Projektdokumentation

Realisierung einer redundanten

Kundennetzanbindung

Version 1.0

Autor:

Jan Wagner
waja --at-- cyconet *dot* org

TMT TeleService GmbH & Co. KG
Nürnberger Strasse 42
95448 Bayreuth
noc@tmt.de

Inhaltsverzeichnis

1 Projektbeschreibung

1.1 Vorstellen des umsetzenden Dienstleistungsunternehmens

Das IT-Dienstleistungsunternehmen „TMT TeleService GmbH & Co. KG“ (im weiteren Verlauf TMT) in Bayreuth umfasst drei wichtige Bereiche moderner Kommunikation. TMT bietet Unternehmen und öffentlich-rechtlichen Kunden eine abgestimmte und ganzheitliche Produktpalette in den Geschäftsbereichen „Webdevelopment und Design“, „Call Center“ und „IT- und Netzwerk-Sicherheit“ und beschäftigt derzeit mehr als 40 Mitarbeiter.

Die Abteilung für „IT- und Netzwerk-Sicherheit“ stellt die Arbeitsgruppe dar, welche mit der Betreuung des Kunden „Schlot GmbH“ [1], in deren Rahmen auch die betriebliche Projektarbeit durchgeführt wird, beauftragt wurde. Das Unternehmen TMT pflegt mehr als 100 Server (davon über 95% mit der Betriebssystembasis Linux[2]), verteilt auf zwei eigene Serverzentren in Bayreuth und einer angemieteten Co-Location in Düsseldorf. Zur Infrastruktur in Bayreuth gehört die Anbindung an das Internet als autonomes System über drei Festverbindungen. Durch diese multihome-Anbindung[3] [4] zu großen deutschen Carriern[5], mit den maximalen Übertragungsraten von zweimal 34 Mbit/s und einmal 155 Mbit/s, werden eine hohe Ausfallsicherheit und Bandbreite sowie kurze Paketlaufzeiten gewährleistet.

1.2 Darstellung des Projektumfeldes

Das Unternehmen „Schlot GmbH“, im weiteren Verlauf „Schlot“ genannt, ist eines der führenden Unternehmen für Schornsteinfertigelemente in Deutschland.

Die Firma Schlot hat deutschlandweit ca. 20 Standorte, welche ihren Warenaus- und -eingang sowie weitere unternehmensinterne Applikationen, über eine von TMT betreute VPN[6]-Infrastruktur, in der Konzernzentrale abwickeln. Weiterhin sind noch ca. 15 weitere ausländische Standorte und 20 „mobile“ Außendienstmitarbeiter in das VPN-Netzwerk integriert.

[1] Der Name der Firma (und ausschließlich dieser) ist frei erfunden, um die getätigten Investitionen von TMT und die Firmeninteressen unseres Kunden zu schützen. Eventuelle Parallelen zu anderen existierenden Firmen sind rein zufällig und nicht beabsichtigt.
[2] siehe auch http://www.kernel.org/links.html
[3] siehe RFC 1122 - Requirements for Internet Hosts -- Communication Layers, S. 60 - 64
[4] siehe Halabi, Bassam/ McPherson, Danny: Internet Routing Architekturen – Verbindungen und Protokolle professionell planen und realisieren, Übersetzung: Cosmos Consulting, München 2001, Markt + Technik Verlag (Orginalausgabe: Cisco Press), S. 149 - 150
[5] hier: Datenleitungsanbieter
[6] Virtual Private Network – „ist ein Computernetz, das zum Transport privater Daten ein öffentliches Netz (zum Beispiel das Internet) nutzt“, http://de.wikipedia.org/wiki/Vpn

1.3 Problemstellung

Die Anforderungen an die Verfügbarkeit des virtuellen Netzwerkes[7] sind sehr hoch, da bei einem Ausfall der Verbindung zum Warenwirtschaftssystem in den Niederlassungen keine Waren von Lieferanten entgegengenommen bzw. an Kunden verkauft werden können. Neben dem Umsatzausfall kommt es dabei an einigen Standorten zu größeren Rückstaus der LKWs bis auf Autobahnen. Weiterhin können an Zeiten gebundene Schwertransporte zum Teil nicht planmäßig durchgeführt werden, Einsatzkräfte (Polizei usw.) werden länger als nötig gebunden. So kann es auf Seiten des Kunden sehr schnell zu größeren finanziellen Schäden bei Verbindungsausfällen kommen.

Die momentane Anbindung an den Rest des Internets der Konzernzentrale erfüllt nicht das gewünschte Maß an Verfügbarkeit. Aus diesem Grund soll die vorhandene Anbindung ersetzt und um eine zusätzliche erweitert werden.

1.4 Ziel

Durch die Mitgliedschaft von TMT im RIPE NCC[8] bietet sich die Möglichkeit, autonome Netzwerkeinheiten, auch „Autonome Systeme“ genannt, zu bilden und von diesen eine Internetverbindung zu mehreren Providern aufzubauen. Dies wird auch als Multihoming[9] bezeichnet. Zu Projektende soll dem Kunde Schlot eine Internetanbindung zur Verfügung stehen, welche aus zwei voneinander unabhängigen Netzzuführungen besteht. Bei Ausfall einer der beiden Leitungen soll der Traffic[10] automatisch und für den Kunden transparent über die verbleibende abgewickelt werden.[11]

[7] hier VPN
[8] nähere Informationen siehe: http://www.ripe.net/info/ncc/about.html
[9] vgl. Internet Routing Architekturen – Verbindungen und Protokolle professionell planen und realisieren, S. 148 - 148
[10] „Als Datenverkehr (auch *Traffic*) bei Computern bezeichnet man den Fluss von Informationen innerhalb von Computernetzwerken.“, http://de.wikipedia.org/wiki/Datenverkehr
[11] siehe auch Anlage 5 – Übersicht BGP4+ Multihoming

2 Fachkonzept und Planung

2.1 Kosten-/Nutzenabwägung

Da, wie schon in der vorausgehenden Projektbeschreibung erwähnt, erhebliche Kosten bei einem Ausfall der IP-Kommunikation des Unternehmens entstehen, ist ein Mehraufwand für eine zusätzliche Anbindung und die Betreuung des BGP4+[12]-Routers eher unwesentlich. Vor der Projektrealisierung lag die Ausfallzeit über einer Stunde pro Monat während der gewöhnlichen Arbeitszeiten des Kunden, in welchen diesem der größte Schaden entsteht.

Uns liegen keine genauen Daten über die exakt entstehenden Kosten bei einem Ausfall der Kommunikationsstruktur des Kunden vor, aber eine grobe Schätzung eines IT-Verantwortlichen des Kunden geht von einem Schaden im fünfstelligen Bereich pro Stunde aus. Die Kosten für die zusätzliche Leitungs- und Netzwerktechnik sowie Betreuung dürften sich im vierstelligen Bereich pro Monat bewegen und damit wesentlich unter den Kosten liegen, welche allein durch Ausfälle aufgrund der Netzanbindung des Firmensitzes verursacht werden (könnten).

2.2 Ist-Analyse

Vor Beginn des Projekts erfolgte die Netzanbindung über eine dedizierte SDSL-Anbindung zu einem kleineren lokalen Provider. Über diese wird ein statisch geroutetes Netz mit der Netzmaske[13] /29[14] [15] zur Verfügung gestellt. Der Provider selbst ist mit einer einzigen Anbindung an das restliche Internet angeschlossen. Vermutlich ist dies einer der Grunde, weshalb auch keine zufrieden stellende Verfügbarkeit bereitgestellt werden konnte, wodurch es in der Vergangenheit immer wieder zu Komplettausfällen der VPN-Infrastruktur bei Schlot kam. Ebenfalls problematisch war das Ausfallmanagement, welches man eigentlich als solches nicht bezeichnen konnte, da meist beim Lieferanten niemand zu erreichen war, nicht einmal an der Telefonzentrale. Somit war TMT jede Handlungsmöglichkeit im Fehlerfall genommen.

[12] siehe RFC 1771 - A Border Gateway Protocol 4 (BGP-4)
[13] „ist eine Bitmaske, die eine IP-Adresse in einen Netzwerk- und einen Hostteil trennt“, http://de.wikipedia.org/wiki/Subnetmask
[14] auch 255.255.255.248, was 6 nutzbaren IPs pro Subnetz entspricht
[15] siehe RFC 950 - Internet Standard Subnetting Procedure, RFC 1518 - An Architecture for IP Address Allocation with CIDR, RFC 1519 - Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy sowie http://de.wikipedia.org/wiki/Classless_Inter-Domain_Routing

2.3 Soll-Konzept

Im Vollausbau sollte die Netzanbindung von Schlot durch zwei voneinander unabhängige Leitungen erfolgen. Das dynamische Routing wird durch das Routingprotokoll BGP4+, einem ASN[16] und einem Netz der Größe /24[17] realisiert.

Im Normalfall ist geplant, den Traffic über eine primäre Leitung abzuwickeln. Sollte es zu einem Ausfall kommen, wird nach außen hin transparent automatisch die zweite Leitung genutzt. Dies geschieht ohne Ausfallzeit, da das genutzte Routingprotokoll einen Leitungsausfall, oder besser gesagt einen Verlust der Routinginformationen, bemerkt und entsprechend der Konfiguration andere (bisher nicht verwendete) Routinginformationen benutzt.[18]

Durch diese Maßnahmen sollen die Verfügbarkeit und Qualität der Netzanbindung drastisch erhöht werden. Ebenfalls kann TMT als Anbieter bei einem Leitungsausfalls direkt auf die Lösung des Problems hinwirken, da die Carrier unmittelbar durch TMT beauftragt sind bzw. selbst an der Problembehebung arbeiten, falls die Ursache innerhalb der von TMT betreuten Infrastruktur auftreten sollten.

[16] vgl. RFC 1930 - Guidelines for creation, selection, and registration of an Autonomous System (AS)
[17] auch 255.255.255.0, was 254 nutzbaren IPs pro Subnetz entspricht
[18] siehe Anlage 5 – Übersicht BGP4+ Multihoming

3 Phasen der Umsetzung

3.1 Anforderungsabgabe für die Netzzuleitungen an den Einkauf

Folgende inhaltliche Anforderungen wurden bezüglich der Leitungsbeschaffung an den Einkauf übermittelt:

Eine der beiden Zuleitungen soll vom Standort des Kunden zum Standort des TMT Rechenzentrums, wo sich die Backbone-Infrastruktur befindet, geführt werden. Eine weitere Leitung soll von Schlot zu einem ISP freier Wahl führen. Von diesem wird dann ein BGP4+ Sitzung mit dem BGP4+ Router beim Kunden aufbaut und für die Netzbereiche von Schlot Transit[19] gewährt. Vorzugsweise würde sich ein ISP in geografischer Nähe anbieten, weil hier die Kosten für die Leitung günstiger sein sollten als bei einer größeren Entfernung.

Beide Leitungen müssen am Standort des Kunden als Medium (Fast-)Ethernet aufweisen, da die Netzwerkschnittstellen am Router vom Typ „FastEthernet“ sind.

3.2 Kosten-/Nutzenanalyse[20]

	Zustand Alt	Zustand Neu	
		Leitung 1	Leitung 2
Leitungskosten	638,-	800,-	463,07
Traffickosten	800,- bis 1500,-[21]	150,- bis 300,-[22]	incl.[23]
Monatl. Kosten	1438,- bis 2138,-	1413,07 bis 1563,07	

Zusätzlich entstanden dem Kunden einmalige Kosten von 1254,53 Eur für den Kauf der Hardware des BGP4+ Routers incl. Installation und Erstkonfiguration. Für den Kunden rechnet sich die Investition in den Kauf des BGP4+-Routes bereits, wenn durch diesen ein Komplettausfall von einer Stunde verhindert wird.

Die monatlichen Kosten für die Internetanbindung liegen entgegen der Annahme in 2.1 bei der neuen Anbindungsvariante sogar unter den Kosten für die alte Anbindung.

Der zusätzliche Wartungsaufwand dieser Lösung für TMT liegt im Wesentlichen unter den Kosten für das Problem-Management bei einem kompletten Leitungsausfall.

[19] hier: Durchleitungsgewährung von IP-Traffic
[20] Einkaufspreise TMT in Eur
[21] 10,00 Eur/Gigabyte
[22] 7,50 Eur/Gigabyte
[23] Flatrateanbindung

In der Gesamtbetrachtung sind die monatlichen Kosten für den Kunden und auch für TMT gesunken. Für TMT steigt in diesem Fall die Marge für den Supportvertrag. Die Investition in den Routerkauf rechnet sich wie schon erwähnt, sobald ein einstündiger Ausfall verhindert wurde.

Weiterhin ist die Ausfallsicherheit der Anbindung gestiegen und es stehen bessere und mehr Möglichkeiten zum Leitungs- und Bandbreitenmanagement zur Verfügung.

3.3 Registrierung eines ASN und Assignment eines geeigneten Netzwerkes

Um mittels einem BGP4+ Routers eine Sitzung des dynamischen Routingprotokolls aufzubauen, wird zwingend eine eindeutige Nummer für das „Autonome System" (kurz AS) benötigt[24]. Ähnlich wie bei IP-Adressen gibt es hier einen „offiziellen" Bereich, der durch die RIRs (Regional Internet Registries) verwaltet und vergeben wird. Es gibt aber auch einen „inoffiziellen" Bereich, welcher zur freien Verfügung steht. Allerdings gilt hier auch, analog zu den IP-Adressen, dass diese nicht im realen Internetbetrieb verwendet werden dürfen und ggf. bei einer BGP4+ Sitzung mit einem „offiziellen" AS durch dieses aus den Routingpfaden herausgefiltert werden müssen.[25]

Der zuständige RIR für den Kunden und auch für TMT ist das RIPE NCC. Bei diesem ist TMT als LIR (Local Internet Registy) Mitglied. In seiner Funktion als LIR ist TMT berechtigt bei seinem RIR sowohl eine ASN[26] als auch IP-Netzwerke zu beantragen.

Aufgrund der immer weiter steigenden Zahl der in der Routingtabelle geführten Netzwerke, werden durch nahezu alle im Internet involvierten ISPs, die ein eigenes ASN unterhalten, Prefixe[27] kleiner als /24 aus der BGP-Tabelle gefilter. Deshalb ist die Mindestgröße für ein Netzwerk auch /24, welches durch ein ASN verbreitet (bzw. announced) werden kann, um garantiert von überall erreichbar zu sein.

Durch das RIPE NCC wurde TMT ein Netzwerk der Größe /20 zur Verfügung gestellt, das für die eigene Versorgung und die der Kunden genutzt werden kann. Allerdings muss für jedes Netz, das einem endgültigen Verwendungszweck zugeführt wird, ein Antrag auf Benutzung von IP-Adressen ausgefüllt werden. Darin wird der Grund und der Nutzungszweck des Endkunden für die Netzgröße gegenüber dem RIPE NCC dargelegt. Da die aktuelle Größe zur zustimmungsfreien Inbetriebnahme (Allocation Window, kurz AW) durch TMT /24 beträgt, muss das Formular zur IP-Nutzung zwar ausgefüllt und dem RIR auf Verlangen vorgelegt werden, allerdings kann die Inbetriebnahme ohne die

[24] vgl. RFC 1930, S. 1 - 8
[25] vgl. RFC 1930, S. 9
[26] weiterführende Informationen zu AS Nummern unter: http://www.ripe.net/rs/as/index.html
[27] hier: Netzmaske

Zustimmung des RIRs erfolgen[28]. Dies geschieht im Allgemeinen durch die Eintragung[29] in die „RipeDB“[30], in unserem Fall ist dies 108.217.145.0/24.[31]

Da zur Beantragung eines ASN[32] beim RIR ein Netzwerk angegeben werden muss, welches durch das AS gehostet wird, ist es nötig dieses bereits in der RipeDB anzulegen, bevor das ASN beim RIR beantragt werden kann und dann ebenfalls in die RipeDB eingetragen wird[33]. Für die Erstellung der beiden Objekte in der RipeDB sind ebenfalls Personenobjekte und Organisationsobjekte für die administrativen Ansprechpartner notwendig, die in den beiden Objekten verwendet werden.

Eine weitere Funktion des ASN-Objektes in der RipeDB ist, dass import- und export-Angaben zur automatischen Erstellung von Filtern bei den Providern benutzt werden. In diesen wird veröffentlicht, von welchem AS welche Routinginformationen in die BGP4+ Tabelle importiert werden und welchem AS welche Routinginformationen angeboten (announced) werden.

Dieses Verfahren setzt also einen korrekten Eintrag mit der aktuellen Import-, Exportkonfiguration des AS voraus, damit ein fehlerfreier Betrieb gewährleistet werden kann.

3.4 Multihop Problematik

Leider konnte der Einkauf keine Anbindungen beschaffen, die direkt als Ethernet terminiert werden. Auf Kundenseite sind an beiden Leitungen Router installiert, die ihrerseits eine (Fast-)Ethernetschnittstelle zur Verfügung stellen.

Dies ist deshalb problematisch, weil normalerweise bei einer BGP4+ Sitzung beide Router ohne weitere zwischen geschaltete Router direkt miteinander verbunden sind. Bei den meisten BGP4+ Implementation ist zwar die Möglichkeit einer „Multihopsitzung“ gegeben, diese erhöht aber den Wartungsaufwand unter Umständen erheblich. Bei bestimmten Konstellationen ist gar keine Umsetzung möglich, da alle Routinginformationen aus dem externen Routingprotokoll auch auf die dazwischen liegenden Router übertragen werden müssen.[34]

In unserem Fall ließ sich das Problem lösen, da im Kunden-AS nur ein Prefix[35] gehostet und auch kein Transit zu anderen AS gegeben wird[36]. Auf den Routern[37] zwischen den BGP4+-Routern wurde die Standardroute auf das Interface gelegt,

[28] mehr Informationen sind http://www.ripe.net/ripe/docs/ipv4-policies.html zu entnehmen
[29] siehe Anlage 2 – Auszug des Netz-Objekts in der RipeDB
[30] Informationen zu Sinn und Nutzungsmöglichkeiten sind http://www.ripe.net/db/index.html zu entnehmen
[31] Sämtliche IP-Adressen und die ASN von Schlot wurden geändert, um die getätigten Investitionen von TMT und die Firmeninteressen unseres Kunden zu schützen.
[32] siehe auch: http://www.ripe.net/ripe/docs/asn-assignment.html
[33] siehe Anlage 4 – Auszug des ASN-Objekts in der RipeDB
[34] vgl. Internet Routing Architekturen – Verbindungen und Protokolle professionell planen und realisieren, S. 189 - 190
[35] 108.217.145.0/24
[36] D.h. Schlot stellt Dritten keine Internetverbindung mittels dynamischen Routingprotokollen zur Verfügung.
[37] bei der Anbindung TMT ist dies „acryl“ und „schornstein“

das vom Kunden „wegzeigt“ und die Route zum Kundenprefix natürlich in Richtung Kunde.

Dies möchte ich beispielhaft am Router „schornstein“ auf Schlot-Seite, der die primäre Multiplexanbindung für die Anbindung TMT terminiert, erklären. Wie Anlage 7 – Anbindungsschema TMT zu entnehmen, ist die Standardroute auf die serielle Schnittstelle gelegt[38], über die die Verbindung zu TMT hergestellt wird. Somit werden auf diesem Router alle Pakete zu TMT geroutet, wenn sie nicht an IPs des Netzbereichs 108.217.140.160/29[39] gerichtet sind oder eine spezifischere[40] Route vorhanden ist. Solch eine ist für 108.217.145.0/24 konfiguriert[41]. Als Gateway wird hierfür der BGP4+-Router[42] des Kunden benutzt. Durch diese Konfiguration werden auf dem Router alle Pakete standardmäßig vom Kunden weg und nur Pakete mit Zieladresse aus dem Kundennetzbereich[43] zum Kunden hin geleitet.

Die Peeringpartner[44] mussten nun nur noch als Multihop konfiguriert werden und arbeiteten in unserem Fall problemlos über diese logische Verbindung miteinander[45]. Falls Schlot in Zukunft weitere Netze benötigt, müssen neben seinem BGP4+ Router auch die Router zwischen den BGP4+ Routern in der Konfiguration entsprechend angepasst werden.

3.5 Inbetriebnahme der Netzwerkkomponenten

Als Hardware für den BGP4+ Router auf Kundenseite wurde ein „Nexgate NSA-1030-R“ benutzt. Dies ist ein vollwertiger PC im 19“ Format, mit einer Bauhöhe von 1HE und halber Bautiefe. Dadurch ist es möglich, in einem Rack auf einer Höheneinheit zwei dieser Einheiten zu verbauen. An der Frontseite des Gerätes sind drei Netzwerkschnittstellen angebracht und der Rückseite befindet sich das Netzteil incl. Lüfter.[46]

Auf dieser Hardware wurde ein „Debian GNU/Linux ‚sarge‘“[47] in der Grund-Installation, ein Routing-Daemon namens „Quagga“[48] sowie das iptables-Frontend, „Shorewall“[49], mittels des Frontends der distributionseigenen Paketverwaltung „apt-get“[50] installiert. Der Paketfilter ist allerdings nicht zur Projektrealisierung nötig und wurde nach Ende des Projekts konfiguriert.

[38] durch die Anweisung „ip route 0.0.0.0 0.0.0.0 Serial0“
[39] Lokaler Netzbereich, mit dem die Ethernet-Schnittstelle konfiguriert wurde.
[40] genauere
[41] durch die Anweisung „ip route 108.217.145.0 255.255.255.0 108.217.140.162“
[42] mit der IP 108.217.140.162
[43] 108.217.145.0/24
[44] BGP4+-Router bei Schlot und TMT bzw. Mnet
[45] zur Verdeutlichung siehe Anlage 6 – Grobübersicht Router
[46] nähere Informationen zu den Hardwarespezifikationen sind Anlage 1 – Hardwarespezifikationen Kundenrouter zu entnehmen.
[47] für nähere Informationen zum Debian Release siehe: http://www.debian.org/releases/sarge/
[48] für nähere Informationen zu Quagga siehe: http://www.quagga.net/docs/quagga.pdf
[49] für eine Einführung in Shorewall siehe: http://www.shorewall.net/Introduction.html
[50] für nähere Informationen zu apt-get siehe: http://www.debian.org/doc/manuals/apt-howto/index.de.html

Da die jetzige Leitung vom Kunden zum alten ISP zu dem ISP geschwenkt werden sollte, welcher auch den BGP4+-Fullfeed zur Verfügung stellt, musste der Router vor Inbetriebnahme der neuen Leitungen zum Kunden verschickt werden, um eine hinreichend kurze Ausfallzeit zu erreichen. Dieser wurde dann vor Ort im Serverraum des Kunden durch Kundenpersonal unter telefonischer Anleitung installiert.

Damit der Router „core1“ auch zur Fernwartung und Konfiguration bis zur Inbetriebnahme der eigentlichen Anbindungen erreichbar war, wurde der Router auf der ersten Netzwerkschnittstelle (eth0) mit einer IP des alten Netzes konfiguriert und an den externen Switch gesteckt.

3.5.1 Inbetriebnahme der Leitung „TMT-Kunde“

Die Leitung zwischen TMT und dem Kunden war als erste verfügbar und wurde über einen Dienstleister realisiert, mit dem schon einige weitere Lösungen im Bereich „Access“ umgesetzt wurden. Die Verbindung wurde bei Schlot durch einen primären Multiplexanschluss zu einem „Point of Presence“ (PoP) des ISPs hergestellt. Die Zuführung auf der Seite TMT wurde über einen L2TP[51]-Tunnel durchgeführt. Letztlich wurde eine logische serielle Verbindung über das ppp-Protokoll zwischen dem Router „schornstein“ auf der Kundenseite und dem für das Layer2-Tunneling zuständigen Router („acryl“) im TMT-Backbone aufgebaut. Der Router auf Kundenseite ist ein Cisco 1601[52] mit einem Ethernetinterface und einer S2M Schnittstelle, welche an die Primärmultiplex-Zuführung gekoppelt ist.[53] „acryl“ läuft auf einem „Cisco 2620“[54] [55], die entsprechende Konfiguration wurde durch den Kollegen (Maintainer) durchgeführt, der den Backbone-Router wartet. Nach außen hin arbeitet diese Lösung transparent wie eine „normale“ Multihop-Verbindung. „schornstein“[56] wurde mit Hilfe des an „acryl“[57] angebundenen Radius-Servers über die PPP-Sitzung die IP „108.217.140.145/25“ zugewiesen.[58]

3.5.2 Worst Case Szenario

Da in der Ausgangssituation Mnet als Carrier fungierte und lediglich der Datenverkehr über den alten ISP abgewickelt wurde, sollte der Leitungsendpunkt nur zu einem Mnet PoP geschwenkt werden. Dies hat zur Folge, dass der Kunde nach dem Leitungsschwenk nicht mehr unter seinen alten Netzadressen erreichbar ist.

[51] vgl. RFC 2661 - Layer Two Tunneling Protocol "L2TP"
[52] als Software wird „IOS (tm) 1600 Software (C1600-Y-L), Version 12.0(27), RELEASE SOFTWARE (fc1)“ eingesetzt.
[53] für nähere Informationen zu Cisco 1601 Routern siehe:
http://www.cisco.com/en/US/products/hw/routers/ps214/ps216/index.html
[54] Software-Release „IOS (tm) C2600 Software (C2600-IS-M), Version 12.2(8)T5, RELEASE SOFTWARE (fc1)“
[55] Informationen zur Cisco 2600 Serie:
http://www.cisco.com/univercd/cc/td/doc/product/access/acs_mod/cis2600/26xx_qsg/2600_qsg.htm
[56] Serieller Router auf Schlot-Seite
[57] L2TP-Endpunkt bei TMT
[58] zur Verdeutlichung siehe Anlage 6 – Grobübersicht Router und Anlage 7 – Anbindungsschema TMT

Leider kam es durch eine Kombination ungünstiger Umstände zu einem Schwenk durch Mnet zu einem Zeitpunkt, als das alte Netz noch in Produktivbetrieb war, das neue Netz jedoch noch nicht.

Deshalb wurde kurzfristig die Leitung „TMT-Kunde“ in Betrieb genommen, da alle dafür notwendigen Objekte bereits korrekt in der RipeDB eingetragen und die Leitung funktionstüchtig gewesen ist. Ebenfalls war bereits die BGP4+ Sitzung so konfiguriert, dass ein Notbetrieb ermöglicht werden konnte. Problematisch dabei war, dass sich dies Freitagvormittag ereignete, also mitten in der Geschäftszeit des Kunden. Somit konnte über einen längeren Zeitraum keine Verbindung zwischen den Zweigniederlassungen und der Konzernzentrale aufgebaut werden, was natürlich erhebliche Probleme verursachte.

Die technische Herausforderung bestand darin, die VPN-Zugangsserver und die VPN-Clients mit Hochdruck umzukonfigurieren[59]. Die Umnummerierung der Kundeninfrastruktur konnte, aufgrund der stehenden Leitung zu TMT, durch Fernzugriff durchgeführt werden und beschränkte sich auf drei Rechner. Die VPN-Clients waren durch eine Datenbank ihrer aktuellen „Einwahl-IPs“ erreichbar und konnten ebenfalls umgestellt werden. Vom Ausfallzeitpunkt, über Problemanalyse, Inbetriebnahme der Leitung, Renummerierung und Rekonfigurierung sind über 8 Stunden vergangen. Details zur Vorgehensweise bei der Sitzungs-Konfiguration folgen im weiteren Verlauf.

3.5.3 Inbetriebnahme der Leitung „Mnet-Kunde“

Wie schon erwähnt, wurde die Leitung zwischen Mnet und Kunden vom alten ISP zu Mnet geschwenkt. Hierbei wurden lediglich zwei Transfernetze von Mnet am alten Router auf Kundenseite durch TMT konfiguriert[60], einmal zwischen dem Router von Mnet und dem Router zum Leitungsabschluss der symmetrischen DSL-Leitung[61] und einmal zwischen dem BGP4+-Router des Kunden und dem DSL-Router. Auch hier wurde, wie oben schon hingewiesen, eine Multihop-Verbindung zwischen den BGP4+-Routern von Mnet und Kunden aufgebaut. Als DSL-Router kommt ein Gerät der Marke „IBM“ zum Einsatz. Da dieser allerdings schon seit mehreren Jahren vor Ort in Betrieb ist und deshalb auch keine Unterlagen mehr zu dem Gerät auffindbar waren, sind die genauen Spezifikationen unbekannt.[62]

[59] die Rekonfiguration der eigentlichen VPN-Struktur erfolgte durch den betreuenden Kollegen
[60] erfolgte durch den, den Router betreuenden, Kollegen
[61] Durchsatz von 2 MBit/s
[62] zur Verdeutlichung siehe Anlage 6 – Grobübersicht Router und Anlage 8 – Anbindungsschema Mnet

3.6 Erstellung des Routeobjektes

Wie schon in 3.2 erwähnt, werden Objekte der RipeDB bzw. der Datenbanken der anderen RIRs, zur Erstellung von Filterregeln an den BGP4+-Routern herangezogen. Ein weiteres dieser Objekte ist das so genannte Route-Objekt.

Solange für ein Netzwerk kein Route-Objekt existiert, wird es aus den Routingtabellen herausgefiltert. Neben dem Netz ist auch die ASN Bestandteil des Objektes, von welchem das Netzwerk veröffentlicht werden darf (Beginn des AS-Pfads). Solange die Informationen, welche in den Route-Objekten enthalten sind, nicht mit denen der BGP4+-Tabelle übereinstimmen, werden diese bei nahezu allen Providern herausgefiltert. Aus diesem Grund ist die Korrektheit des Objektes von großer Bedeutung.[63]

3.7 Konfiguration des dynamischen Routings

Nachdem alle Netzwerkkomponenten bezüglich der Leitung „TMT-Kunde“ in Betrieb genommen und alle notwendigen Maßnahmen bezüglich des RIPE NCC und RipeDB abgewickelt waren, konnte die Konfiguration auf BGP4+-Ebene erfolgen. Wichtig ist, wie unter 3.4 schon erwähnt, dass auf „schornstein“[64] und „acryl“[65] die Standardroute auf die Schnittstellen „in Richtung“ TMT zeigen, in welche natürlich der gesamte Traffic geleitet werden soll, der nicht an das Kundennetz adressiert ist. Umgekehrt muss natürlich die Route zum Kundennetz auf die Schnittstellen „in Richtung“ Kunde konfiguriert werden[66]. Nur so kann diese Lösung verwirklicht werden, da die dazwischen liegenden Router keine dynamischen Routingprotokolle beherrschen.

Auf „core1“[67] wird als nächstes die Routingsuite Quagga gestartet, genauer gesagt die Serverprozesse „bgpd“ und „zebra“. Der „bgpd“ ist die BGP4+ Implementierung, welche mit dem Prozess „zebra“ kommuniziert. Dieser wiederum übernimmt die Kommunikation mit dem (Betriebssystem-)Kernel, um die Routingtabelle zu manipulieren. Weiterhin können hier den Schnittstellen IP-Adressen usw. zugewiesen und statische Routen gesetzt werden.

Die Command Line Interface(s) (CLI) können per telnet über Port 2601 (zebra) und 2605 (bgpd) erreicht und Änderungen an den Konfigurationen vorgenommen werden. Auf der Seite des Cisco Routers „sackgasse“ gibt es keine getrennte Implementierung und die CLI ist über Standardport 23 erreichbar.

Am „core1“ muss zunächst „IP Forwarding“ aktiviert sein, damit IP-Pakete weitergeleitet werden, wie bei einem Router nötig. Als nächstes werden die ASN[68]

[63] weiterführende Informationen entnehmen Sie bitte: http://www.ripe.net/ripe/docs/ripe-252.html#1.2.16
[64] Serieller Router auf Schlot-Seite
[65] L2TP-Endpunkt bei TMT
[66] siehe Anlage 7 – Anbindungsschema TMT
[67] BGP4+-Router Schlot
[68] vgl. Doyel, J. und DeHaven Carrel, J.: Routing TCP/IP Band II, München 2002, Markt + Technik Verlag (Originalausgabe: Cisco Press), S. 182

und die Router-ID[69] konfiguriert. Auf „core1“ wird angegeben, welches Netzwerk über BGP4+ veröffentlicht (announcet) werden soll[70], in unserem Fall das Kundennetz „108.217.145.0/24“. Ebenfalls müssen der BGP-Nachbar inklusive zugehöriges ASN[71] und, wie mehrfach erwähnt, bgp-multihop[72] mit, in unserem Fall, max. 5 Hops zwischen den Nachbarn eingerichtet werden[73]. Die „Update-Source“[74] ist nur am „core1“ zu konfigurieren.

Mittels dieser Konfiguration ist bereits ein Produktivbetrieb möglich, allerdings werden auf „core1“ Netze kleiner als /24, welche von „sackgasse“ angeboten werden und alle anderen BGP-Pfade, welche nicht am „core1“ beginnen, aus dem „Announcement“ zu „sackgasse“ herausgefiltert[75].

Andererseits werden auf „sackgasse“ aus dem „Announcement“ zu „core1“ BGP-Pfade mit privatem AS gefiltert[76]. Netze kleiner /24 und BGP-Pfade[77], die nicht mit dem AS des Kunden beginnen, werden auf „sackgasse“ vom „core1“ verworfen.

Nach dem Leitungsschwenk des Anbieters Mnet, welcher näher unter 3.5.2 beschrieben wurde, konnte die Nachbarverbindung zum zweiten Nachbarn aufgebaut werden. Auf dem IBM-Router, welcher die Leitung zu Mnet terminiert, wurden ebenfalls analog zu „schornstein“ und „acryl“ Standardrouten zu Mnet und eine Route des Kundennetzes auf die Schnittstelle zum Kunden konfiguriert.

Die BGP-Nachbarschaftskonfiguration zwischen „core1“ und dem BGP4+-Router von Mnet wurde ebenfalls analog zum Nachbarn TMT eingerichtet.[78] Allerdings musste hier nur eine Konfiguration auf dem Router „core1“ vorgenommen werden. Die Konfiguration am BGP4+-Router von Mnet wurde durch einen mnet-eigenen Techniker durchgeführt.

3.8 Testbetrieb und eventuelle Nachkonfiguration / Überführung in den Produktivbetrieb

Außer einem Leitungstest zwischen Kunden und TMT konnte aufgrund des in 3.5.2 angesprochenen Problems kein eigentlicher Testbetrieb durchgeführt werden.

Der Anbindungsvertrag über Mnet beinhaltet eine Traffic-Pauschale, unabhängig der tatsächlich anfallenden Menge, was meist auch als Flatrate bezeichnet wird. Aufgrund dieser Tatsache wurde entschieden, die Anbindung als

[69] vgl. Routing TCP/IP Band II, S. 186 - 188
[70] vgl. Routing TCP/IP Band II, S. 193 - 194
[71] vgl. Routing TCP/IP Band II, S. 182
[72] vgl. Routing TCP/IP Band II, S. 219
[73] Für Fallbeispiel siehe Routing TCP/IP Band II, S. 217 - 220
[74] vgl. Routing TCP/IP Band II, S. 217
[75] siehe Routing TCP/IP Band II, S. 257 - 262
[76] siehe Routing TCP/IP Band II, S. 257 - 262
[77] siehe Routing TCP/IP Band II, S. 257 - 262
[78] nähere Informationen siehe Anlage 8 – Anbindungsschema Mnet

Standardanbindung und die Anbindung über TMT nur bei Ausfall der Standardanbindung zu benutzen.

Deshalb wurde die Konfiguration zum Nachbarrouter von TMT so verändert, dass nur noch „0.0.0.0/0“ als Route in der BGP-Tabelle angenommen wird und im Kernel-Space wurde der Router der TMT-Anbindung als Standardgateway konfiguriert. Dies erfolgte in Anlehnung an eine Empfehlung in „Routing TCP/IP Band II“[79]. Da ja über die BGP4+ Verbindung zu Mnet spezifischere Routen in die Kernel-Routingtabelle gelangen, werden diese genutzt. Nur bei Ausfall dieser Verbindung oder Nichtverfügbarkeit einer entsprechenden Route, wird ausgehend, aufgrund der Standardroute, auf die Verbindung zu TMT zurückgegriffen.

Gegenüber dem BGP4+-Router von TMT wird durch die Route-Map „tmt-out“ der angebotene AS-Pfad um vier weitere Einträge[80] des Kunden-ASN verlängert[81], wie es auch in „Routing TCP/IP Band II“[82] erläutert ist. Da die Länge des AS-Pfades als Hauptmerkmal innerhalb der BGP4+-Logik benutzt wird, wird mit hoher Wahrscheinlichkeit bei den meisten ISPs die Anbindung über Mnet genutzt, da der AS-Pfad zum Kunden-AS über Mnet durch das „Prepending“ der TMT-Sitzung bei nahezu allen Teilnehmern kürzer sein wird, solange die Anbindung über Mnet nicht beeinträchtigt ist.

3.9 Abschließende Bemerkungen

Eine Dokumentation für Schlot wurde im Rahmen des Projektes nicht angefertigt, da die Veränderung für diesen transparent ist, d.h. eine Veränderung des Verhaltens gegenüber dem Kunden erfolgte nicht. Es wurde lediglich die Verfügbarkeit erhöht. Aus diesem Grund und weil TMT die Internetanbindung komplett betreut, wurde keine Kundeneinweisung und Abnahme im eigentlichen Sinne mit dem Kunden durchgeführt. Sehr wohl wurde die TMT-interne Anbindungsdokumentation angepasst und aktualisiert, diese kann allerdings aus unternehmenspolitischen Gründen nicht veröffentlicht werden und wurde aus Zeitmangel nach Projektende durchgeführt.
Eine Übergabe des funktionstüchtigen Systems erfolgte TMT-intern mit Ende des Projekts an den Leiter der Abteilung „IT- und Netzwerk-Sicherheit“. Dabei wurden auch wechselseitige Leitungsausfälle durch Abschalten der jeweiligen Netzwerkschnittstellen erfolgreich simuliert

[79] Doyel, J. und DeHaven Carrel, J., Routing TCP/IP Band II, München 2002, S. 98 - 114
[80] siehe Anlage 7 – Anbindungsschema TMT
[81] AS Path Prepending
[82] Routing TCP/IP Band II, S. 127.

4 Zeitlicher Ablauf

Datum	Phase	Tätigkeit	Dauer
04.04.05	Kosten-/Nutzenabwägung	Evaluierung der Projektzweckmässigkeit	1 h
	Ist-Analyse	Analyse der bestehenden Infrastruktur	2 h
05.04.05	Soll-Konzept	Grobkonzept	3 h
		Planung der zu verwendenden Hard- und Software	2 h
		Vorüberlegungen zur Anbindungsentscheidung	1 h
06.04.05	Umsetzungsphase	Abstimmung mit Einkauf (Anforderungen)	1 h
11.04.05		Kosten-/Nutzenanalyse	1 h
		Registrierung von Netz und ASN bei RIPE NCC	2,5 h
18.04.05		Inbetriebnahme BGP4+-Router	1 h
19.04.05		Inbetriebnahme der Leitung „TMT-Kunde“	1 h
		Erstellung des Routeobjekts	0,5 h
20.04.05		Konfiguration des dynamischen Routings zu TMT (Testbetrieb)	8 h
22.04.05	Inbetriebnahme	Leitungsschwenk der Mnet-Anbindung (incl. Fehlersuche)	2 h
		Überführung der Anbindung TMT in den Produktivbetrieb	2 h
25.04.05		Konfiguration des dynamischen Routings zu Mnet	3 h
		Überführung der Anbindung MNet in den Produktivbetrieb	2 h
29.04.05		Nachkonfiguration	2 h
Gesamtdauer			35 h

5 Anlagen

Anlage 1 – Hardwarespezifikationen Kundenrouter[83]

Construction		
Construction		Heavy-Duty Steel Chassis
Mounting Configuration		19" Rackmount
Height, Units		1 U
CPU Card		
CPU Board Form Factor		5.25" SBC
Processor	CPU	VIA C3 EBGA Package
	Number of CPU	1
	CPU Socket	CPU On Board
	CPU Frequency	800 MHz
System Memory	Maximum Memory Size	512 MB
	Memory Type	SDRAM
	Memory Sockets	1xDIMM 168pin
	ECC	No
Solid State Disk	Extension Memory	DiskOnModule (DOM)
Ethernet	Ethernet Channels	3
	100Mbps	3
	Controller	3 x Realtek RTL8139C 10 Base T/100 Base TX
	Connector	3xRJ45
IDE	Channels	2xE-IDE
	Mode	Ultra DMA 33/66/100 Mb/s
COM	Ports, total	1
	RS-232 Ports	1
Watch Dog Timer		Software Programming
Disk Drive Bays		
3.5" Drive Bays	All	1
	Hidden	1
Front Panel		
LEDs		1xOn/Off LED
Connectors		3xRJ45 Etherne
Cooling		
Fans		2x40 mm
Hot Swappable Fan		No
Power Supply		
Type of Power Supply		ATX

[83] http://www.ipc2u.de/catalog/0/MB.html

Output Power		180 W
Input Voltage Range	AC	90...264 V
Operation		
Operation	Temperature	0...40 °C
	Humidity	10...95 %
Storage		
Storage	Temperature	-20...80 °C
	Humidity	5...95 %
Dimensions and Weight		
Dimensions	Width	426 mm
	Height	43.5 mm
	Depth	228 mm
Certifications		
Certifications		CE,FCC Class B

Anlage 2 – Auszug des Netz-Objekts in der RipeDB (ohne Person- und Roleobjekte)

inetnum: 108.217.145.0 - 108.217.145.255
netname: TMT-DE-SCHLOT-NET-1
descr: Schlot GmbH
country: DE
admin-c: HB1750-RIPE
tech-c: TMT1
status: ASSIGNED PA
mnt-by: TMT-MNT
mnt-lower: TMT-MNT
source: RIPE

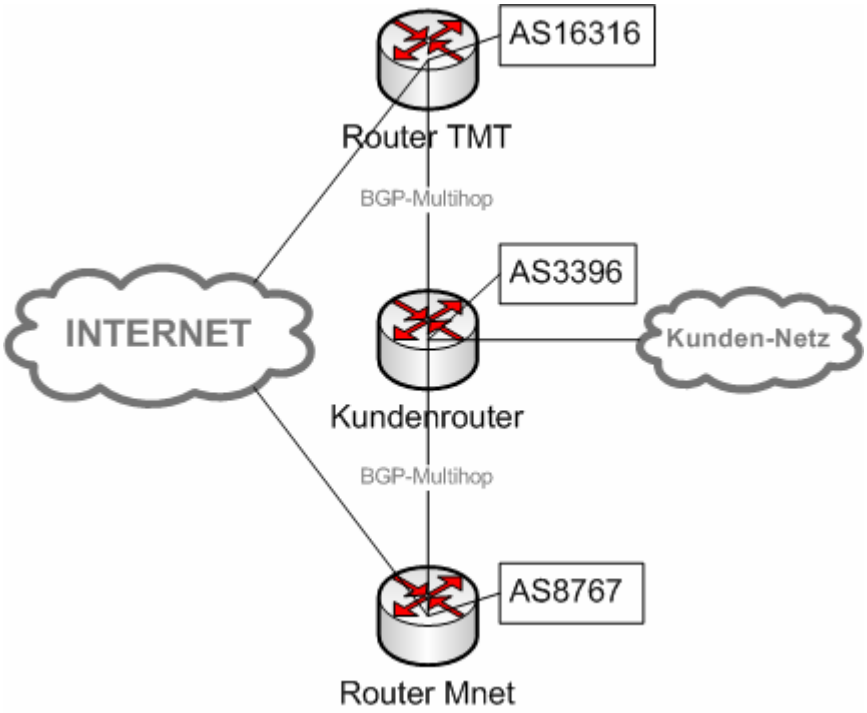
Anlage 3 – Auszug des Routeobjekts in der RipeDB (ohne Person- und Roleobjekte)

route: 108.217.145.0/24
descr: TMT-DE-AS3396-108_217_145_0-24
descr: DE
origin: AS3396
mnt-by: TMT-MNT
mnt-lower: TMT-MNT
source: RIPE

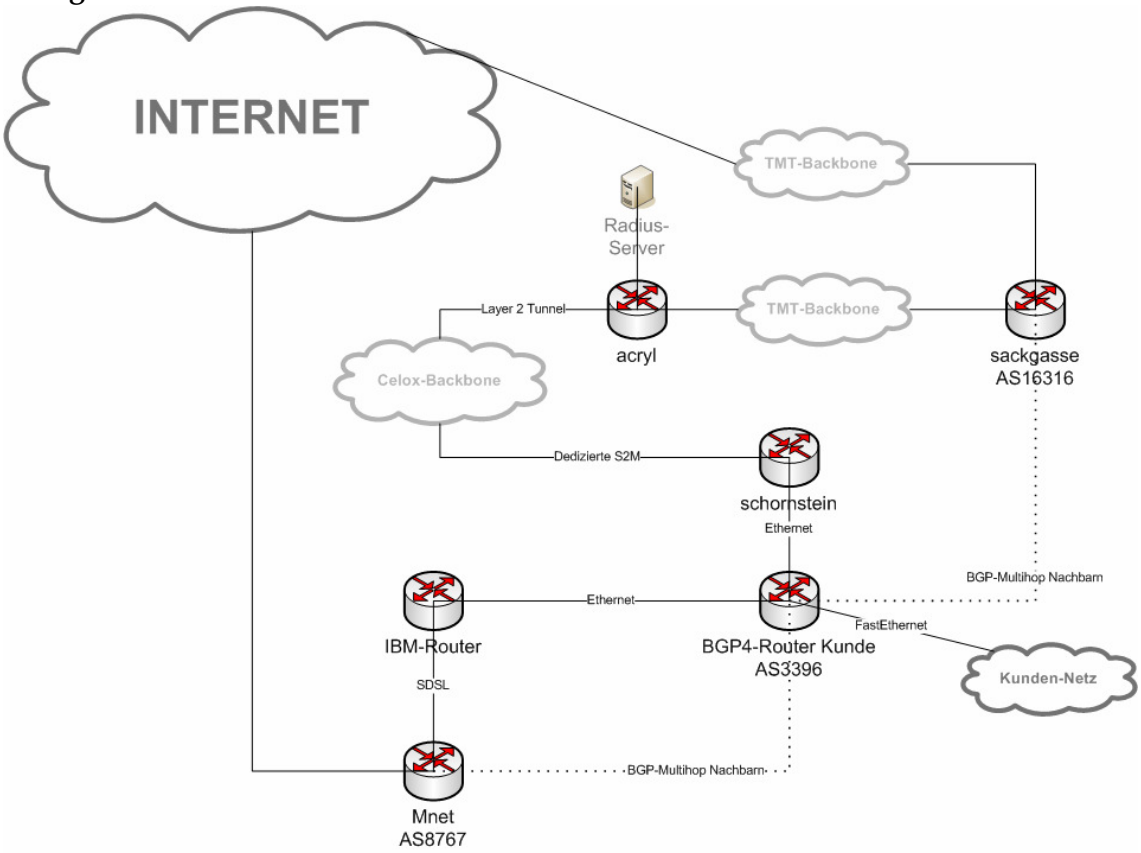
Anlage 4 – Auszug des ASN-Objekts in der RipeDB (ohne Person- und Roleobjekte)

aut-num: AS3396
as-name: SCHLOT-AS
descr: Schlot GmbH
descr: Lerchenstrasse 9
descr: D-80995 Muenchen
import: from AS8767 action pref=100; accept ANY
import: from AS16316 action pref=200; accept ANY
export: to AS8767 announce AS3396
export: to AS16316 announce AS3396
org: ORG-SA104-RIPE
default: to AS8767 action pref=100; networks ANY
admin-c: PM1
tech-c: TMT1
mnt-routes: TMT-MNT
mnt-by: TMT-MNT
source: RIPE

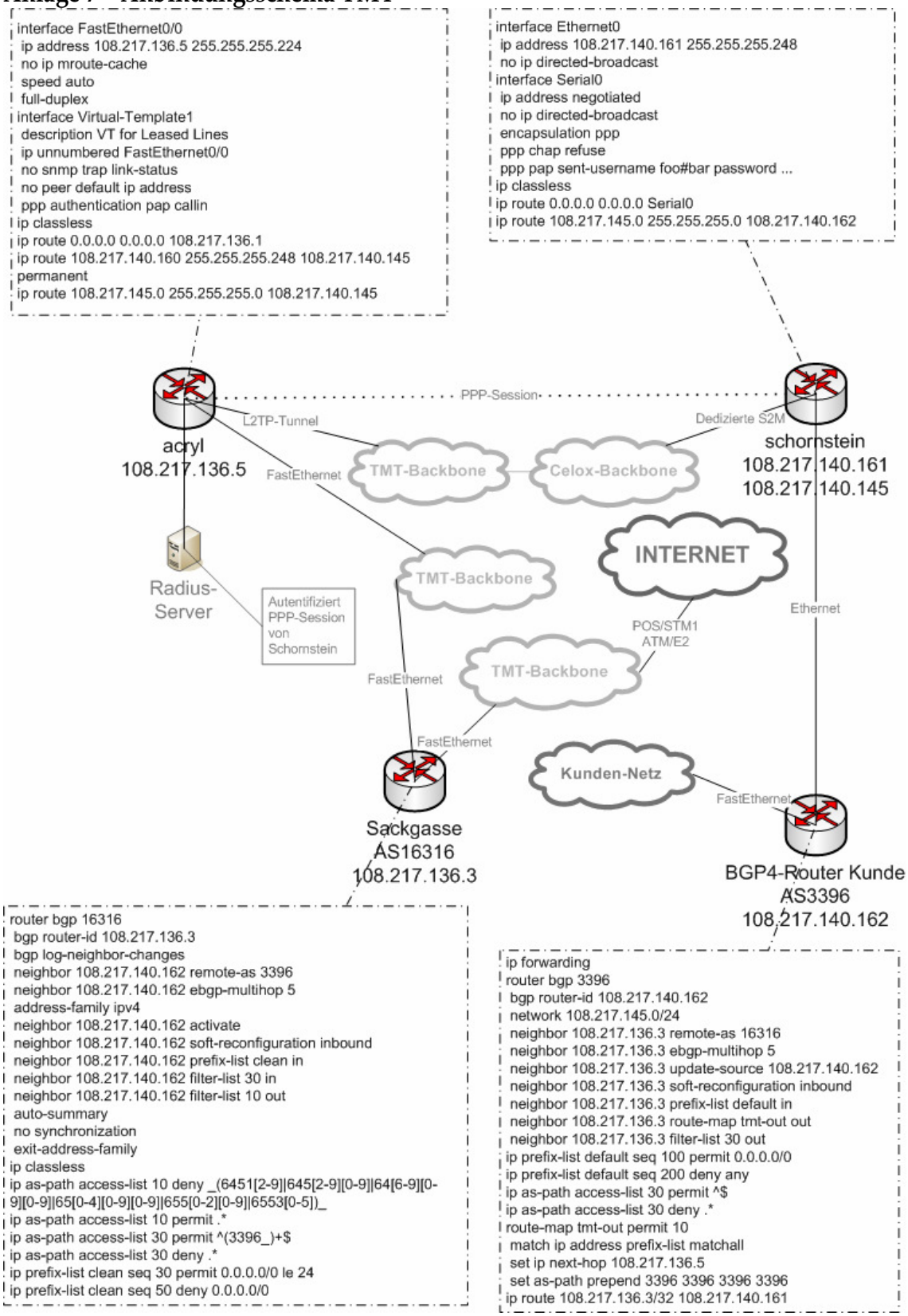
Anlage 5 – Übersicht BGP4+ Multihoming



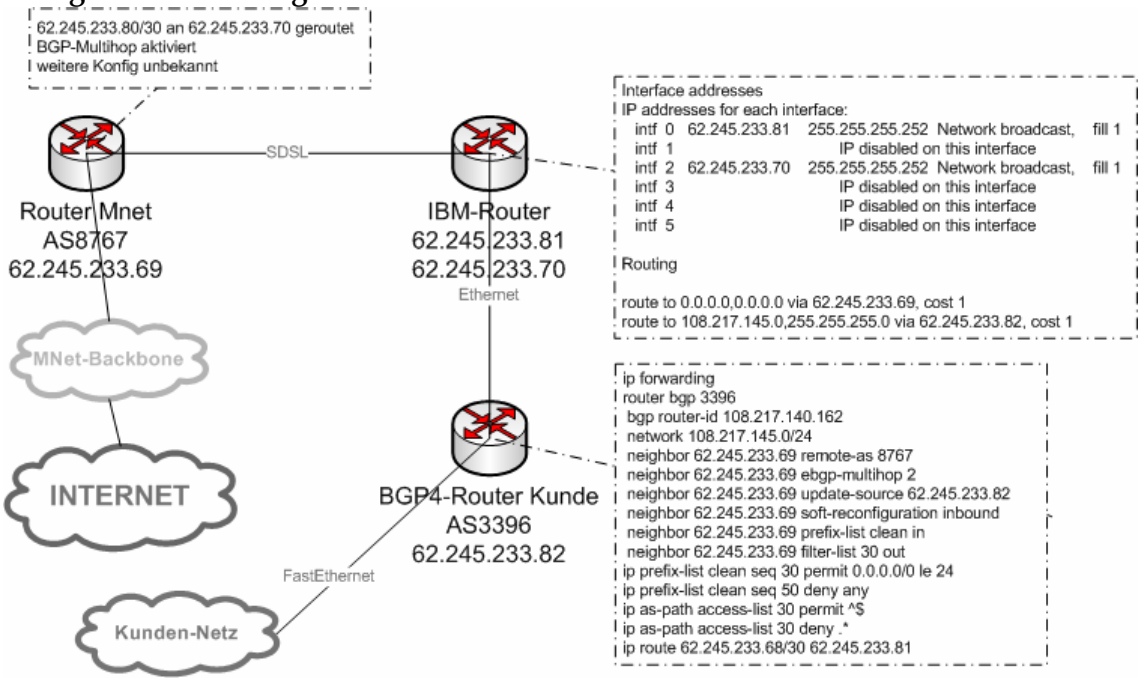
Anlage 6 – Grobübersicht Router



Anlage 7 – Anbindungsschema TMT



Anlage 8 – Anbindungsschema Mnet



Anlage 9 – Quellen

Die Kenntnisse, welche zur Projektrealisierung nötig waren, habe ich mir unter anderem durch das Studium folgender Informationsmaterialien angeeignet:

- CCNA V 3.1 Kursmaterial Semester 1-3 auf http://cisco.netacad.net
- Suchfunktion auf http://www.cisco.com
- Doyel, J. und DeHaven Carrel, J.: Routing TCP/IP Band II, München 2002, Markt + Technik Verlag (Originalausgabe: Cisco Press)
- Halabi, Bassam/ McPherson, Danny: Internet Routing Architekturen – Verbindungen und Protokolle professionell planen und realisieren, Übersetzung: Cosmos Consulting, München 2001, Markt + Technik Verlag (Orginalausgabe: Cisco Press)
- Lewis, Chris: Cisco TCP/IP-Routing, 2. überarbeitete Auflage, Übersetzung: Stiels, Rebecca und Mauser, Matthias, München 2001, Addison-Wesley Verlag
- Hunt, Craig: TCP/IP Netzwerk-Administration, 2.Aufl., Übersetzung und deutsche Bearbeitung: Klicman, Peter, Köln 1998, O´Reilly Verlag
- Mason, Andrew G. und Newcomb, Mark J.: Cisco Secure – Sicherheitslösungen für das Internet, Übersetzung und Lokalsierung: Ring, Uwe, München 2002, Markt + Technik Verlag (Originalausgabe: Cisco Press)
- Akin, Thomas: Hardening Cisco Routers, First Edition, Sebastopol, CA 95472 2002, O´Reilly & Associates, Inc.
- Ballew, Scott M.: Managing IP Networks with Cisco Routers, First Edition, Sebastopol, CA 95472 1997, O´Reilly & Associates, Inc.

Anlage 10 – CD-Inhalt

Folgende Dokumente sind auf der beiliegenden CD enthalten:

- c1600_pl.pdf – Produktinformationen zur Cisco 1600 Serie
- 2600_qsg.pdf – Produktinformationen zur Cisco 2600 XM Serie und Cisco 2612 Router
- quagga_20050512.pdf – Quagga Info page
- rfc950.txt – Internet Standard Subnetting Procedure
- rfc1122.txt – Requirements for Internet Hosts -- Communication Layers
- rfc1518.txt – An Architecture for IP Address Allocation with CIDR
- rfc1519.txt – Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy
- rfc1771.txt – A Border Gateway Protocol 4 (BGP-4)
- rfc1930.txt – Guidelines for creation, selection, and registration of an Autonomous System (AS)
- rfc2661.txt – Layer Two Tunneling Protocol "L2TP"
- ripe-252.pdf – RIPE Database Reference Manual
- ripe-263.pdf – Autonomous System (AS) Number Assignment Policies and Procedures
- ripe-324.pdf – IPv4 Address Allocation and Assignment Policies for the RIPE NCC Service Region
- ripe-as.txt – Ausgabe von „whois AS3396“
- ripe-prefix.txt – Ausgabe von “whois 108.217.145.0/24” incl. Route-Object
- Projektdokumentation.pdf – diese Dokumentation

Anlage 11 – Selbstständigkeitserklärung

Hiermit versichere ich, dass ich diese Projektdokumentation selbstständig und unter ausschließlicher Nutzung der angegebenen Quellen und Hilfsmittel erstellt habe.
Sollten Teilarbeiten, welche zur Projektrealisierung notwendig waren, durch weitere Personen durchgeführt worden sein, so ist dies ausdrücklich gekennzeichnet.

Jan Wagner