreiber kein unmittelbarer Einsatz von Verkehrsklassen im MWN geplant. Vor diesem Hintergrund besteht das Ziel dieses Abschnitts nicht darin, ein konkretes Umsetzungskonzept zu entwerfen. Stattdessen wird die Rolle untersucht, die MPLS bei der technischen Realisierung von Verkehrsklassen auf der vorhandenen Cisco-Hardware einnehmen würde, und welche Möglichkeiten sich in diesem Zusammenhang im MWN böten.

Prinzipielle QoS-Ansätze mit MPLS

Allgemein sind zwei unterschiedliche Ansätze zur Umsetzung von Verkehrsklassen mit MPLS denkbar. Der am häufigsten eingesetzte Ansatz basiert auf MPLS CoS und setzt bei dem Weiterleitungsvorgang an. Dazu werden bei der Übertragung mehrere Warteschlangen eingesetzt, die sich in bestimmten QoS-Parametern unterscheiden. Diese Unterschiede werden sichtbar, sobald sich die betreffende Übertragungsstrecke ihrer Kapazitätsgrenze nähert und nicht mehr alle Verkehrsklassen vollständig bedient werden können. Niedrigpriore Datenpakete werden dann mit größerer Wahrscheinlichkeit verworfen als höherpriore Pakete.

Ein alternativer Ansatz zur Trennung von Verkehrsklassen mit MPLS-TE besteht darin, den Datenverkehr unterschiedlicher Klassen über verschiedene Teile eines Netzes zu lenken, die unterschiedliche Charakteristika in Bezug auf QoS-Parameter aufweisen. Beispielsweise können mittels TE höherpriore Pakete über einen gesonderten Pfad geführt werden, der eine hohe Kapazität und niedrige Latenz besitzt, während niederpriore Pakete über stärker ausgelastete oder längere Pfade geführt werden. Dies kann insbesondere dann effektiv sein, wenn in einem Netz unterschiedliche Transporttechnologien parallel genutzt werden. Da dies aber im MWN nicht der Fall ist und die Übertragungsstrecken im Backbone alle als gleichwertig gelten können, ist dieser Ansatz derzeit eher nicht von Bedeutung.

Wirkungsweise von MPLS CoS

Die technische Umsetzung von Verkehrsklassen gliedert sich in mehrere spezifische Abläufe:

- Zunächst müssen die Datenströme unterschiedlicher Klassen auf geeignete Weise differenziert werden. Dies umfasst die in Kapitel 2 eingeführten Konzepte der Klassifikation und Markierung. Dieser Prozess findet i.A. am Rand des Netzes statt. Zur Klassifikation können schichtenübergreifend vielfältige Informationen herangezogen werden. Typische Kriterien sind z.B. Quell-/Zieladresse, die empfangende Schnittstelle und Portnummern bzw. die damit korrespondierenden Protokolle.

- Wirksam werden Verkehrsklassen schließlich durch eine klassenspezifische Verarbeitung der Datenströme beim weiteren Transport. Diese umfasst die Bereiche Überlastvermeidung (Congestion Avoidance) und Überlasthandhabung (Congestion Management). Diese Prozesse finden i.A. im Kern des Netzes statt. Proaktive Mechanismen wie z.B. WRED sind ersterem Bereich zuzuordnen, während das Scheduling betreffende Verfahren wie z.B. WFQ oder Priority Queueing zu letzterem gehören.

MPLS CoS betrifft von den genannten Abläufen im Wesentlichen nur die Markierung. Alle anderen Vorgänge sind von MPLS unabhängig. Insbesondere ist durch MPLS CoS nicht näher definiert, in welcher Weise sich die Unterschiede zwischen den Verkehrsklassen konkret manifestieren. Der Grund dafür ist, dass die Paketweiterleitung auf modernen Routern praktisch vollständig von der Hardware umgesetzt wird. Technische Einflussfaktoren bzgl. der Charakteristik der Weiterleitung wie z.B. das interne Management, die verfügbare Anzahl und die eingesetzten Algorithmen der Warteschlangen sind zumeist proprietär und von den konkreten Schnittstellenmodulen abhängig [Cis09].

Zur Markierung von MPLS-Rahmen existieren die Techniken der E-LSPs und L-LSPs. Bisher wurden L-LSPs jedoch nur für das sog. Cell Mode MPLS (MPLS auf Basis von ATM) implementiert [Ike02]. Für die Markierung von MPLS-Rahmen auf den Routern im MWN

Exp-Bits	Verkehrsklasse
5	VoIP
4	Multimedia
0	Standard
3	Out-of-Profile
2	Backup
1	Filesharing

Tabelle 4.2.: Beispiel für eine mögliche Ausgestaltung von Verkehrsklassen im MWN (Priorität von oben nach unten sinkend)

sind daher derzeit lediglich die Exp-Bits verfügbar, so dass die Anzahl simultan realisierbarer Verkehrsklassen auf 8 beschränkt ist. Dies ist aber unproblematisch, da im MWN derzeit nur potentiell bis zu 6 Prioritätsstufen identifiziert werden können (Tabelle 4.2).

Das QoS-Management auf Cisco-Routern wird von der sog. Policy Feature Card (PFC) übernommen. Diese verwendet intern ein eigenes System von Klassen (Internal DSCP). Die Abbildung von Exp-Bits auf Internal DSCP und auf spezifische Parameter von Schnittstellen-Warteschlangen ist durch weitreichende Konfigurationseinstellungen in hohem Maße anpassbar. Die numerische Ordnung der Exp-Bits muss dabei nicht notwendigerweise eine Ordnung bzgl. der Verkehrsklassen repräsentieren, d.h. der Standardwert 0 muss beispielsweise nicht unbedingt die niedrigste Priorität darstellen.

4.2.5. Lastverteilung

Die Umsetzung einer Lastverteilung wird zumeist durch zwei Ziele motiviert. Zum einen sollen Hot-Spots soweit wie möglich eliminiert werden, d.h. das Maximum der prozentualen Auslastungen der beteiligten Übertragungsstrecken im Netz soll minimiert werden. Zum anderen sollen die Netzressourcen effizient genutzt werden, d.h. die mittlere Auslastung über alle beteiligten Übertragungsstrecken soll minimiert werden.

Eine Schwierigkeit bei der Netzplanung ergibt sich dadurch, dass diese beiden Ziele häufig in Konkurrenz zueinander stehen. Zu berücksichtigen sind außerdem zeitliche Schwankungen der Auslastungen. Weiterhin sollen die Hop-Anzahlen von Verkehrsbeziehungen, obgleich diese kein entscheidendes Kriterium wie bei konventionellem SPF-Routing darstellen, möglichst gering gehalten werden.

MPLS kann zur Lastverteilung eingesetzt werden, indem an geeigneten Stellen im Netz TE-Tunnel platziert werden. Auf Cisco-Routern kommt zur Signalisierung der LSPs RSVP-TE zum Einsatz; CR-LDP wird derzeit nicht unterstützt.

Beim Entwurf eines Konzepts stellt sich zunächst die Frage, in welchem Umfang der Einsatz einer Lastverteilung im MWN praktikabel wäre und sich lohnen würde. Hierfür sind mehrere Möglichkeiten denkbar.

Netzweite automatische Lastverteilung

Als Idealmodell kann man eine netzweite Lastverteilung betrachten, die hochgradig automatisch gesteuert wird. Theoretisch könnte man diesem Ziel auf der vorhandenen Hardware unter Ausnutzung einiger spezieller Verfahren recht nahe kommen. Grundlage dafür wäre ein IntServ-ähnliches Kommunikationsmodell auf Basis von MPLS.

Vereinfacht dargestellt sähe das derart aus, dass beim Aufbau jedes LSP die vsl. dafür benötigte Bandbreite vorab zu reservieren wäre. Da RSVP-TE zur Signalisierung von TE-LSPs verwendet wird, würden diese Informationen ohnehin erfasst und verwaltet. Unter der Voraussetzung, dass der gesamte Datenverkehr im Netz über MPLS weitergeleitet wird, könnten dann für jede Übertragungsstrecke die zu einem bestimmten Zeitpunkt verfügbaren Netzkapazitäten berechnet werden. Diese Information könnte dann bei der Wegewahl berücksichtigt werden, indem sie über IGP-Erweiterungen (z.B. durch OSPF-TE) zunächst netzweit propagiert und dann in den CSPF-Algorithmus einbezogen wird. CSPF wählt standardmäßig aus den verfügbaren Wegen denjenigen mit der höchsten freien Bandbreite aus und, falls das Kriterium nicht zu einem eindeutigen Ergebnis führt, daraus denjenigen mit der geringsten Hop-Anzahl [Cis02].

Entscheidend für die Flexibilität des Konzepts wäre schließlich ein Mechanismus, um automatisch auf Veränderungen der benötigten Bandbreite eines LSP zu reagieren. Hierfür eignete sich die Cisco-proprietäre Funktion "MPLS Auto-Bandwidth", die periodisch das tatsächliche Verkehrsaufkommen von TE-LSPs misst und bei Abweichungen jenseits definierbarer Schwellenwerte die ursprünglichen Reservierungen angleicht. Falls notwendig, könnte ein Rerouting einzelner LSPs vorgenommen werden, um eine Überlastung zu verhindern [MW09]. Für die Nutzer wären diese Abläufe vollkommen transparent.

Eine netzweite Lastverteilung hätte den bedeutenden Nutzen, dass die Nutzung der vorhandenen Netzressourcen global optimiert werden könnte. Obgleich technisch vsl. machbar, erscheint das Konzept mit Hinblick auf das Szenario und aus einer Reihe konzeptioneller Gründe z.Zt. jedoch eher unpraktikabel:

- Bei einer Ring- bzw. ringähnlichen Topologie, wie sie im MWN derzeit noch vorherrschend ist, ist nur ein begrenzter Nutzen zu erwarten, da in einer solchen keine hinreichend große Auswahl an möglichen Alternativrouten verfügbar ist. Es lassen sich in einem Ring sogar leicht Szenarien konstruieren, bei denen eine Lastverteilung kontraproduktiv wäre (Abbildung 4.3).

- Für eine netzweite Lastverteilung müssten alle Router, die Zugangsnetze (d.h. potentielle Quellen und Senken) bedienen, jeweils paarweise mit TE-LSPs untereinander verbunden werden. Die gegebene Topologie des Backbone ist jedoch nicht hierarchisch in ein Core Layer und ein Distribution Layer unterteilt, weshalb eine Vollvermaschung aus TE-LSPs resultieren würde. Die damit verbundene Komplexität und der Aufwand wären unverhältnismäßig hoch.

- Es gibt derzeit keine Übertragungsstrecken im Backbone, die exklusiv aus Redundanzgründen existieren und ohne eine netzweite Lastverteilung ungenutzt blieben.

- Der Einrichtungsaufwand wäre beträchtlich, aber angesichts der derzeitigen Verhältnisse bei der Netzauslastung ergäbe sich kurzfristig kein messbarer Nutzen.

- Durch die feedbackgesteuerten selbständigen Anpassungen entstünde ein Regelkreis. Es ist unklar, ob sich dies negativ auf die Stabilität des Netzes auswirken würde oder gar zu Oszillationen führen könnte.

- Es kann nicht sicher angenommen werden, dass im Backbone zukünftig nicht auch Router anderer Geräteherstellers als Cisco zum Einsatz kommen. Ein Konzept mit

derart tiefgreifendem Einfluss auf die Netzplanung sollte daher möglichst nicht auf proprietären Funktionen gründen.

- Die globale Optimierung der Allokation von Netzressourcen ist ein nichttriviales Problem. Es ist unklar, wie gut die dafür eingesetzten Algorithmen skalieren.

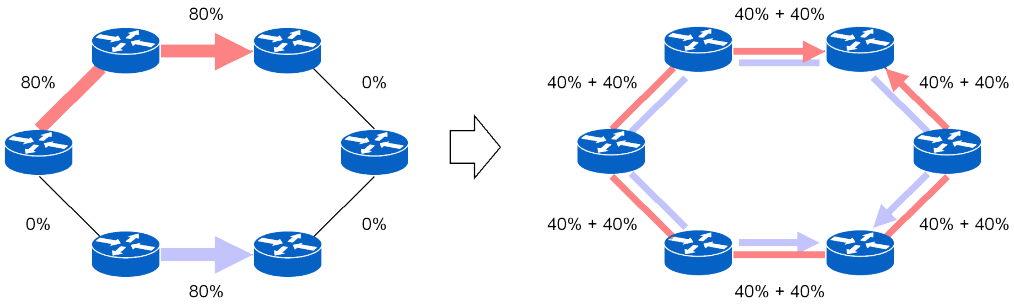

Abbildung 4.3.: Kontraproduktive Wirkung multipler Übertragungspfade

Netzweite manuelle Lastverteilung

Anstatt die TE-Tunnel zwischen allen PE-Routern durch CBR automatisch platzieren zu lassen, könnte auch eine manuelle Einrichtung vorgenommen werden. Um einen geeigneten Verlauf der Pfade entwerfen zu können, wäre es zunächst erforderlich, eine Verkehrsmatrix zu erstellen, d.h. das Die Verkehrscharakteristik und das mittlere Volumen der Datenflüsse zwischen den PE-Routern zu erfassen.

Problematisch dabei ist jedoch, diese Größen nicht konstant bleiben und dass Netz daher in regelmäßigen Zeitabständen – ebenfalls manuell – neu optimiert werden müsste. Dabei wäre jede Iteration vsl. mit einem erheblichen Planungsaufwand verbunden. Kurzfristige und flexible Reaktionen des Netzbetreibers bei Störungen im Netz wären vermutlich schwer zu bewerkstelligen.

Desweiteren gelten einige der o.g. Gegenargumente für die netzweite automatische Lastverteilung auch für diesen Fall. In jedem Fall erscheint ein Lastverteilungsmodell, welches mit einem derartigen Aufwand verbunden wäre, mit Hinblick auf die aktuellen Auslastungswerte im MWN nicht angemessen.

Selektive Lastverteilung

Sinnvoll machbar erscheint derzeit lediglich ein gezielter Einsatz von Lastverteilung für ausgewählte Verkehrsbeziehungen, d.h. die Partitionierung des Datenverkehrs zwischen zwei designierten Netzknoten zur Weiterleitung über multiple Pfade anstatt nur über den Shortest-Path. Ein umfassender Einbezug von dynamischen Parametern und Ressourcenanforderungen scheidet dabei aus. Diese Merkmale könnten nur bei einer netzweiten Lastverteilungsmodell da diese Informationen für den von der Lastverteilung nicht betroffenen Verkehr nicht erfasst werden können.

Wie in den nächsten Abschnitten dargestellt wird, ist das Funktionsprinzip zur Erreichung einer solchen Lastverteilung konventionellen IGP-basierten Methoden ähnlich, ihnen aber durch seine hohe Flexibilität überlegen, da keine Änderungen an den IGP-Kosten bestehender Übertragungsstrecken vorgenommen werden müssen.

Modell

Da LSPs unidirektional sind, kann eine Lastverteilung zwischen zwei Netzknoten aus der Perspektive jedes LSR jeweils getrennt betrachtet werden. Im Folgenden sollen die beiden Netzknoten als LSR A und LSR B bezeichnet werden, wobei o.B.d.A. die Perspektive von LSR A eingenommen wird.

Für jeden zu verwendenden Pfad von A nach B ist mindestens ein TE-LSP auf A zu errichten. Über die Definition von CBR-Constraints lassen sich die Tunnel über die gewünschten Teile des Netzes lenken. Solche Constraints können beispielsweise die Festlegung von zwingend zu durchlaufenden Netzknoten oder umgekehrt der Ausschluss bestimmter Knoten aus dem Pfad sein. Es bestehen mehrere Möglichkeiten, um den Datenverkehr über die LSPs zu leiten:

- Die triviale Methode ist statisches Routing auf A. Sofern die Lastverteilung nur auf einen Teil des Datenverkehrs angewendet werden soll, ist dieses Vorgehen oft unumgänglich; ansonsten sollte es aber aufgrund der mangelnden Flexibilität vermieden werden.

- Soll der gesamte Datenverkehr zwischen von A nach B pauschal einbezogen werden, so ist es am einfachsten, die LSPs auf A in das IGP zu importieren. Die Tunnel werden dann wie konventionelle Übertragungsstrecken im SPF-Algorithmus berücksichtigt. Um den Datenverkehr über die Tunnel zu leiten, weist man ihnen für die TE-Metrik z.B. einen geringeren Kostenwert zu als die IGP-Kosten des Shortest-Path von A nach B.

Eine gleichmäßige Gewichtung des Datenverkehrs zwischen den Pfaden wird erzielt, indem die TE-Metriken der einzelnen Tunnel identisch gewählt werden. Der grundsätzliche Mechanismus ist folglich derselbe wie bei ECMP-Routing. Eine Lastverteilung zwischen Pfaden mit verschiedenen Kosten wird von OSPF nicht unterstützt. Eine ungleichmäßige Gewichtung kann jedoch trotzdem erreicht werden, indem sie auf eine gleichmäßige abgebildet wird, d.h. indem für bestimmte Pfade proportional zusätzliche LSPs mit der gleichen TE-Metrik errichtet werden.

Desweiteren kann auf Cisco-Hardware eine ungleichmäßige Aufteilung erreicht werden, indem die RSVP-Bandbreitenreservierungen der TE-Tunnel in einem entsprechenden Verhältnis gewählt werden [Pep07d]. Die in OSPF importierten TE-Metriken der Tunnel bleiben hierbei zwar identisch, jedoch erfolgt die Aufteilung IOS-intern proportional zu den reservierten Bandbreitenvolumina.

Betrachtet man nun die Situation bidirektional, so besteht die naheliegendste Vorgehensweise zur Aufteilung des Datenverkehrs darin, auf den Pfaden zwischen A und B jeweils einen LSP pro Übertragungsrichtung anzulegen. Da A und B jedoch autonom voneinander agieren können, sind aber auch asymmetrische Varianten denkbar (Abbildung 4.4).

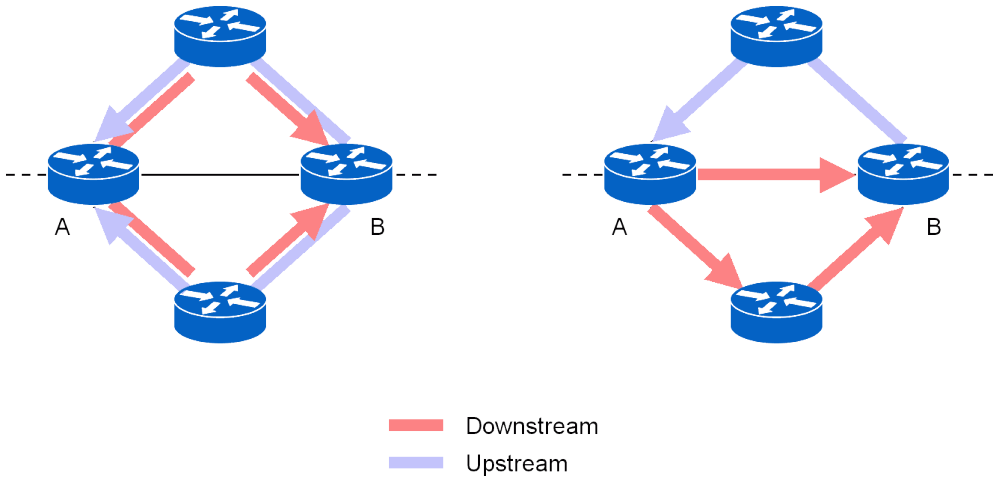

Abbildung 4.4.: Symmetrische (links) und asymmetrische (rechts) LSP-Konstellationen

Kommunikationsablauf

Der Datenverkehr wird am LSP-Ingress nach Flows zwischen den einzelnen LSPs aufgeteilt. Dies geschieht durch Berechnung eines Hashwertes aus der Kombination von Quell- und Zieladresse, über den der LSP zur Weiterleitung ermittelt wird. Zusätzlich fließt in die Berechnung des Wertes ein Hash-Seed ein, der beim Start des Routers zufällig bestimmt wird; die konkrete Aufteilung variiert daher mit jedem Router-Neustart.

Für alle LSPs ist B der LSP-Egress. Von dort aus werden die Pakete auf konventionelle Weise weitergeleitet und sind nicht mehr nach dem Pfad, den sie genommen haben, unterscheidbar.

Von einer Flow-unabhängigen (d.h. paketorientierten) Lastverteilung ist abzuraten. Zwar wäre dies auch möglich, jedoch entstünden dadurch keine Vorteile, da sich beim Vorhandensein einer genügend großen Anzahl von Flows eventuelle Ungleichmäßigkeiten im Verkehrsaufkommen über die LSPs hinweg statistisch ausgleichen. Im Gegenzug würde aber vsl. die Reihenfolge der Pakete bei der Weiterleitung unnötig durcheinander gebracht, was negative Auswirkungen auf die Performance von TCP und vergleichbaren Protokolle zur Folge hätte. Dies wiederum würde mit hoher Wahrscheinlichkeit einen zusätzlichen Jitter induzieren.

Inbetriebnahme und Wartung

Die Konfiguration der TE-LSPs erfolgt jeweils am LSP-Ingress. Im bidirektionalen Fall sind entsprechend A und B, ansonsten nur A von der Konfiguration betroffen. Die CBR-Constraints für das Routing müssen überprüft und ggf. angepasst werden, falls Veränderungen an LSRs vorgenommen werden, die in der Definition der Constraints verwendet werden.

Fehlerverhalten

Fällt von mehreren verfügbaren LSPs in eine Richtung einer aus, so wird dieser nach Erkennung des Ausfalls automatisch umgeleitet, falls ein Alternativpfad verfügbar ist, der die CBR-Constraints dieses LSP erfüllt. Die Zeitdauer bis zur Umleitung hängt von den

Timeout-Werten des IGP ab (sofern der Ausfall nicht durch anderweitige Mechanismen schneller erkannt wird). In der Zwischenzeit geht der Datenverkehr über den defekten Pfad weitergeleitete Datenverkehr verloren.

Kann der LSP nicht wiederhergestellt werden, so verteilt sich der gesamte Verkehr zwischen A und B auf die verbleibenden LSPs. Sind alle LSPs ausgefallen, so wird, sofern über das IGP noch Konnektivität besteht, der Verkehr auf konventionelle Weise weitergeleitet, d.h. direkt über IP bzw. über mittels LDP signalisierte LSPs.

Sofern keine entsprechenden Constraints definiert wurden, kann es auch passieren, dass nach einem Ausfall mehrere an der Lastverteilung beteiligte LSPs über denselben Pfad führen. Der Verkehr verteilt sich dann (ggf. in der vorgesehenen Gewichtung) auf diese LSPs, wird aber über denselben physischen Pfad im Netz übertragen.

Eine problematische Situation kann sich bei asymmetrischen Lastverteilungsmustern ergeben. In bestimmten Konstellationen könnte eine Störung dazu führen, dass bei einem Ausfall der betroffene LSP nicht umgeleitet werden kann, obwohl eigentlich ein Alternativpfad verfügbar wäre. Dies ist genau dann der Fall, wenn alle gemäß den definierten CBR-Constraints zulässigen Pfade in der Topologie ausfallen und eine Rückfall auf konventionelles Routing nicht möglich ist (z.B. wenn die Einleitung des Datenverkehrs in den LSP über eine statische Route erfolgt). Bei der Definition von CBR-Contraints ist daher zu bedenken, dass sie die Topologie einschränken und die Ausfallsicherheit reduzieren können.

4.2.6. Verbesserung der Ausfallsicherheit

Ob MPLS-FRR oder MPLS Path Protection zum Schutz von LSPs im MWN besser geeignet ist, kann nicht pauschal beantwortet werden. Wegen der besseren Skalierbarkeit ist MPLS-FRR i.A. zu bevorzugen. Für eine Ringtopologie kann MPLS Path Protection jedoch deutlich bessere Resultate als MPLS-FRR liefern (Abbildung 4.5). Im Backbone des MWN könnten aufgrund des niedrigen Vermaschungsgrades der Topologie durch MPLS-FRR tendenziell ungünstige Strukturen entstehen.

Die einfachste Variante zur Absicherung des gesamten Datenverkehrs im Backbone mit Backup-LSPs wäre, MPLS-FRR zusammen mit den in Kapitel 2 beschriebenen 1-Hop TE-LSPs einzusetzen. Sonstige TE-LSPs müssten unabhängig davon separat abgesichert werden. Bei diesem Modell wäre jedoch nur der Einsatz von FRR Link Protection praktikabel, da kein vergleichbares Verfahren für FRR Node Protection existiert [MW09].

Nutzenanalyse

Da mit der Einrichtung und Pflege von Backup-LSPs ein zusätzlicher Aufwand verbunden wäre und die Netzstruktur einen neuen Komplexitätsfaktor erhalten würde, erscheint zunächst eine Betrachtung des potentiellen Nutzens angebracht, bevor weitere Planungsschritte unternommen werden. Obgleich die Gewährleistung einer hohen Ausfallsicherheit im MWN zu den wichtigsten Zielen des Netzbetreibers gehört, wird hierbei erkennbar, dass kaum gewichtige Argumente vorliegen, die für die erforderlichen Investitionen sprechen. Teilweise ist dies darin begründet, dass bereits verschiedene andere Redundanzmaßnahmen ergriffen wurden:

- Neben VoIP-Telefonie können derzeit keine Anwendungen identifiziert werden, die substantiell von einem der o.g. Verfahren profitieren würden. Für bestimmte Systeme wie

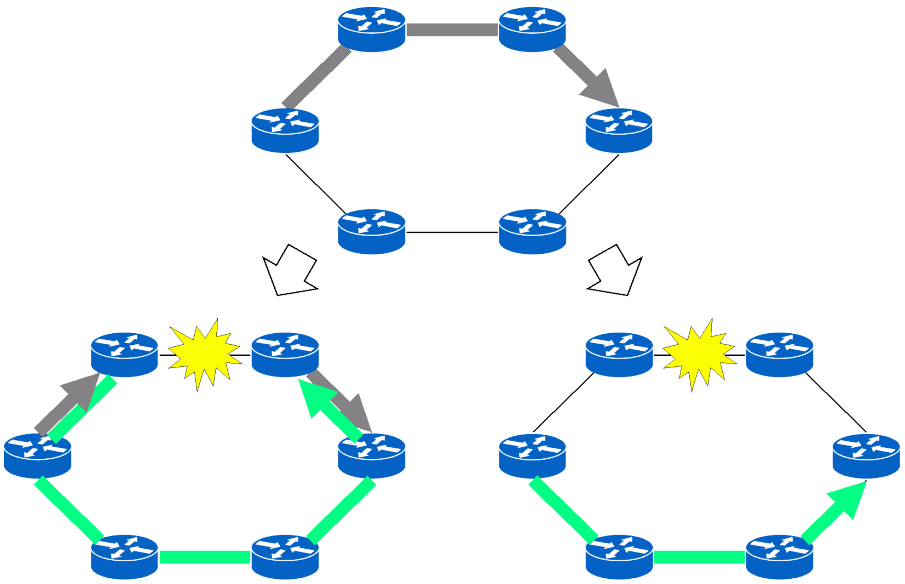

Abbildung 4.5.: MPLS-FRR (links) vs. MPLS Path Protection (rechts)

z.B. den NAT-o-MAT wird als Redundanzmechanismus bereits das Cisco-proprietäre HSRP eingesetzt.

- Ausgehend von vergangenen Beobachtungen [lrz] sind fehlerbedingte Ausfälle im Backbone (d.h. von Hauptverbindungsstrecken und Routersystemen) als selten einzustufen. Nach Einschätzung des Netzbetreibers tritt bei der überwiegenden Mehrheit der Fälle der Ausfall in einem Netzknoten auf. Diese sind, soweit dies mit einem vertretbaren Aufwand machbar ist, bereits redundant ausgelegt. Die Eintrittswahrscheinlichkeit einer Störung kann daher als gering angesehen werden.

- Der Netzbetreiber nimmt an, dass Fehler, die nicht sofort von der Hardware erkannt werden, sehr selten sind. Der worst-case hoher Timeout-Werte des IGP kommt daher kaum zum Tragen, so dass die Unterbrechungsdauern im Regelfall nur von den Konvergenzzeiten abhängen und damit im Bereich weniger Sekunden liegen.

- Zwar stellt VoIP-Telefonie eine geschäftskritische Anwendung dar; eine Unterbrechung von wenigen Sekunden wird vom Netzbetreiber jedoch nicht als bedeutender Schaden eingeschätzt. Diese Einschätzung gilt gleichermaßen für den übrigen Datenverkehr im MWN.

Das Risiko, welches durch den Einsatz von MPLS-FRR oder MPLS Path Protection im MWN eliminiert werden könnte, ist angesichts dessen als verhältnismäßig gering zu bewerten. Unter diesen Umständen erscheint die Anwendung der o.g. Verfahren derzeit eher nicht gerechtfertigt.

	VPLS	Bandbreitenbegrenzung	Verkehrsklassen	Lastverteilung
VPLS		-	W	W
Bandbreitenbegrenzung	A		W	-
Verkehrsklassen	W	W		-
Lastverteilung	W	-	-	

A: Abhängigkeit (gerichtet von Zeile nach Spalte)
W: Wechselwirkung (ungerichtet)

Tabelle 4.3.: Matrix der Abhängigkeiten zwischen den Anwendungsfällen

4.3. Interoperabilität der Anwendungsfälle

Die Anwendungsfälle bzw. die dafür gefundenen Lösungsansätze sind größtenteils unabhängig voneinander (Tabelle 4.3). Unabhängig bedeutet in diesem Kontext, dass es für die Funktionsfähigkeit eines Lösungsansatzes unerheblich ist, ob er separat oder in Kombination mit anderen implementiert wird, und dass bei einer inkrementellen Migration prinzipiell keine Einschränkungen bzgl. der Reihenfolge bestehen. Einzige Ausnahme ist die Bandbreitenbegrenzung, die in diesem Konzept zusammen mit VPLS zur Anwendung kommt.

Allerdings sind zwischen verschiedenen Anwendungsfällen, wenn sie gemeinsam in demselben Netzbereich auftreten, Wechselwirkungen zu erwarten. Wie die Konfigurationen der einzelnen Fälle miteinander zu kombinieren sind, kann an dieser Stelle nicht pauschal erörtert werden, da dies von der jeweiligen konkreten Zielsetzung bei der Netzplanung abhängt.

4.4. Integration in die existierende Infrastruktur

Generell sind bei einer Einführung von MPLS im MWN, abgesehen von einigen erforderlichen Grundeinstellungen, keine erheblichen Anpassungsarbeiten oder gravierenden Nebenwirkungen hinsichtlich der bestehenden Netzstrukturen zu erwarten.

MPLS kann theoretisch parallel mit konventionellem IP-Routing betrieben werden, d.h. es ist zur Nutzung von MPLS nicht zwingend notwendig, die Weiterleitung im Netz komplett darauf umzustellen. Voraussetzung für eine solche Konstellation ist, dass alle LSPs im Netz über MPLS-TE eingerichtet werden und LDP nicht aktiviert wird.

Normalerweise ist LDP als Signalisierungsprotokoll erwünscht, da es die Konfiguration von MPLS deutlich erleichtern kann. LDP signalisiert entlang der aus dem IGP bekannten Routen automatisch LSPs zwischen allen beteiligten Routern, die für verschiedene Dienste und Funktionen von MPLS gemeinsam genutzt werden können. Für viele Einsatzbereiche von MPLS ist es nicht notwendig, die Eigenschaften und/oder Pfade der beteiligten LSPs

genauer zu spezifizieren.

LDP unterstützt folgende zwei Betriebsmodi zur Verteilung von Labels:

- Im Downstream Unsolicited-Modus vergibt ein LSR selbständig Labels für alle ihm bekannten FECs.
- Im Downstream on Demand-Modus vergibt ein LSR erst nach Aufforderung durch einen LDP-Peer ein Label für eine FEC.

Die Cisco-Router im MWN-Backbone unterstützen nur den Downstream Unsolicited-Modus. Sind die LSPs schließlich etabliert, wird automatisch der gesamte Datenverkehr über MPLS weitergeleitet.

Bei einer vollständigen Umstellung der Weiterleitung im MWN auf MPLS wären ohne Vorkehrungen allerdings einige Seiteneffekte zu erwarten. Laut Netzbetreiber existieren im MWN einige Policy-basierte Routen, bei denen die Umleitung des betreffenden Datenverkehrs nicht unmittelbar an dem Eintrittspunkt in das Backbone initiiert wird, sondern der Verkehr erst an einem späteren Punkt seines Weges durch Filter selektiert und von dort aus statisch geroutet wird. Der Grund dafür ist, dass auf diese Weise mehrere Vorkommen derselben Policy aggregiert verarbeitet werden können.

Diese Konstruktion trifft aber auch implizite Annahmen darüber, auf welche Weise ein entsprechendes Datenpaket bis zu diesem Punkt durch das Backbone transportiert wird. Diese Annahmen würden durch den Einsatz von LDP verletzt. Das Problem liegt darin, dass das Paket am genannten Eintrittspunkt in eine FEC klassifiziert würde, die dem Shortest-Path entspricht. Da es danach in einen MPLS-Rahmen verkapselt würde, könnte der auf dem Pfad zwischenliegende Filter – und damit die Policy – nicht wirksam werden. Das Resultat wäre also eine inkorrekte Weiterleitung.

Eine partielle Nutzung von LDP derart, nur in einer Untermenge aller Backbone-Router LDP zu aktivieren, ist faktisch nicht machbar, da dadurch die Konnektivität für MPLS-Funktionen eingeschränkt würde. Daher erscheint es sinnvoll, bei einer Einführung von MPLS zu Beginn auf LDP gänzlich zu verzichten. Dies ist jedoch nicht bei allen Anwendungsfällen praktikabel.

Langfristig empfiehlt sich eine Bereinigung der bestehenden Netzstrukturen von den o.g. Abhängigkeiten. Dies kann einfach dadurch erreicht werden, dass die Filterung und Umleitung des betreffenden Datenverkehrs an den Eintrittspunkt zum Backbone verlegt wird, so dass die FEC die Policy widerspiegelt. MPLS TE-Tunnel bieten sich hier als probates Werkzeug an. Die Konfiguration kann dadurch zwar etwas umfangreicher werden, jedoch stellt dies in jedem Fall eine robustere und flexiblere Lösung dar.

4.5. Zusammenfassung des Konzeptentwurfs

In der Anforderungsanalyse wurden sechs potentielle Anwendungsfälle für MPLS identifiziert. Während der Betrachtungen dieses Kapitels hat sich gezeigt, dass unter den derzeitigen Voraussetzungen im MWN in drei Fällen eine gewinnbringende Anwendung machbar und sinnvoll ist.

Von einer sofortigen Umstellung des MWN-Backbone auf die standardmäßige Weiterleitung mit MPLS sollte abgesehen werden, da hier problematische Auswirkungen zu erwarten wären. Bei einer praktischen Umsetzung der drei Anwendungsfälle sollte daher vorerst

MPLS-TE zur Signalisierung von LSPs verwendet werden. Dadurch kann MPLS während der Übergangszeit auf eine isolierte Weise genutzt werden. Da auch hier das dynamische Routing einbezogen werden kann, sollte dies weder die Robustheit noch die Flexibilität beeinträchtigen. Die derart erstellten Konfigurationen können auch bei einem späteren Einsatz von LDP beibehalten werden.

Der Anwendungsfall zu Traffic Policing ist eng mit dem Anwendungsfall zu VPLS verbunden, so dass diese gemeinsam implementiert werden sollten. Ansonsten bestehen bzgl. der Reihenfolge der Umsetzung keine Einschränkungen.

Hinsichtlich der Komplexität der Implementierung ist anzunehmen, dass ein Anwendungsfall problemlos innerhalb eines Wartungszeitfensters konfiguriert werden kann. Eine Migration im laufenden Betrieb ist eher nicht empfehlenswert.

5. Implementierung

Im praktischen Teil der Arbeit wird nun die konkrete Umsetzung des Konzeptentwurfs auf der im MWN vorhandenen Hardware betrachtet. Dazu wird für jeden Anwendungsfall eine Musterkonfiguration erstellt und deren relevante Einzelheiten erörtert. Weiterhin werden auch Aspekte der Migration betrachtet wie z.B. die Vorgehensweise bei der Konfiguration, geeignete Kontroll- und Testmöglichkeiten sowie die Werkzeuge zur Fehleranalyse.

5.1. Testumgebung

Das LRZ besitzt ein Hardware-Labor, in dem zwei Cisco Catalyst 6509-Geräte mit kompletter Ausstattung für Testzwecke verfügbar sind. Daneben sind noch eine Reihe weiterer Geräte vorhanden, die jedoch nicht MPLS-fähig sind. Bei den Vorüberlegungen zur Implementierung ist schnell klar geworden, dass zwei MPLS-Router nicht ausreichend sind, um auf realistische Weise verschiedene technische Verfahren zu testen bzw. zu beobachten. Zur Abbildung von sinnvollen Topologien sind wenigstens drei oder besser vier Netzknoten erforderlich. Da keine Abhilfe mit zusätzlichen Geräten geschaffen werden kann, wird auf die Durchführung von Tests im Labor verzichtet. Stattdessen wird zur Überprüfung der praktischen Funktionsfähigkeit von Konfigurationen der Router-Emulator Dynamips zusammen mit dem Frontend Graphical Network Simulator (GNS3) herangezogen.

Dynamips unterscheidet sich von anderen gebräuchlichen Emulatoren dadurch, dass es nicht ein IOS CLI nachbildet, sondern stattdessen eine echte Version von Cisco IOS auf einer virtuellen Maschine (VM) ausführt. Ein klarer Vorteil dieses Ansatzes ist, dass die CLI ein identisches Verhalten wie auf der tatsächlichen Hardware zeigt und alle Eigenschaften der betreffenden IOS-Version im Emulator verfügbar sind. Neben verschiedenen Typen von Cisco-Routern kann Dynamips weitere Netzkomponenten wie Firewalls, Ethernet- und ATM-Switches emulieren.

Zu Simulation von Hostrechnern wird das Tool Virtual PC Simulator (VPCS) verwendet. VPCS kann bis zu 9 Hostrechner simulieren, von denen aus allerdings lediglich einige ICMP-Funktionen ausgeführt werden können (`ping` und `tracert`). Es kann entsprechend nur für Tests der Konnektivität und des Routings verwendet werden; dies ist allerdings hier vollkommen ausreichend.

Zur Analyse des Datenverkehrs zwischen den virtuellen Routern wird der Protokollanalysator Wireshark verwendet. Alle diese Programme sind unter kostenfreien Lizenzen nutzbar und können über das WWW bezogen werden.

5.1.1. Einschränkungen

In der emulierten Testumgebung können zwar hinreichend komplexe Topologien nachgebildet werden, jedoch ergeben sich in anderer Hinsicht einige Beschränkungen:

- Dynamips unterstützt derzeit nur bestimmte Router-Serien, zu denen keine MLS wie der Catalyst 6509 gehören. Im Testszenario werden daher emulierte Cisco 7200-Router

verwendet. Die dabei eingesetzte IOS-Software (*12.2(33)SRC2 Advanced IP Services*) ist der im MWN eingesetzten Version vom Funktionsumfang her zwar relativ ähnlich, aber es bestehen dennoch einige signifikante Inkompatibilitäten. Die im MWN eingesetzte Version ist auf diesen Routern nicht lauffähig.

- Die Emulation von Routern und Hosts ist ein enorm rechen- und speicherintensiver Prozess, der einen durchschnittlich modernen Arbeitsplatzrechner ohne weiteres vollständig auslasten kann. Obgleich durchaus eine brauchbare Anzahl von Geräten parallel emuliert werden kann, muss das Testszenario daher so simpel wie möglich gestaltet werden. Das Hinzufügen bzw. Entfernen von Prozessen auf den virtuellen Routern hat einen deutlichen Einfluss auf die Performanz der Emulation.

- Zeitabhängige Netzeigenschaften können in dem emulierten Netz nicht auf realistische Weise untersucht werden. Beispielsweise ist der Paketdurchsatz zwischen den Netzknoten vergleichsweise niedrig und – ebenso wie die Round-Trip-Time (RTT) – starken Schwankungen unterworfen.

5.1.2. Testszenario

Als Testszenario für die einzelnen Anwendungsfälle wurde eine einfache teilvermaschte Netzstruktur in GNS3 erzeugt (Abbildung 5.1). Die Topologie wurde dabei zwar identisch wie das MWN-Backbone gewählt, ansonsten haben die Netzstrukturen allerdings recht wenig gemeinsam, da die Konfigurationen der virtuellen Router vollständig neu erstellt wurden. Dies war zum einen sinnvoll, um das Testszenario möglichst einfach zu halten; als wesentlich zwingender erwies sich jedoch der Umstand, dass sich die Netzstruktur des MWN-Backbone aufgrund der o.g. Inkompatibilitäten auch in vereinfachter Form nicht analog nachbilden lässt.[1]

Über das Netz wurden insgesamt 9 VPCS-Hostrechner verteilt, die größtenteils analog wie im MWN jeweils einem VLAN zugeordnet sind. Konkrete Adressen und Zahlenwerte der im Folgenden betrachteten Konfigurationsausschnitte beziehen sich auf dieses Szenario.

5.2. Grundeinstellungen von MPLS

Zu Beginn der Einrichtung von MPLS sind unabhängig von den Anwendungsfällen einige grundlegende Konfigurationsschritte durchzuführen. Auf Cisco-Geräten muss zur Nutzung von MPLS zunächst das sog. Cisco Express Forwarding (CEF), eine proprietäre Schnittstellenfunktion, eingeschaltet werden. Auf den Routern im MWN-Backbone ist dies die Standardeinstellung. Anschließend kann die globale MPLS-Funktionalität des Routers aktiviert werden.

```
Router{config}# ip cef
Router{config}# mpls ip
```

[1] Die physischen Schnittstellen im Backbone sind alle ausnahmslos als sog. Switchports konfiguriert; dieser Modus wird aber nur von MLS, nicht jedoch von konventionellen Routern wie dem Cisco 7200 unterstützt.

Dieser Schritt ist auf allen Routern durchzuführen, auf denen MPLS genutzt werden soll. Anschließend ist auf ihnen mindestens ein Signalisierungsprotokoll für LSPs zu aktivieren, d.h. entweder LDP, RSVP-TE oder auch beide.

5.2.1. Basiskonfiguration LDP

Um die Verteilung von Labels mit LDP zu bewirken, muss das Protokoll zunächst gestartet werden. Dabei muss der Identifikator des Routers für die LDP-Nachrichten spezifiziert werden, wobei hierfür üblicherweise eine Loopback-Schnittstelle herangezogen wird (diese muss nicht separat definiert werden; beispielsweise kann dieselbe Schnittstelle wie für den OSPF-Prozess verwendet werden). Schließlich muss auf den gewünschten Schnittstellen die standardmäßige Paketweiterleitung mittels MPLS aktiviert werden. Die Verteilung von Labels mittels LDP über diese Schnittstellen hängt unmittelbar damit zusammen.

```
Router{config}# mpls label protocol ldp
Router{config}# mpls ldp router-id Loopback0

Router{config-if}# mpls ip
```

LDP sollte nur auf den Kernnetz-internen Schnittstellen aktiviert werden. Der Betrieb von LDP auf Schnittstellen, an denen Zugangsnetze angeschlossen sind, stellt ein Sicherheitsrisiko dar [Pep07c]. Dies ist im Zusammenhang mit LDP Autoconfiguration zu bedenken, einer Konfigurationsoption, die LDP automatisch global auf allen Schnittstellen eines Routers aktiviert. Diese Funktion ist nur auf P-Routern problemlos einsetzbar; da im MWN jedoch alle Backbone-Router Zugangsnetze bedienen, sollte folglich darauf eher verzichtet werden.

Da die physischen Schnittstellen im MWN-Backbone als Switchports deklariert sind, muss LDP dort ggf. auf den entsprechenden VLAN-Schnittstellen der Router-zu-Router-Kommunikation aktiviert werden.

Die Konfiguration sollte schließlich überprüft werden. Am wichtigsten ist hierbei die MPLS Forwarding-Tabelle, die die Zuordnungen zwischen FECs und Labels enthält.

```
Router> show mpls interfaces
Router> show mpls ldp neighbor
Router> show mpls forwarding-table
```

5.2.2. Basiskonfiguration MPLS-TE

Um MPLS-TE auf einem Router nutzen zu können, muss dieses sowohl global aktiviert werden als auch auf jeder Schnittstelle, die von einem TE-Tunnel durchlaufen werden darf. Zusätzlich muss auf jeder betroffenen Schnittstelle RSVP-TE aktiviert werden; dabei sind die für Reservierungen zulässigen maximalen Bandbreiten zu definieren [Lak05] (global für alle Tunnel und maximale je Tunnel, im Beispiel jeweils 100000 kbps).

Weiterhin müssen die TE-Erweiterungen des Routing-Protokolls aktiviert werden, um CBR zu ermöglichen. Analog wie bei LDP muss dabei der Identifikator des Routers für die TE-Signalisierung spezifiziert werden, wofür i.A. ebenfalls eine Loopback-Schnittstelle verwendet wird. Im Fall von OSPF muss zudem die Area angegeben werden, in der MPLS-TE eingesetzt wird.

```
Router{config}# mpls traffic-eng tunnels

Router{config-if}# mpls traffic-eng tunnels
Router{config-if}# ip rsvp bandwidth 100000 100000

Router{config-router}# mpls traffic-eng router-id Loopback0
Router{config-router}# mpls traffic-eng area 0
```

Wie bereits bei LDP kann MPLS-TE im MWN-Backbone nicht auf den physischen Schnittstellen aktiviert werden, sondern muss auf den entsprechenden VLAN-Schnittstellen konfiguriert werden.

5.3. Musterkonfigurationen

Aufgrund der Einschränkungen der Testumgebung konnten die Musterkonfigurationen für die Anwendungsfälle nicht hinsichtlich ihrer Kompatibilität mit dem Catalyst 6509 getestet werden. Einige Konfigurationsteile wurden mangels Alternative ausschließlich aus der Herstellerdokumentation von Cisco abgeleitet. Eventuell vorhandene Unterschiede beschränken sich aber vsl. lediglich auf Abweichungen in der Syntax.

5.3.1. VPLS

Das hier als Beispiel erstellte VPN verbindet im Testszenario die Hostrechner VPCS1, VPCS4 und VPCS8. Der Konfigurationsausschnitt bezieht sich auf den virtuellen Router `csr1-0q1`.

Zur Erstellung des VPN ist zunächst die VFI zu definieren. Dabei muss ihr eine netzweit gültige VPN-ID (hier: 123) zugewiesen werden. Auf jedem PE-Router müssen die Adressen der anderen PE-Router eingegeben werden. Bei diesen Adressen handelt es sich jeweils um deren Router-IDs, die bei ihrer Konfiguration von LDP bzw. MPLS-TE angegeben wurden (d.h. die Adressen der Loopback-Schnittstellen). Anschließend ist die VFI an die Schnittstelle zu binden, an der das VPN-Teilnetz angeschlossen ist.

```
Router{config}# l2 vfi MY_VPN manual
Router{config-vfi}# vpn id 123
Router{config-vfi}# neighbor 10.0.1.5 encapsulation mpls
Router{config-vfi}# neighbor 10.0.1.6 encapsulation mpls

Router{config}# interface vlan101
Router{config-if}# xconnect vfi MY_VPN
```

Diese Konfiguration ist auf jedem PE-Router unter Angabe derselben VPN-ID zu wiederholen. Danach sollte die Konnektivität der PWE3-Verbindungen überprüft werden.

```
Router> show mpls l2transport vc
```

Mit wachsender Anzahl von VPN-Teilnetzen wird die Konfiguration zunehmend aufwendiger. Sie kann daher auch zentralisiert werden, indem einer der PE-Router als “Master” designiert wird. Auf ihm werden dann die Adressen aller anderen PE-Router eingegeben.

Der Master kann dann über die auf BGP basierende Funktion “VPLS Autodiscovery” die Adressen an alle anderen PE-Router verteilen. Da der Master aber einen SPoF darstellt, sollte eine Zentralisierung, wenn möglich, eher vermieden werden.

Die gezeigte Konfiguration basiert auf mittels LDP signalisierten LSPs. Soll MPLS-TE verwendet werden, müssen für jeden benachbarten PE-Router sowohl die PWE3-Verbindung als auch der zugehörige TE-Tunnel manuell konfiguriert werden. Für den Tunnel ist dabei die zu reservierende Bandbreite sowie die Art der Wegewahl anzugeben.

```
Router{config}# pseudowire-class MY_VPN_VC1
Router{config-pw-class}# encapsulation mpls
Router{config-pw-class}# preferred-path interface Tunnel0

Router{config}# interface Tunnel0
Router{config-if}# ip unnumbered Loopback0
Router{config-if}# tunnel destination 10.0.1.5
Router{config-if}# tunnel mode mpls traffic-eng
Router{config-if}# tunnel mpls traffic-eng bandwidth 10000
Router{config-if}# tunnel mpls traffic-eng path-option 1 dynamic
```

Darüber hinaus muss die VFI-Konfiguration angepasst werden. Dabei ist der PWE3-Verbindung eine lokal eindeutige VC-ID zuzuweisen (hier: 345).

```
Router{config-vfi}# neighbor 10.0.1.5 345 pw-class MY_VPN_VC1
```

Die Konfiguration von PWE3-Verbindungen und TE-Tunneln ist für alle weiteren PE-Router entsprechend zu wiederholen.

5.3.2. Traffic Policing

Das Traffic Policing kann für jedes VPN-Teilnetz eigenständig konfiguriert werden. Das folgende Beispiel beschränkt die nutzbare Bandbreite für VPCS1. Der Konfigurationsausschnitt bezieht sich auf den virtuellen Router `csr1-0q1`.

Zunächst ist eine entsprechende Policy zu erstellen, wobei die mittlere Datenübertragungsrate, das zulässige maximale Burst-Volumen und die Aktionen für Konformität bzw. Nichtkonformität des Datenverkehrs zu definieren sind.

```
Router{config}# policy-map VPN_LIMIT
Router{config-pmap}# class class-default
Router{config-pmap-c}# police 1000000 50000 conform-action transmit exceed-action drop
```

Anschließend ist die Policy an die Schnittstelle zu binden, an der das VPN-Teilnetz angeschlossen ist.

```
Router{config}# interface vlan101
Router{config-if}# service-policy input VPN_LIMIT
```

5.3.3. Lastverteilung

Die Übertragungsstrecke zwischen den Routern `csr1-kb1` und `csr1-2wr` gehört zu den höher ausgelasteten Strecken im MWN-Backbone. Hier wird auch in Zukunft vom Netzbetreiber noch eine deutliche Zunahme der Auslastung erwartet. Daher wird diese Strecke im Testszenario als Demonstrationsbeispiel herangezogen. Der Konfigurationsausschnitt bezieht sich auf den virtuellen Router `csr1-kb1`.

Da das Prinzip unabhängig von der Anzahl der verwendeten LSPs ist, werden für das Beispiel der Einfachheit wegen zwei herangezogen. LSP 1 wird dabei über den Shortest-Path und LSP 2 über eine durch CBR zu bestimmende Route geführt. Für jeden Pfad wird ein TE-Tunnel konfiguriert.

```
Router{config}# interface Tunnel0
Router{config-if}# ip unnumbered Loopback0
Router{config-if}# tunnel destination 10.0.1.1
Router{config-if}# tunnel mode mpls traffic-eng
Router{config-if}# tunnel mpls traffic-eng autoroute announce
Router{config-if}# tunnel mpls traffic-eng bandwidth 10000
Router{config-if}# tunnel mpls traffic-eng path-option 1 dynamic

Router{config}# interface Tunnel1
Router{config-if}# ip unnumbered Loopback0
Router{config-if}# tunnel destination 10.0.1.1
Router{config-if}# tunnel mode mpls traffic-eng
Router{config-if}# tunnel mpls traffic-eng autoroute announce
Router{config-if}# tunnel mpls traffic-eng bandwidth 5000
Router{config-if}# tunnel mpls traffic-eng path-option 1 explicit name ALT_LSP
Router{config-if}# tunnel mpls traffic-eng path-option 2 dynamic
```

Über die Funktion “MPLS Autoroute Announce” werden die TE-Tunnel in das IGP importiert. Die TE-Metrik wird dabei auf Cisco-Routern standardmäßig auf den Kostenwert des Shortest-Path zwischen den Tunnelendpunkten festgelegt, so dass keine weiteren Anpassungen notwendig sind, damit die Tunnel zur Weiterleitung von IGP verwendet werden.

Da für den zweiten Tunnel nur die halbe Bandbreite reserviert wurde, wird der Datenverkehr zwischen dem ersten und dem zweiten Tunnel im Verhältnis 2 : 1 aufgeteilt. Für den zweiten Tunnel wird außerdem ein CBR-Contraint definiert, so dass der Shortest-Path nicht von ihm verwendet werden kann. Sein Pfad führt daher über den virtuellen Router `csr1-kw5`.

```
Router{config}# ip explicit-path name ALT_LSP enable
Router{config-expl-path}# exclude-address 10.0.0.9
```

Schließlich sollte verifiziert werden, dass beide Pfade auch wie gewünscht in die Weiterleitung einbezogen werden.

```
Router> show ip route 10.0.1.1
Router> show ip cef 10.0.1.1
Router> show ip cef 10.0.1.1 internal
```

Der gesamte Vorgang ist für den virtuellen Router `csr1-2wr` zu wiederholen.

5.4. Troubleshooting

Zwei nützliche Werkzeuge zur Fehlersuche in LSPs sind die Funktionen MPLS Ping und MPLS Traceroute. Diese Befehle sind beide unter IOS verfügbar. Der Unterschied gegenüber ihren ICMP-Pendants ist, dass sie nur über LSPs hinweg funktionieren. Dies kann nützlich sein, da die konventionellen Varianten von Ping und Traceroute im Zusammenhang mit MPLS oft keine eindeutigen Schlüsse über die verwendeten Routen bzw. LSPs zulassen.

```
Router> ping mpls ipv4 10.0.0.24/30
Router> traceroute mpls ipv4 10.0.0.24/30
```

Beide Werkzeuge arbeiten mit erweiterten MPLS-spezifischen Fehlercodes. MPLS Traceroute liefert neben dem Verlauf eines LSP auch die unterwegs verwendeten Labels zurück.

Zu beachten ist, dass die IP-Adressen im Beispiel FECs darstellen, weshalb die Angabe der entsprechenden Netzmaske erforderlich ist. Neben FECs können auch PWE3-Verbindungen oder TE-Tunnel als Ziel angegeben werden.

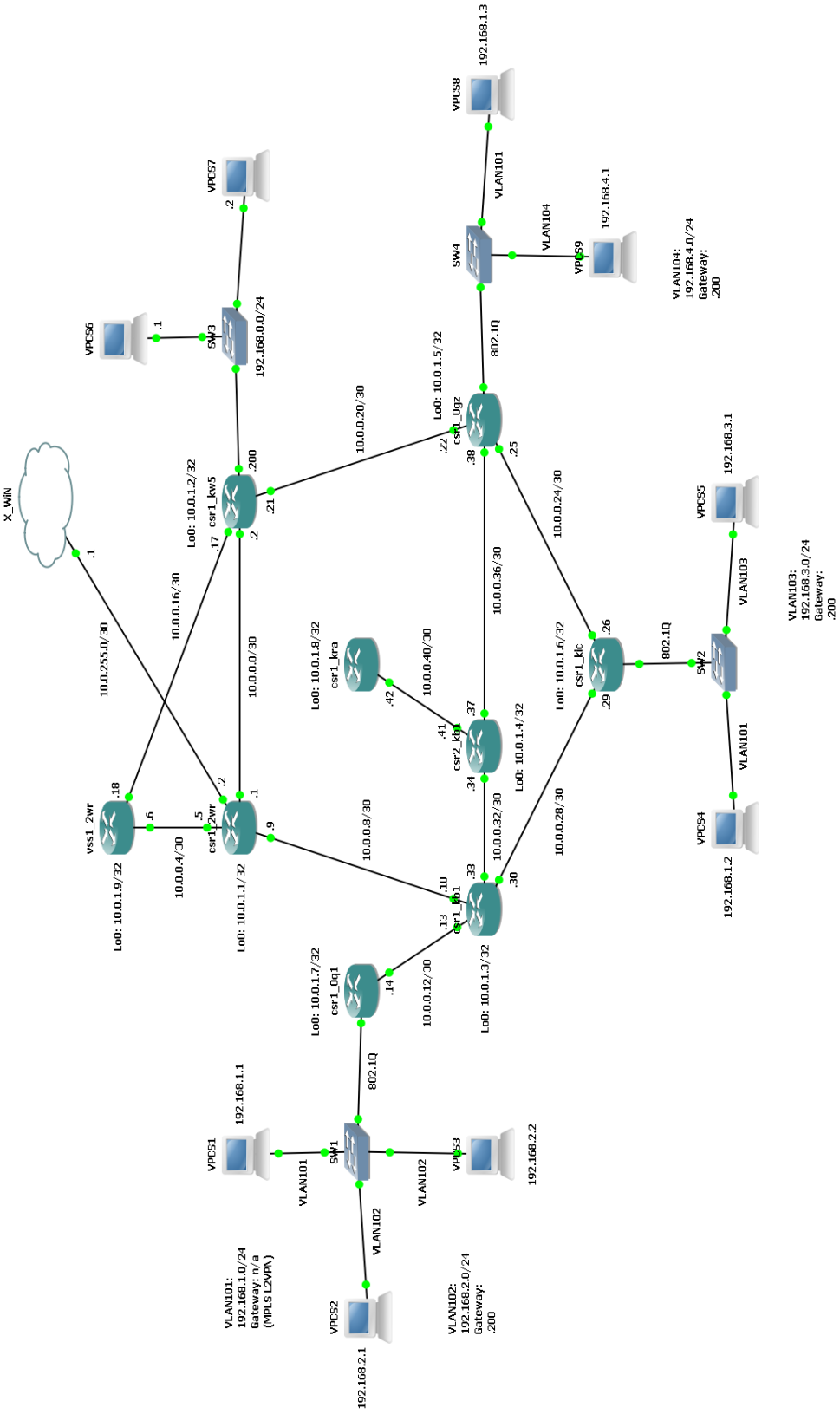

Abbildung 5.1.: Testszenario aus GNS3/Dynamips

6. Zusammenfassung

In dieser Diplomarbeit wurde untersucht, wie MPLS im Münchner Wissenschaftsnetz auf sinnvolle Weise eingesetzt werden kann. Dazu wurden das Netz und sein Umfeld analysiert und mehrere mögliche Anwendungsbereiche ermittelt. In der Planungsphase der Arbeit wurden diese Anwendungsbereiche eingehend untersucht und verschiedene Ansätze zur Umsetzung betrachtet. Einige der Anwendungsbereiche wurden schließlich in einer Testumgebung realisiert.

6.1. Evaluation

Die wesentlichen Ziele der Arbeit, die Erforschung von Einsatzmöglichkeiten für MPLS im MWN und der Entwurf eines Einführungskonzepts, wurden erreicht. Einige der ursprünglichen Problemstellungen konnten durch MPLS nicht gelöst werden. Insbesondere eine Vereinfachung des Policy-basierten Routing wäre für den Netzbetreiber eine Erleichterung gewesen.

Im praktischen Teil der Arbeit mussten aufgrund eingeschränkter Testmöglichkeiten einige Abstriche in der Überprüfung von Konfigurationsoptionen gemacht werden. Dennoch konnten alle gewünschten Anwendungsfälle prototypisch umgesetzt werden.

Eine wesentliche Schwierigkeit der Arbeit bestand darin, die komplexe Infrastruktur des MWN zu erfassen und die relevanten Aspekte in einer kompakten Form herauszukristallisieren. Desweiteren decken die verschiedenen Anwendungsfälle – ebenso wie die Technologie MPLS selbst – ein breites thematisches Feld ab, was in allen Phasen der Arbeit eine große Herausforderung war.

6.2. Ausblick

Es ist anzunehmen, dass Traffic Engineering in der Netzplanung mit dem weiteren Wachstum des MWN an Bedeutung zunehmen wird. Die Fähigkeit, auf einfache Weise über die physischen Verbindungsstrecken hinweg eine Overlay-Topologie errichten und sich dabei gleichzeitig der Robustheit dynamischer Routing-Protokolle bedienen zu können, macht MPLS zu einem mächtigen Werkzeug, das in der zukünftigen Netzplanung im MWN eine Schlüsselfunktion einnehmen kann.

A. Anmerkungen zum Testszenario

In den Testkonfigurationen tauchen einige von den Standardwerten abweichende Einstellungen auf, die nicht mit MPLS zusammenhängen oder in irgendeiner Weise für die Anwendungsfälle bedeutsam sind. Im Wesentlichen dienen diese dazu, eine zusätzliche Rechenbelastung bei der Emulation der Testumgebung zu vermeiden, indem das Aufkommen an Management-Datenverkehr explizit verringert wird:

- Deaktivierung von Ethernet-Keepalives,
- Deaktivierung von CDP,
- Verlängerung des Hello-Intervalls von OSPF, und
- Verlängerung des Hello-Intervalls von LDP.

Folgende Testkonfigurationen wurden hinsichtlich Konnektivität, Routing- und Fehlerverhalten untersucht:

- EoMPLS, jeweils auf Basis von LDP und TE (VPLS wird auf Cisco 7200-Routern nicht unterstützt)
- Lastverteilung, jeweils mit LDP aktiv und inaktiv

Glossar

Application Specific Integrated Circuit (ASIC)

Ein Hardwareschaltkreis, der speziell zur Unterstützung einer bestimmten Anwendung entwickelt wurde. (Nicht zu verwechseln mit dem Begriff Astronomical Scale Integrated Circuit, der ebenfalls mit ASIC abgekürzt wird.)

Bidirectional Forwarding Detection

BFD ist ein von Cisco und Juniper entwickeltes Protokoll, um die Erkennung von Ausfällen von Netzelementen zu beschleunigen. Es ist insbesondere in Kombination mit Protokollen der Sicherungsschicht interessant, die keine eigenen Mechanismen hierfür bereitstellen (z.B. Ethernet) und daher normalerweise auf die Ausfallerkennung durch höhere Schichten (z.B. das IP-Routing) angewiesen sind, die jedoch für QoS-sensiblen Datenverkehr i.A. zu langsam reagieren. BFD ist prinzipiell nichts anderes als ein einfaches Hello-Protokoll.

Control Plane Policing (CoPP)

Unter dem Begriff CoPP werden Sicherheitsmaßnahmen zusammengefasst, die den Datenverkehr zur Control/Management Plane eines Routers restringieren (d.h. Nachrichten, die vom Router selbst zu verarbeiten sind). Damit soll sichergestellt werden, dass der Routing-Prozessor nicht überlastet wird (z.B. durch einen DoS-Angriff), was die Stabilität des Netzes beeinträchtigen könnte.

Differentiated Services Code Point (DSCP)

Der DSCP ist eine Datenstruktur zur Markierung von Paketen, die dazu dient, ein Paket während des Transports in eine Verkehrsklasse und Priorität einordnen zu können. In einem verbindungslosen Netz wird den einzelnen Knoten dadurch eine differenzierte Behandlung des Datenverkehrs ermöglicht. Bei IP DiffServ wird der DSCP im TOS-Feld des IP-Headers mitgeführt.

Dynamic Trunking Protocol (DTP)

DTP ist ein proprietäres Layer 2-Protokoll von Cisco zur automatischen Konfiguration von VLAN-Trunking auf einem Link zwischen zwei Switches (Auto-Negotiation). Das Protokoll regelt, ob ein Link zu einem VLAN-Trunk wird und welche Methode bzw. welches Protokoll zur Enkapsulierung ggf. verwendet wird (z.B. IEEE 802.1Q oder – das ebenfalls Cisco-proprietäre – Inter-Switch Link, ISL).

Equal Cost Multi Path (ECMP)

In einem verbindungslosen Netz mit einer vermaschten Topologie können zwischen einer Quelle und einer Senke mehrere Pfade mit den gleichen Kosten bestehen. Anstatt

willkürlich einen dieser Pfade als den "besten" festzulegen und alle Pakete darüber weiterzuleiten, kann der Verkehr gleichmäßig über die verschiedenen Pfade aufgeteilt werden. Diese Strategie wird als ECMP-Routing bezeichnet und stellt ein einfaches Verfahren zur Erreichung einer Lastverteilung dar.

Hot Standby Router Protocol (HSRP)

HSRP ist ein proprietäres Protokoll von Cisco zur Unterstützung von Hochverfügbarkeitsnetzen. HSRP ermöglicht die Kopplung mehrerer physischer Router zur einem Verbund, der von außen betrachtet wie ein einzelner virtueller Router erscheint. Einer der physischen Router dient dabei als sog. aktiver Router, während ein weiterer als sog. Standby-Router fungiert, auf den bei einer Störung schnell umgeschaltet werden kann. Der virtuelle Router besitzt eigene IP- und MAC-Adressen, die auf den Hosts als die des Standard-Gateway konfiguriert werden. Hieraus ergibt sich der Nutzen von HSRP, da der Ausfall eines physischen Routers somit im Idealfall von den Hosts unbemerkt bleibt.

Interior Gateway Protocol (IGP)

Der Begriff IGP bezeichnet Routing-Protokolle, die zum Routing innerhalb eines Autonomen Systems eingesetzt werden.

Random Early Detection (RED)

RED ist ein Verfahren zur Vermeidung von Überlast im Zusammenhang mit Warteschlangen. Die triviale Vorgehensweise bei Überlauf einer Warteschlange ist das Verwerfen von Elementen, die nicht mehr im Puffer gespeichert werden können (Tail Drop). Demgegenüber agiert RED proaktiv, indem ein neu einzureihendes Element mit einer bestimmten statistischen Wahrscheinlichkeit verworfen wird, auch wenn der Puffer noch nicht voll ist. Der Wahrscheinlichkeitswert berechnet sich z.B. aus der aktuellen Länge der Warteschlange im Verhältnis zu ihrer Gesamtkapazität. Ist die Warteschlange fast leer, so ist der Wert nahe 0; ist die Warteschlange fast voll, so ist der Wert nahe 1. RED erhöht die Fairness gegenüber Tail Drop, da die Burst-Charakteristik beim Verwerfen von Elementen verloren geht. Zwei bekannte Varianten des Verfahrens sind Weighted RED (WRED, Erweiterung auf mehrere Warteschlangen unterschiedlicher Priorität durch Gewichtung) und Adaptive RED (ARED, Erweiterung um feedbackgesteuerte Mechanismen, um die Aggressivität des Verfahrens bei niedriger Auslastung der Warteschlange zu verringern).

Resource Reservation Protocol (RSVP)

Layer 4-Protokoll zur Reservierung von Ressourcen in einem IP-Netz, um bestimmte Qualitätsparameter für eine Datenübertragung sicherzustellen (z.B. minimal verfügbare Bandbreite). Wurde in Vergangenheit meist im Zusammenhang mit IP IntServ eingesetzt.

Service Level Agreement (SLA)

Ein SLA ist eine Vereinbarung zwischen einem Dienstleister und einem Dienstnutzer über spezifische Dienstgütekriterien, die ein zu erbringender Dienst erfüllen soll.

Shortest Path First (SPF)

SPF bezeichnet die Familie von Routing-Verfahren, die aus den verfügbaren Wegen von Quelle zu Senke jeweils den kürzesten zur Weiterleitung auswählen (bzw. den Weg mit den geringsten Kosten); dieser wird i.A. mit dem Dijkstra-Algorithmus berechnet.

Split Horizon

Split Horizon bezeichnet ein Prinzip zur Vermeidung von Schleifen in einer Netztopologie. Das Prinzip fordert, dass ein Knoten ein Datenpaket niemals über die Schnittstelle weiterleitet, über die er es zuvor empfangen hat. (Sofern die Topologie Zyklen enthält, sind weitere Maßnahmen notwendig, um Schleifenfreiheit zu garantieren.)

Traffic Policing

Traffic Policing ist ein Verfahren zur Begrenzung der nutzbaren Bandbreite auf einem Übertragungskanal. Dazu sind der mittlere Datendurchsatz und das maximal zulässige Burst-Volumen zu definieren. Solange diese Werte eingehalten sind, werden zu sendende Pakete schnellstmöglich übertragen; nichtkonforme Pakete werden i.A. verworfen. Die ursprüngliche Verzögerungscharakteristik des Datenverkehrs bleibt erhalten. ("Token-Bucket"-Algorithmus)

Traffic Shaping

Traffic Shaping ist ein Verfahren zur Begrenzung der nutzbaren Bandbreite auf einem Übertragungskanal. Dazu sind der Datendurchsatz und die maximal zu verwendende Puffergröße zu definieren. Zu sendende Pakete werden gepuffert, solange freier Speicher vorhanden ist und aus dem Puffer mit konstanter Rate übertragen. Nichtkonforme Pakete werden verworfen. Die ursprüngliche Verzögerungscharakteristik des Datenverkehrs geht verloren. ("Leaky-Bucket"-Algorithmus)

Type of Service (TOS)

Ein Feld aus dem IP-Header zur Angabe einer Verkehrsklasse. Wird heute nicht mehr nach seiner ursprünglichen Definition verwendet und bei der Paketweiterleitung im öffentlichen Internet üblicherweise ignoriert.

Unicast Reverse Path Forwarding (RPF)

Unicast RPF ist ein optionaler sicherheitsrelevanter Mechanismus in modernen Routern zur Erschwerung von IP-Spoofing und der damit verbundenen Angriffsmethoden. Pakete mit offensichtlich gefälschten Absenderadressen werden nicht weitergeleitet, wenn der Mechanismus aktiviert ist. Die Robustheit des Netzes kann so verbessert werden. Eine Absenderadresse kann beispielsweise als gefälscht eingestuft werden, wenn diese Adresse gemäß Routingtabelle nicht über die Schnittstelle des Routers erreichbar ist, über die das Paket empfangen wurde. (Verschiedene andere Restriktionsgrade sind möglich.)

VLAN Trunking Protocol (VTP)

VTP ist ein proprietäres Layer 2-Protokoll von Cisco zur automatischen netzweiten Konfiguration von VLAN-Trunks. Zunächst sind dazu eine oder mehrere VTP-Domänen zu definieren. Als VTP-Domäne wird eine Gruppe von Switches bezeichnet,

die VLAN-Informationen untereinander austauschen. Hierbei übernimmt (mindestens) einer der Switches die Rolle des VTP-Servers, der eine VLAN-Datenbank verwaltet und über den alle VLANs innerhalb der VTP-Domäne zentral verwaltet werden können. Die VLAN-Datenbank wird vom VTP-Server an die übrigen Switches in der VTP-Domäne verteilt, so dass diese ihre Trunk-Schnittstellen entsprechend konfigurieren können. Um zu vermeiden, dass der Broadcast-Verkehr aus einem VLAN auch an Switches gesendet wird, die keine Access-Schnittstelle in diesem VLAN besitzen, kann das sog. VTP Pruning aktiviert werden.

Weighted Fair Queueing (WFQ)

WFQ ist ein Verfahren zur differenzierten Verarbeitung mehrerer Warteschlangen mit unterschiedlichen Prioritäten. Die Warteschlangen werden gewichtet nach dem Round-Robin-Prinzip bearbeitet, d.h. jede Warteschlange erhält für ihren Durchgang eine bestimmte Menge an Ressourcen (Rechenzeit, Datenübertragungsrate, etc.), die ihrer Priorität entspricht.

Abkürzungsverzeichnis

ACL	Access Control List
AS	Autonomous System
ATM	Asynchronous Transfer Mode
AToM	Any Transport over MPLS
BGP	Border Gateway Protocol
CatOS	Catalyst OS
CBR	Constrained Based Routing
CDP	Cisco Discovery Protocol
CE	Customer Edge
CEF	Cisco Express Forwarding
CLI	Command Line Interface
CoS	Class of Service
CR-LDP	Constraint Based LDP
CSPF	Constrained SPF
DDOS	Distributed DOS
DFN/X-WiN	Deutsches Forschungsnetz
DiffServ	Differentiated Services
DOS	Denial-of-Service
DSL	Digital Subscriber Line
E-LSP	Exp-Inferred-PSC LSP
EoMPLS	Ethernet over MPLS
Exp	Experimental Bits
FEC	Forwarding Equivalence Class
FRR	Fast Reroute
GMPLS	Generalized MPLS
GNS3	Graphical Network Simulator
GRE	Generic Routing Encapsulation
HTTP	Hypertext Transfer Protocol
ICMP	Internet Control Message Protocol
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IntServ	Integrated Services
IOS	Internetworking Operating System
IP	Internet Protocol
IPP	Max-Planck-Institut für Plasmaphysik
IPsec	IP Security
IS-IS	Intermediate System to Intermediate System Protocol
IS-IS-TE	IS-IS Extensions for Traffic Engineering
ISDN	Integrated Services Digital Network
ISO	International Organization for Standardization

ISP	Internet Service Provider
L-LSP	Label-Inferred-PSC LSP
LAN	Local Area Network
LDP	Label Distribution Protocol
LER	Label Edge Router
LMU	Ludwig-Maximilians-Universität
LRZ	Leibniz-Rechenzentrum
LSP	Label Switched Path
LSR	Label Switching Router
LWL	Lichtwellenleiter
MAC	Medium Access Control
MAN	Metropolitan Area Network
MLS	Multi-Layer-Switch
MPG	Max-Planck-Gesellschaft
MPLS	Multi-Protocol Label Switching
MWN	Münchner Wissenschaftsnetz
NAT	Network Address Translation
NVM	Non-Volatile Memory
OAM	Operation, Administration & Maintenance
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
OSPF-TE	TE Extensions for OSPF
P	Provider Core
PE	Provider Edge
PFC	Policy Feature Card
PHB	Per-Hop-Behaviour
PHP	Penultimate Hop Popping
PPP	Point-to-Point Protocol
PSC	PHB Scheduling Class
PVST	Per-VLAN-Spanning-Tree
PWE3	Pseudo Wire Emulation Edge-to-Edge
QoS	Quality of Service
RFC	Request for Comments
RIP	Routing Information Protocol
RSVP-TE	TE Extensions to RSVP for LSP Tunnels
RTT	Round-Trip-Time
SDH	Synchronous Digital Hierarchy
SDM	Spatial Division Multiplexing
SLB	Server Load Balancing
SONET	Synchronous Optical Network
SPoF	Single Point of Failure
SSH	Secure Shell
TC	Traffic Class
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TE	Traffic Engineering
TFTP	Trivial File Transfer Protocol

TTL	Time to live
TUM	Technische Universität München
UDP	User Datagram Protocol
VC	Virtual Circuit
VCI	Virtual Channel Identifier
VFI	Virtual Forwarding Instance
VLAN	Virtual LAN
VM	Virtual Machine
VMPS	VLAN Membership Policy Server
VoIP	Voice-over-IP
VPCS	Virtual PC Simulator
VPI	Virtual Path Identifier
VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
VRF	Virtual Routing and Forwarding
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
WLAN	Wireless LAN
WWW	World Wide Web

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[Abd02] Abdelhalim, Ahmed: *IP/MPLS-Based VPNs: Layer-3 vs. Layer-2*. Technischer Bericht, Foundry Networks, Inc., San Jose, CA, 2002.

[ABG+01] Awduche, D., L. Berger, D. Gan, T. Li, V. Srinivasan und G. Swallow: *RSVP-TE: Extensions to RSVP for LSP Tunnels*. RFC 3209, Internet Engineering Task Force (IETF), Dec 2001. http://www.ietf.org/rfc/rfc3209.txt.

[ADF+01] Andersson, L., P. Doolan, N. Feldman, A. Fredette und B. Thomas: *LDP Specification*. RFC 3036, Internet Engineering Task Force (IETF), Jan 2001. http://www.ietf.org/rfc/rfc3036.txt.

[alc09] *Virtual Private LAN Service (VPLS) Technical Primer*. Technischer Bericht, Alcatel-Lucent, 2009.

[Alv06] Alvarez, Santiago: *QoS for IP/MPLS Networks*, Kapitel 2. Cisco Press, Jun 2006.

[Alv08] Alvarez, Santiago: *MPLS Traffic Engineering Traffic Protection using Fast Re-route (FRR)*. Microsoft PowerPoint Presentation, Aug 2008.

[Bor02] Borowka, Petra: *Netzwerk-Technologien*. mitp-Verlag, Bonn, 2002.

[cisa] *Offizielle Webseite von Cisco Systems, Inc.* http://www.cisco.com.

[Cisb] Cisco Systems, Inc.: *Traffic Policing*. http://www.cisco.com/en/US/docs/ios/12_2t/12_2t2/feature/guide/ftpoli.pdf.

[Cis99] Cisco Systems, Inc.: *MPLS Traffic Engineering*, 1999. http://www.cisco.com/en/US/docs/ios/12_0s/feature/guide/TE_1208S.pdf.

[Cis02] Cisco Systems, Inc.: *Advanced Topics in MPLS-TE Deployment*, 2002.

[cis03] *Cisco IOS(R) MPLS Virtual Private LAN Service (VPLS)*. Technischer Bericht, Cisco Systems, Inc., Dec 2003.

[Cis07] Cisco Systems, Inc.: *Cisco Catalyst 6500/Cisco 7600 Series Supervisor Engine 720 Data Sheet*, 2007. http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps708/product_data_sheet09186a0080159856.pdf.

[Cis09] Cisco Systems, Inc.: *Cisco IOS Software Configuration Guide: Release 12.2 (33) SXH and Later Releases*, 2009.

[Hor00] Horney, Carter: *Quality of Service and Multi-Protocol Label Switching*. Technischer Bericht, Nuntius Systems, Inc., Irvine, CA, Nov 2000.

[ii004] *Quality of Service and MPLS Methodologies.* Technischer Bericht, IP Infusion Inc., San Jose, CA, 2004.

[Ike02] Ikejiri, Yuichi: *MPLS QoS E-LSP? L-LSP?* Microsoft PowerPoint Presentation, Jul 2002.

[IKPS04] Iselt, A., A. Kirstädter, A. Pardigon und Thomas Schwabe: *Resilient Routing Using MPLS and ECMP.* In: *Proceedings of International Workshop on High Performance Switching and Routing (HPSR).* IEEE, Apr 2004.

[Lak05] Lakshman, Umesh: *MPLS Configuration on Cisco IOS Software*, Kapitel 9. Cisco Press, Oct 2005.

[LR08] Leibniz-Rechenzentrum: *Das Münchner Wissenschaftsnetz: Konzepte, Dienste, Infrastrukturen, Management.* Technischer Bericht, LRZ, Garching, Jul 2008.

[LR09] Leibniz-Rechenzentrum: *Jahresbericht 2008.* Technischer Bericht, LRZ, Garching, 2009.

[LR10] Leibniz-Rechenzentrum: *Das Münchner Wissenschaftsnetz: Konzepte, Dienste, Infrastrukturen, Management.* Technischer Bericht, LRZ, Garching, Apr 2010.

[lrz] *Webseite des Netzmanagement-Server am Leibniz-Rechenzentrum.* `http://wwwmwn.lrz-muenchen.de`.

[mc006] *Demystifying Layer 2 and Layer 3 VPNs.* Technischer Bericht, Market Clarity Pty Ltd, Sydney, NSW, Jun 2006.

[Men03] Menga, Justin: *CCNP Practical Studies: Layer 3 Switching.* `http://www.ciscopress.com/articles/article.asp?p=102093`, Nov 2003.

[MW09] Morrow, Monique und Martin Winter: *MPLS Application, Services & Best Practices for Deployment.* Microsoft PowerPoint Presentation, Asia Pacific Regional Internet Conference on Operational Technologies (APRICOT), Feb 2009.

[nor03] *Virtual Private LAN Service (VPLS) using Distributed MPLS.* Technischer Bericht, Nortel Networks, 2003.

[OH09] Obermann, Kristof und Martin Horneffer: *Datennetztechnologien für Next Generation Networks: Ethernet, IP, MPLS und andere.* Vieweg, Wiesbaden, 2009.

[Orl05] Orlamünder, Harald: *Paket-basierte Kommunikationsprotokolle.* Hüthig Telekommunikation, Bonn, 2005.

[PD07] Peterson, Larry L. und Bruce S. Davie: *Computernetze: Eine systemorientierte Einführung.* dpunkt-Verlag, 2007.

[Pep06a] Pepelnjak, Ivan: *CEF load sharing details.* `http://blog.ioshints.info/2006/10/cef-load-sharing-details.html`, Oct 2006.

[Pep06b] Pepelnjak, Ivan: *Perfect Load-Balancing: How Close Can You Get?* `http://www.nil.com/ipcorner/LoadSharingTE/?open`, Nov 2006.

[Pep07a] Pepelnjak, Ivan: *10 MPLS traffic engineering myths and half truths.* `http://searchnetworking.techtarget.com.au/articles/23690-1-MPLS-traffic-engineering-myths-and-half-truths`, Dec 2007.

[Pep07b] Pepelnjak, Ivan: *Making the case for Layer 2 and Layer 3 VPNs.* `http://searchnetworking.techtarget.com.au/articles/21981-Making-the-case-for-Layer-2-and-Layer-3-VPNs`, Nov 2007.

[Pep07c] Pepelnjak, Ivan: *MPLS LDP autoconfiguration.* `http://blog.ioshints.info/2007/08/mpls-ldp-autoconfiguration.html`, Aug 2007.

[Pep07d] Pepelnjak, Ivan: *Unequal cost load-sharing.* `http://blog.ioshints.info/2007/02/unequal-cost-load-sharing.html`, Feb 2007.

[RR99] Rosen, E. und Y. Rekhter: *BGP/MPLS VPNs.* RFC 2547, Internet Engineering Task Force (IETF), Mar 1999. `http://www.ietf.org/rfc/rfc2547.txt`.

[RR06] Rosen, E. und Y. Rekhter: *BGP/MPLS IP Virtual Private Networks (VPNs).* RFC 4364, Internet Engineering Task Force (IETF), Feb 2006. `http://www.ietf.org/rfc/rfc4364.txt`.

[RTF+01] Rosen, E., D. Tappan, G. Fedorkow, Y. Rekhter, D. Farinacci, T. Li und A. Conta: *MPLS Label Stack Encoding.* RFC 3032, Internet Engineering Task Force (IETF), Jan 2001. `http://www.ietf.org/rfc/rfc3032.txt`.

[RTM99] Roth, Rudolf, Jens Tiemann und Lutz Mark: *MPLS-Study.* Technischer Bericht, GMD FOKUS Research Institute for Open Communication Systems, Berlin, 1999.

[RVC01] Rosen, E., A. Viswanathan und R. Callon: *Multiprotocol Label Switching Architecture.* RFC 3031, Internet Engineering Task Force (IETF), Jan 2001. `http://www.ietf.org/rfc/rfc3031.txt`.

[SMS06] Shah, Paresh, Utpal Mukhopadhyaya und Arun Sathiamurthi: *Overview of QoS in Packet-based IP and MPLS Networks.* Microsoft PowerPoint Presentation, North American Network Operators' Group (NANOG), 2006.

[SQ] Sawant, Asha Rahul und Jihad Qaddour: *MPLS DiffServ: A Combined Approach.* Technischer Bericht, Illinois State University.

[Ste98] Stephenson, Ashley: *Diffserv and MPLS: A Quality Choice.* Data Communications, 27(17):73–77, Nov 1998.

[SV01] Shaughnessy, Tom und Toby J. Velte: *Cisco-Systeme.* mitp-Verlag, Bonn, 2001.

[Wel03] Welcher, Pete: *Condensed QoS for MPLS.* Microsoft PowerPoint Presentation, 2003.

[Wu03] Wu, Tim: *MPLS VPNs: Layer 2 or Layer 3? Understanding the Choice.* Technischer Bericht, Riverstone Networks, Inc., Santa Clara, CA, Mar 2003.