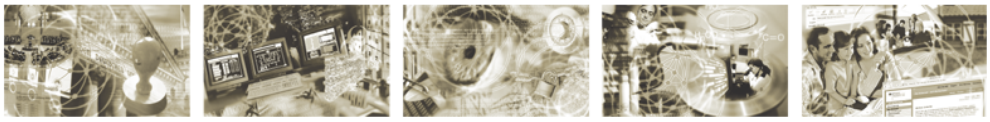
Bundesamt
für Sicherheit in der
Informationstechnik



Band G, Kapitel 3: Netzwerk

Bundesamt für Sicherheit in der Informationstechnik
Postfach 20 03 63
53133 Bonn
Tel.: +49 22899 9582-0
E-Mail: hochverfügbarkeit@bsi.bund.de
Internet: https://www.bsi.bund.de

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1 Netzwerkarchitektur

Das wirkungsvolle Zusammenspiel von IT-Komponenten in einem Netzwerk ist fast immer zur Durchführung von Geschäftsprozessen erforderlich. Das Netzwerk bildet dabei die notwendige Grundlage, um mit anderen Stellen zu kommunizieren, auf Dienste zuzugreifen oder Daten zu speichern.

Da eine große Zahl von unterschiedlichen Technologien, Komponenten, Übertragungswegen und -arten in modernen Netzen verwendet wird, ergibt sich bei der Planung hochverfügbarer Anwendungen eine hohe Herausforderung, die Prinzipien der Hochverfügbarkeit systematisch und konsequent umzusetzen (siehe Beitrag „Prinzipien der Verfügbarkeit“ im HV-Kompendium).

Das elementare Ziel der Netzplanung bezüglich HV muss sein, einen SPoF zu vermeiden. Die Daten durchlaufen in Netzen auf ihrem Weg vom Ausgangspunkt (Source, Quelle) zum Zielpunkt (Destination, Senke) oft eine große Zahl von Einzelelementen. Dabei kann bei jedem dieser Elemente eine Unterbrechung entstehen, wenn nicht durch die Nutzung alternativer Wege die Kommunikation aufrechterhalten werden kann.

Die redundante Auslegung aller Netzwerkkomponenten muss aber nicht in allen Fällen erforderlich sein. Das Prinzip der Priorisierung (ausgehend von den Geschäftsprozessen) führt zu einer wirtschaftlich gerechtfertigten hochverfügbaren Ausprägung. Ein Beispiel hierfür ist die redundante Ausführung der Komponenten zur Aufrechterhaltung der Außenkommunikation.

Neben der reinen Erreichbarkeit des Zielpunktes für die Daten ist die Einhaltung gewisser Qualitätsmerkmale der Netzverbindungen notwendig. Aus diesem Grunde fließen in die Planung der HV-Maßnahmen Parameter wie die dauerhaft garantierte mittlere Durchsatzleistung, die Latenz und die Bitfehlerrate einer Verbindung als Leistungsmerkmale ein.

Die Anforderung, unter welchen normalen und besonderen äußeren Bedingungen (Lagen) die Netzkommunikation garantiert sein muss, charakterisiert die geforderte Robustheit des Netzes.

Unter diesen Randbedingungen wird letztlich das Netzwerk implementiert oder erweitert, um die Daten über die erforderlichen Entfernungen verlässlich zu transportieren.

In diesem Beitrag werden die Standard-Maßnahmen zur Steigerung der Verfügbarkeit, wie sie z. B. in den IT-Grundschutzkatalogen aufgeführt sind, als bekannt vorausgesetzt und nicht weiter ausgeführt. Beispielsweise sind die sorgfältige Dokumentation eines Netzwerkes oder die klare Festschreibung von Kompetenzen und Maßnahmen, die sich direkt auf die Verfügbarkeit von Netzwerken auswirken, hier aber nicht weiter beschrieben werden.

1.1 Prinzipien der Verfügbarkeit im Netzwerk

Die folgenden Abschnitte befassen sich mit der Anwendung der Prinzipien der HV insbesondere auf Netzwerke.

Die grundsätzlichen Prinzipien der HV finden sich auch bei der Implementierung von Netzwerken wieder. Dies wird in den nächsten Abschnitten anhand der Schichten des OSI-Modells genauer erläutert. Übergreifend sollten vorab folgende Prinzipien aufgeführt werden:

Robustheit/Fehlertoleranz: Die Auswahl von Netzkomponenten für HV-Umgebungen bevorzugt qualitativ hochwertige Komponenten bezüglich der zugedachten Funktionalität, auf die sie minimalisiert (Härten) sein müssen. Ungeplant integrierte Bausteine, die diesen hohen Ansprüchen nicht genügen (z. B. billige Hubs, komplexe teure Server für Spezialaufgaben mit nicht erforderlicher Funktionalität), gefährden die Zielstellung hinsichtlich der Verfügbarkeit. Hochwertige Netzkomponenten besitzen oft schon integrierte Redundanzen (z. B. doppelt ausgeführte Netzteile) und sind robuster gegenüber Spitzenlasten im Netzverkehr.

Skalierbarkeit: Ein gut geplantes und implementiertes Netz sollte von sich aus schon skalierbar sein. Die Anforderungen an die Netze wachsen gegenwärtig exponentiell mit der Zeit, neue Applikationen und Dienste verschlingen immer mehr Bandbreite. Dies lässt sich auf längere Sicht nur mit skalierbaren Konzepten erfüllen.

Autonomie: Zurzeit gibt es in der Praxis noch keine Realisierung von echter Autonomie. Es existieren in der Forschung vielversprechende Ansätze, wie z. B. sich selbst organisierende Netzwerke (Ad-hoc-Netzwerke).

Virtualisierung: Dieses Prinzip findet auch in Netzwerken eine hohe Verbreitung. So werden Netzwerke untereinander virtuell abgetrennt (Virtuelles LAN, sog. VLAN) um Überlastungen zu vermeiden und die Verfügbarkeit zu erhöhen. Absicherung gegen Angriffe oder auch Überflutung von außen bieten sog. „Virtuelle Private Netze“ (VPN). Virtuelle Maschinen simulieren ein Netzwerk mittels Software und steigern die Verfügbarkeit durch den Verzicht auf weitere fehleranfällige Hardware. Inzwischen gibt es Anstrengungen, ganze Netzwerkkomponenten sowie deren Verbindungen in eigens dafür konzipierten Komponenten zu virtualisieren.

1.2 Verbindungskette in einer typischen Netzwerkarchitektur

Entsprechend dem Grundsatz, den Ausfall jeder einzelnen Komponente in einer Verbindungskette in Betracht zu ziehen und für diesen Fall Redundanzen anzubieten, sollte bei der Planung zunächst der komplette Weg von der Quelle zur Senke analysiert werden. Der Weg, den die Daten in einer typischen Netzwerkarchitektur durchlaufen, stellt sich wie folgt dar:

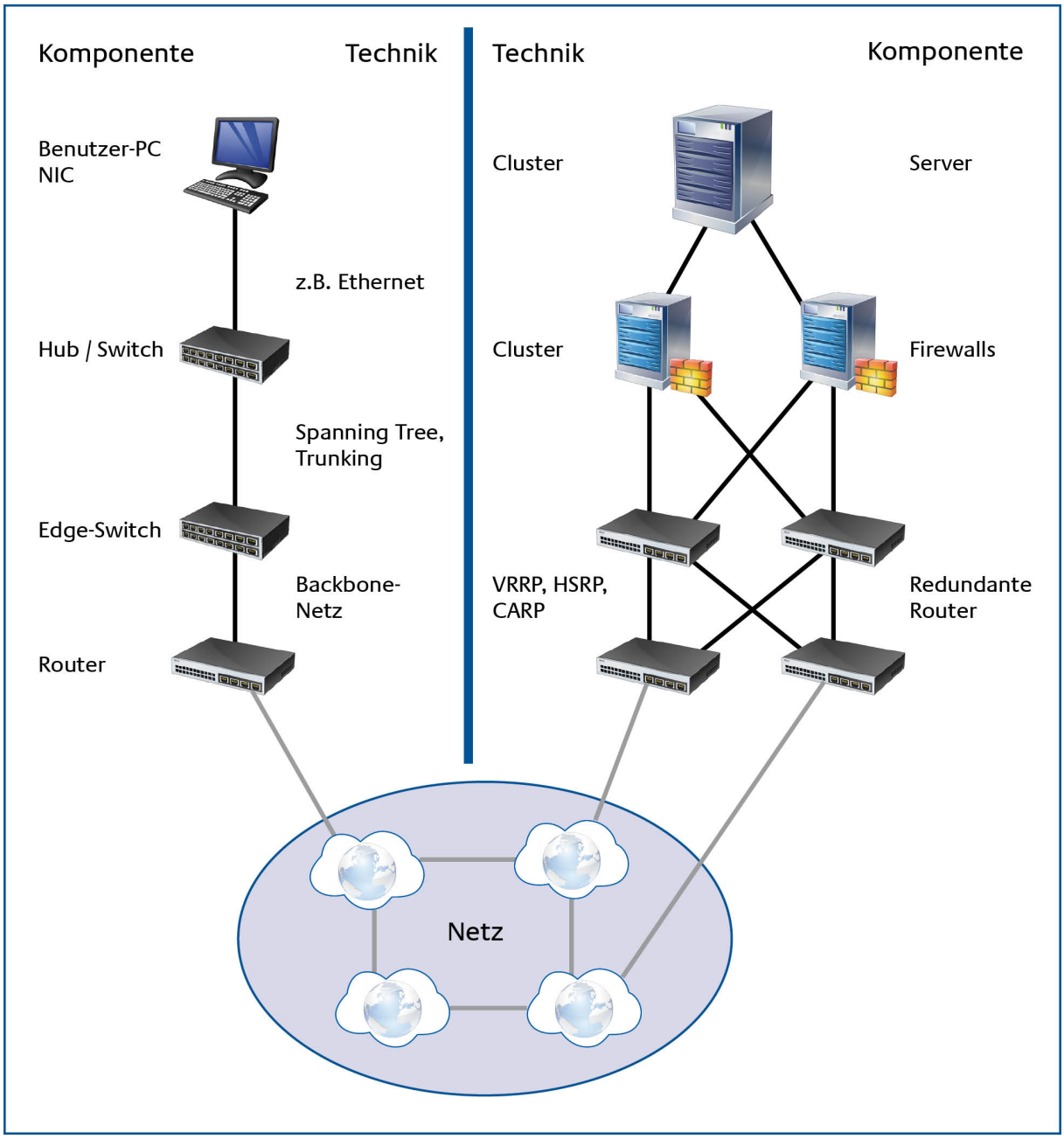


Abbildung 1: Physikalischer Weg, den die Daten in einer typischen Architektur durchlaufen

1. Absenden der Daten vom Quellsystem durch ein logisches Netzinterface an eine im IT-System eingebaute Netzkarte (engl. NIC, Network Interface Card).
2. Übertragung der Daten durch ein lokales Datenübertragungsmedium (LAN). Dies geschieht häufig über ein Kabel der Kategorie CAT-5-6-7 mit den Protokollen Ethernet, Fast-Ethernet oder Gigabit-Ethernet.
3. Verteilung zum nächsten aktiven Netzgerät durch ein aktives Netzgerät wie einen Switch, eine Bridge oder einen Hub.

4. Weiterverteilung an weitere LAN-Komponenten über ein geeignetes Ethernet-Protokoll, ggf. durch LAN-Backbone-Netze wie FDDI oder ATM durch Etagen- oder Edge-Switches.
5. Zur Strukturierung von Netzen und zur Verkehrsleitung werden Router eingesetzt, die Datenpakete beispielsweise in andere Gebäude oder an komplett andere Standorte weiterleiten.
6. Zur Anbindung eines Standortes werden verschiedene Techniken der WAN-Anbindung genutzt, dazu gehören Standleitungen, ADSL- und SDSL-Anschlüsse und verschiedene andere Technologien. Teilweise spielen auf dieser Strecke auch Wählverbindungen über analoge Telefonleitungen oder ISDN eine Rolle. Über diesen Weg werden die Daten zum nächsten Vermittlungspunkt geleitet.
7. Vermittlungspunkte werden durch spezielle Dienstleister, nämlich Internet Service Provider (ISPs) betrieben.
8. ISPs kommunizieren untereinander über weitere WAN-Leitungen entweder direkt oder über nationale oder internationale Exchange-Points.
9. Ab diesem Punkt wiederholen sich die einzelnen durchlaufenen Stellen in umgekehrter Reihenfolge, bis die Daten schließlich das Zielsystem, die Datensenke erreicht haben.

Nicht alle Daten durchlaufen den kompletten Weg durch alle hier beschriebenen Instanzen. Wenn sich Quelle und Senke am gleichen Standort befinden, durchlaufen die Daten im Allgemeinen nur die Stufen 1 – 5 und umgekehrt.

Häufig werden noch zusätzliche Netz-Komponenten durchlaufen, die nicht in erster Linie dem Transport der Daten dienen, sondern andere Qualitätsmerkmale bereitstellen oder sichern. Dazu gehören unter anderem:

- Firewalls, die das Durchleiten illegitimer Daten verhindern (Stichwort: Zugriff, Access),
- VPN-Gateways, die durch Einsatz von Verschlüsselungstechnologien die Offenlegung der übertragenen Inhalte verhindern (Stichwort: Vertraulichkeit, Privacy),
- Applikationsgateways (etwa Proxys), die zu einem gewissen Grad die Inhalte der durchgeleiteten Daten auf Konformität mit festgelegten Richtlinien überprüfen. So können etwa unerwünschte Inhalte, Viren oder Würmer abgewehrt werden (Stichwort: Integrität, Integrity).

Zusätzlich werden auf dem Weg von der Quelle zum Ziel häufig Hilfsdienste in Anspruch genommen. Diese Hilfsdienste befinden sich nicht unmittelbar in der Verbindungskette, sie sind jedoch für die Adressierung oder Weiterleitung der Daten notwendig. Dazu zählen unter anderem:

- DHCP-Dienst, um dem lokalen Interface eine Adresse zuzuweisen.
- ARP-Dienst, um die MAC-Adresse eines Netzwerkknotens im LAN festzustellen.
- DNS-Dienst, um die IP-Adresse des adressierten Kommunikationspartners festzustellen.
- Routing-Dienst über diverse Routing-Protokolle, darunter RIP, OSPF und BGP, um die Wege innerhalb des Netzes zum Zielnetz festzustellen.

In jedem dieser Teilbereiche liegen Gefahren für die Verfügbarkeit der Datenübertragung, wenn die Daten den jeweiligen Teilbereich durchlaufen. Ist ein Teilbereich nicht verfügbar, so ist die gesamte Verbindungskette ausgefallen. Aus dieser Überlegung ergibt sich unmittelbar die Folgerung, die Anzahl der beteiligten Komponenten in der Verbindungskette zwischen Quelle und Senke so gering wie möglich zu gestalten, um eine hohe Verfügbarkeit sicherzustellen.

1.3 Realisierung von HV in den Schichten des Netzwerkes

In den folgenden Abschnitten erfolgt orientiert an dem ISO/OSI-Schichtenmodell eine Darstellung der in der Verbindungskette vorkommenden Komponenten. Es werden Möglichkeiten für Redundanzen beschrieben. Diese Redundanzen können durch die Nutzung geeigneter Netzwerkprotokolle und durch Mechanismen für die automatische Wegewahl im Fehlerfall zu einer Fehlertoleranz des Systems führen.

1.3.1 Schicht 1: Übertragungsebene

Die Übertragungskomponenten umfassen die Codierung und den tatsächlichen Transport der Daten von einem Ort zum anderen von den Ebenen 0 und 1 des ISO/OSI-Schichtenmodells.

Hierbei kommen verschiedene Medien und Protokolle zum Einsatz, welche unterschiedliche Auswirkungen auf die HV-Gesamtarchitektur haben.

1.3.1.1 LAN

Neben dem Einsatz von hochwertigen Kabeln (z. B. CAT-5-6-7) sollte insbesondere in Rechenzentren der Einsatz von strukturierter Verkabelung gemäß EN50173-5 zum Einsatz kommen. Dies dient nicht nur einer besseren Übersichtlichkeit, es minimiert auch die Anzahl der verlegten Kabel und damit der Fehlerquellen.

Zur Realisierung der Robustheit ist es in HV-Umgebungen erforderlich, eine redundante Verkabelung zu realisieren. Die Leitungen sollten dabei möglichst über unterschiedliche Wegstrecken verlegt werden (Separation), um die Nichtverfügbarkeit durch mechanische Unterbrechungen zu minimieren. So könnten z. B. Erdbauarbeiten ein Kabel durchtrennen oder es könnte ein Brand in einem Gebäudeteil, in dem sich die Datenleitungen befinden, die Netzanbindung unterbrechen. Voraussetzung für eine redundante Kabelführung ist der Einsatz von Mehrfach-Netzwerkkarten, diese werden gesondert im Kapitel Server (Querverweis) behandelt. Neben der redundanten Kabelführung ist es notwendig, diesen Umstand auf höheren Schichten des ISO/OSI-Schichtenmodells zu berücksichtigen (vgl. Abs. 1.3.3), damit mittels Automatismen (z. B. durch die Wahl geeigneter Protokolle) diese Redundanzen auch zum Tragen kommen.

Eine gesonderte Betrachtung kommt verschiedenen kabellosen Übertragungstechniken (insbesondere Wireless LAN nach IEEE 802.11a/b/g und WiMax) zu. Auch Richtfunkstrecken, extraterrestrischer Funk über Satellitenübertragungen oder laser- oder mikrowellengestützte Verfahren gehören zu diesem Bereich. Für die normalen HV-Lösungen kommt der Einsatz solcher Lösungen jedoch nur in Sonderfällen in Betracht, da, obgleich die durchschnittlichen Datenraten teilweise durchaus mit kabelgebundener Übertragung von Ethernet oder Fast Ethernet vergleichbar sind, die Verlässlichkeit der Verbindung vielen nur schwer zu kalkulierenden Faktoren unterliegt. So können der Betrieb nicht ordnungsgemäß entstörter Geräte, Veränderungen in Gebäuden (auch im Umfeld der Betriebsumgebung) oder sogar atmosphärische Störungen Einfluss auf die Übertragungsqualität haben. Im Rahmen einer Notfall-Lösung können solche Alternativen durchaus sinnvoll sein, da hier in der Regel andere Infrastrukturen als im Normalfall genutzt werden sollte (vgl. [ALKO08]). So ist z. B. die Kommunikation via Satellit relativ robust gegen weiträumige Stromausfälle oder Naturereignisse. Der Einsatz von Funknetzen (mobile Richtfunkstationen oder auch selbstorganisierende WLAN-Netze) bietet in Notfällen ähnliche Vorteile, da hier ad hoc ein Netzwerk ohne vorhandene Verkabelung aufgebaut werden kann. Umfassende Tests der Qualitätsmerkmale in unterschiedlichen Situationen (etwa bei Nebel, Regen, Gewitter oder während

besonderer Aktivität von Solarprotuberanzen) müssen im Voraus erfolgen (z. B. im Rahmen von Notfallübungen).

Das im lokalen Umfeld (LAN) häufig verwendete Ethernet (IEEE 802.3) ist auf eine gute Kabelqualität entsprechend der eingesetzten Leistungsstufe angewiesen. Grundsätzlich ist Standard-Ethernet durch das zugrunde liegende CMSA/CD-Verfahren bei Bus-Topologien anfällig für Kollisionen und damit für einen Datenverlust, der zumeist durch höhere Protokollebenen ausgeglichen wird. Inzwischen wird hier praktisch nur noch die sternförmige Netztopologie (Twisted pair Ethernet) in Verbindung mit modernen Switches (siehe Abschnitt 1.3.2) eingesetzt.

Im LAN werden gelegentlich noch einige weitere weniger verbreitete Netzprotokolle für Spezialanwendungen verwendet. So sind z. B. vereinzelt noch Netze, die auf Token Ring (IEEE 802.5) basieren, im Einsatz. Solche Technologien haben oft spezifische Eigenschaften, die für Verfügbarkeitsüberlegungen durchaus relevant sind. Aufgrund der Vielzahl von Protokollen und ihrer geringen Relevanz im breiten Einsatz werden diese hier nicht weiter betrachtet.

Sind bei Storage-Clustern harte Garantien für den Netzdurchsatz nötig, so können spezialisierte Übertragungsprotokolle wie Fibre Channel oder iSCSI zum Einsatz kommen, die solche Qualitätsmerkmale (QoS) besser garantieren können. Diese Protokolle werden in den nachfolgenden Abschnitten näher beschrieben, in denen die zugehörigen Speichertechnologien und Hardware-Server-Architekturen untersucht werden (siehe Beitrag „Speichertechnologien“ im HV-Kompendium).

1.3.1.2 WAN

Eine hochverfügbare Anbindung an das WAN sollte redundant über getrennte Wege geführt werden. Der Ausfall einer Verbindung wird dann normalerweise auf höherer Protokollebene aufgefangen.

Im WAN-Bereich werden heute praktisch nur noch Glasfaserverbindungen eingesetzt. Eine Ausnahme bilden die oben angesprochenen Verfahren per Satellit oder Richtfunk. Über die Glasfaser werden zum Transport unterschiedliche Übertragungsprotokolle genutzt.

Dark-Fibre (*dt. „dunkle Faser“*) ist eine LWL-Leitung[1] die unbeschaltet, also ohne Endgeräte, verkauft oder vermietet wird. Die Faser erlaubt dabei eine Punkt-zu-Punkt Verbindung zwischen zwei Standorten. Für die Nutzung dieser Verbindung müssen an den beiden Standorten entsprechende Endgeräte angeschlossen werden, die eine Datenübertragung über Lichtwellenleiter erlauben.

Die Dark-Fibre besteht meist aus einer Singlemode-Faser (auch Monomode genannt), wobei der Kern einen solch kleinen Durchmesser besitzt, dass sich nur eine Mode (ein "Strahl") ausbreiten kann. Dadurch sind Singlemode-Faser optimal für lange Strecken und sehr hohe Übertragungsraten geeignet.

Für die Überbrückung von kurzen Distanzen oder für die Inhouse LAN Verkabelung werden statt den Singlemode-Fasern vorwiegend sogenannte Multimode-Fasern verwendet. Diese Fasern verfügen über einen deutlich größeren Kerndurchmesser, sodass hier die optische Ansteuerung in den Endgeräten auch mit herkömmlichen LEDs (Light Emitting Diode, Leuchtdioden) statt den deutlich teureren Laser Dioden, die für die Singlemode-Faser erforderlich sind, verwendet werden können.

1 LWL-Leitung = Lichtwellenleiter-Leitung

Die maximale Reichweite ergibt sich zum einen aus dem verwendeten Faser Typ, dem gewählten Übertragungsprotokoll, den optischen Verlusten auf der gesamten Strecke. (Dämpfung), der Leitung des Senders bzw. der Empfindlichkeit des Empfängers sowie den für die Übertragung verwendeten Wellenlängenbereich. Im MAN- (Metropolitan Area Network) Bereich bieten sich verschiedene Varianten des Ethernet-Protokolls an. Dabei sind typischerweise die folgenden Reichweiten ohne einen zwischengeschalteten Repeater möglich:

- Multimode (MM): 850 nm, bis 1,5 km
- Singlemode (SM): 1310 nm, bis 10 km
- Singlemode (SM): 1550 nm, bis 100 km

City-LANs: Unter dem Begriff City-LAN wird ein Weitverkehrsnetz mit einer lokalen Ausdehnung auf ein Stadtgebiet oder einen entsprechenden Ballungsraum verstanden. Da innerhalb eines Stadtgebietes meist ein hoher Bedarf an ausreichend dimensionierter Kommunikationsinfrastruktur existiert, gibt es in der Regel eine Vielzahl von Anbietern die Hochgeschwindigkeitsnetze auf der Basis von ATM, MPLS, SDMS, WDM oder weiteren Protokollen und Technologien anbieten.

Die Auswahl an Providern und das ausreichende Angebot von Hochgeschwindigkeitsleitungen sind optimale Voraussetzungen für die Einrichtung von redundanten Rechenzentren. Neben einer direkten Punkt-zu-Punkt Verbindung zwischen zwei geografisch verteilten Rechenzentren über eine Dark-Fibre besteht häufig auch noch die Möglichkeit einer Anschaltung an die bestehende City LAN Infrastruktur eines Providers. Durch die in Ballungsgebieten relativ kurzen Wege < 15km zu einen City LAN Provider kann das Risiko eines Ausfalls oder einer Störung bei einer längeren Punkt-zu-Punkt Verbindung minimiert werden, da die Verbindungen innerhalb des City LANs meist als mehrfach redundante Ringe realisiert sind.

SDH: Eine weitere Option für Weitverkehrsnetze ist die SDH - *Synchrone Digitale Hierarchie*. SDH ist ein Übertragungsverfahren auf Basis von redundanten Glasfaserringen, mit denen ausfallsichere Transportverbindungen realisiert werden können. Das SDH-Verfahren ist eine Weiterentwicklung des früheren PDH - Plesiochrone Digitale Hierarchie Verfahrens, welches ursprünglich für die Digitalisierung des öffentlichen analogen Telefonnetzes eingeführt wurde.

Ein wesentliches Merkmal der SDH-Technik ist dabei die automatische Umschaltung im Fehlerfall. Ein redundanter Ring dient dabei als Ersatzweg. Bei einer Störung des primären Ringes schaltet das APS (Automatic Protection System) vom primären Ring auf den redundanten Ring. Diese Topologie ist unter der Bezeichnung 4-Faser MS-SPRing (Multi-Section-Shared-Protection-Ring) ab STM-16 aufwärts standardisiert. Nach dem SDH-Standard müssen die Ersatzschaltmaßnahmen nach dem Erkennen einer Störung automatisch innerhalb von 50 Millisekunden abgeschlossen sein.

Mit SDH können neben Sprachdaten auch ATM-Zellen und IP-Daten über ein international verknüpftes Netz übertragen werden. In der Netzstruktur sind meist vordefinierte Ersatzstrecken für den Fehlerfall definiert, wobei innerhalb von Sekunden nach dem Ausfall eines Verbindungsabschnitts automatisch auf eine Alternativstrecke umgeschaltet werden kann.

WDM: Das Wellenlängenmultiplexverfahren (*engl.: Wavelength Division Multiplexing*) ist ein optisches Frequenzmultiplexverfahren, das bei der Übertragung von Daten (Signalen) über ein Glasfaserkabel verwendet wird. Beim Wellenlängenmultiplexverfahren werden die aus verschiedenen Spektralfarben bestehende Lichtsignale zur Übertragung in einem Lichtwellenleiter verwendet.

Jede dieser so erzeugten Spektralfarben bildet somit einen eigenen Übertragungskanal, auf den man die Daten (Signale) eines Senders modulieren kann. Die so modulierten Daten (Signale) werden dann durch optische Koppelelemente gebündelt und gleichzeitig, unabhängig voneinander, übertragen. Am Ziel dieser optischen Multiplexverbindung werden die einzelnen optischen Übertragungskanäle durch optische Filter oder wellenlängensensible opto-elektronische Empfängerelemente wieder getrennt.

Die Interface-Module der WDM-Endgeräte unterstützen meist alle gängigen Schnittstellen. Damit können alle bisherigen Systeme, die sowohl optisch als elektrisch arbeiten, ihre Signale über WDM-Strecken und Ringe übertragen. Auch lassen sich damit Verbindungen über Fibre Channel und ESCON-Systeme über sehr weite Entfernungen realisieren.

Obwohl dieses Verfahren z. B. eine angemietete Dark-Fibre zwischen zwei Standorten optimal nutzen kann, bringt dieses Verfahren im Hinblick auf eine Hochverfügbarkeit keine Vorteile. Durch dieses Verfahren kann unter Umständen auch das Bewusstsein dafür verloren gehen, dass es sich bei dem physischem Medium immer nur noch um eine einzige Glasfaser handelt, obwohl an den Endgeräten das vorhanden sein von mehrere Verbindungen suggeriert wird.

1.3.2 Schicht 2: Verbindungsebene

Verbindungskomponenten sorgen dafür, dass die einzelnen aktiven Netzgeräte über entsprechende Übertragungsmedien miteinander verbunden sind. Dieses sind Geräte der OSI-Schicht 2. Zu ihnen zählen prinzipiell Repeater, Hubs, Bridges und Switches. Funktional dienen alle Geräte dazu, Datenframes von einem Endpunkt zu einem anderen zu transportieren. Um die beim Ethernet technologieimmanente Gefahr der Kollisionen von Datenpaketen zu vermeiden, werden heute fast nur noch Switches im full-duplex-Betrieb eingesetzt. Diese stellen praktisch eine Punkt-zu-Punkt-Verbindung zwischen den Netzgeräten her.

Viele Cluster-Technologien (Querverweis) erfordern, dass die entsprechenden Server sich in einer Broadcast-Domäne befinden, damit die Automatismen entsprechend greifen. Die Automatismen zur Überbrückung eines Ausfalls einer Verbindung oder einer Komponente müssen dementsprechend schon auf Ebene 2 des OSI-Protokolls greifen.

Eine Reihe von prinzipbedingten Gefährdungen bei Repeatern, Hubs und Bridges werden von den erheblich moderneren Switches konzeptionell bereits deutlich entschärft. Hierzu gehören z. B. die oben genannte Kollision von Datenpaketen oder die Flutung von Netzten mit Broadcastpaketen. Aus diesem Grunde sollten in HV-Umgebungen weitgehend leistungsstarke und robuste Switches zum Einsatz kommen. Diese bieten weiterhin auch die Möglichkeit, virtuelle Netze zu definieren, um den Datenfluss logisch und physikalisch zu trennen (Virtualisierung).

Dennoch ist auch ein moderner Switch gewissen Einschränkungen unterworfen. Die meisten Geräte machen gewisse statistische Annahmen über die Anzahl, Größe und Geschwindigkeit von Paketen, die über sie verarbeitet werden können. Diese liegen häufig über den theoretischen maximalen Datenraten, die über die Medien zugeführt werden können. So kann beispielsweise die Zuordnungstabelle der MAC-Adressen an einzelnen Ports eines Switches durch das bewusste Zusenden von vielen zufälligen Adressen zum Überlaufen gebracht werden, was die Funktionsfähigkeit oder zumindest die Durchsatzparameter eines Switches deutlich beeinflussen kann. Regelmäßiges Update und die Beachtung allgemeiner Sicherheitsmaßnahmen können hier helfen.

Topologisch sind geswitchte Netze zunächst sternförmig angeordnet, d. h., alle Teilnehmer im Segment sind über einen Switch zentral miteinander gekoppelt. Dies ist jedoch die klassische Form

eines SPoF und somit eigentlich zu vermeiden. Wenn nicht jedes einzelne Endgerät mittels redundanter Verkabelung an unterschiedliche Switches angeschlossen werden kann (vgl. Abschnitt 1.3.1), so sollten doch zumindest hinreichend viele Endgeräte über jeweils unterschiedliche Switches angebunden sein, um durch Wechsel des Arbeitsplatzes auf eine Anwendung zugreifen zu können. Für Server-Systeme sollten pro IT-System mehrere Netzanschlüsse vorgesehen sein, die über dedizierte Kabel zu unabhängigen Switches führen. Die nachfolgende Abbildung 2 veranschaulicht eine mögliche Topologie zur Vermeidung von SpoFs.

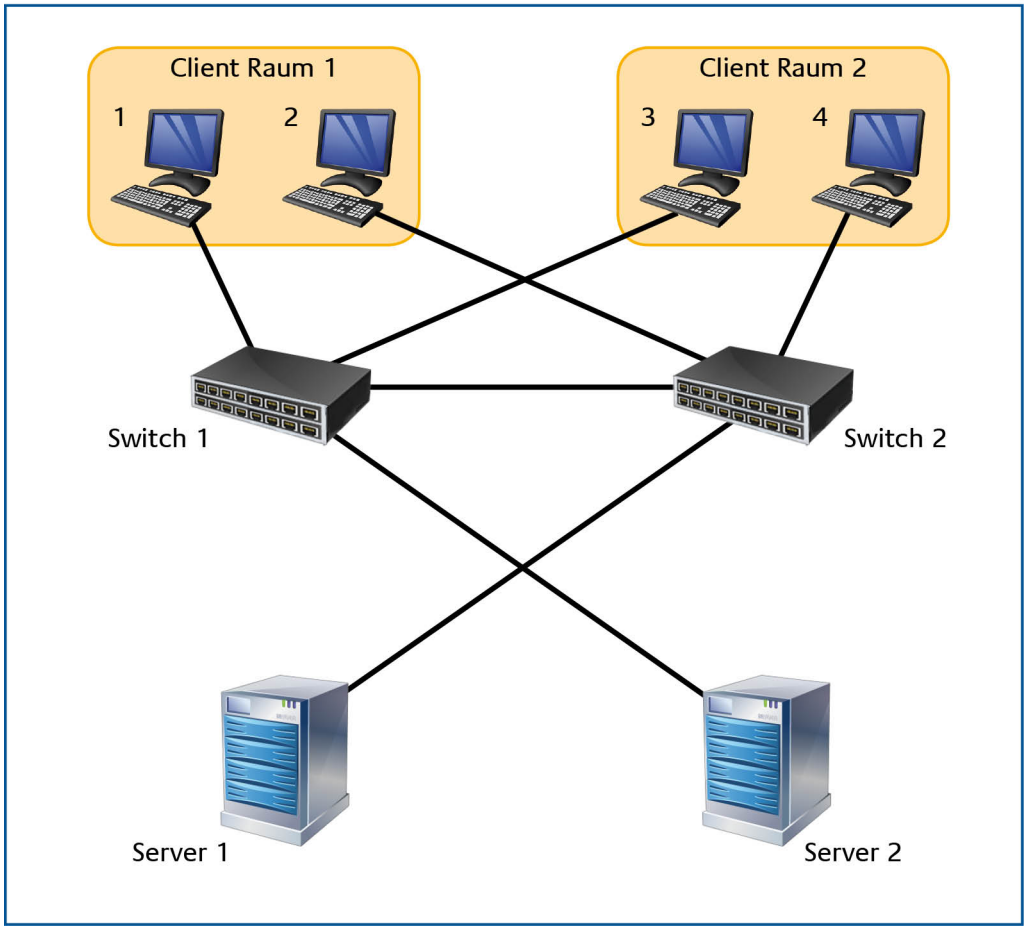


Abbildung 2: Redundante Verkabelung von Server und Benutzer-PC

Weiterhin treten aufgrund der hohen Beanspruchung gelegentlich Hardware-Ermüdungen insbesondere im Bereich der Signalwandler auf, die die digitalen Signale eines Frames in HF- oder Lichtsignale des Mediums umwandeln müssen. Somit können einzelne Switch-Ports ausfallen.

Dieses Problem tritt übrigens nicht nur bei den Koppelgeräten zwischen den Netzgeräten auf, sondern kann durchaus auch bei den einzelnen Netzkarten (NIC) geschehen. So bieten einige Hersteller mittlerweile umfangreiche Diagnose-Schnittstellen in ihren NICs an, um die Funktionsbereitschaft (etwa ein Carrier-Signal) an das darüber liegende Betriebssystem weiterzumelden.

Aus den genannten Gründen greifen bei dem Design von Netzgeräten in HV-Umgebungen normalerweise die in Kapitel 2 dargestellten Prinzipien der strukturellen und funktionalen Redundanz (Geräteredundanz) sowie das Prinzip der Pfadredundanz (Mehrpfadigkeit).

Neben dem Standard IEEE802.3 ad (Link Aggregation) bieten verschiedene Hersteller auch proprietäre Lösungen an, um insbesondere breitbandige Uplinks nicht nur über einen einzelnen, sondern über gleich mehrere Ports miteinander zu koppeln. Dieses oft als Trunking genannte Verfahren ermöglicht es dem Switch, die Verbindungen über mehrere Ports und mehrere physikalische Medien gleichzeitig zu führen. Zumeist ist die Logik zum Erkennen und Ausführen eines Fail-Overs bereits in der Hardware realisiert und wird praktisch in Echtzeit übernommen (Hot-Stand-By).

Darüber hinaus ist es auch möglich, Switches miteinander zu koppeln, um so komplexere Netztopologien aufzubauen. Dieses Verfahren wurde insbesondere früher bei Bridges häufig angewandt, ist aber auch noch bei Switches zu finden. So ist es möglich, mehrere alternative Pfade zwischen zwei Netzgeräten anzubieten. Aufgrund der Ethernet-Technologie dürfen Netzsegmente auf der OSI-Ebene 2 jedoch nur die Form eines spannenden Baumes haben, Dies bedeutet, dass zwischen zwei beliebigen Endpunkten immer nur genau eine aktive Verbindung bestehen darf. Andernfalls besteht die Gefahr, dass Pakete unendlich lange im Netz kreisen und somit die Übertragungsmedien immer stärker „verstopfen“ (*engl. network congestion*).

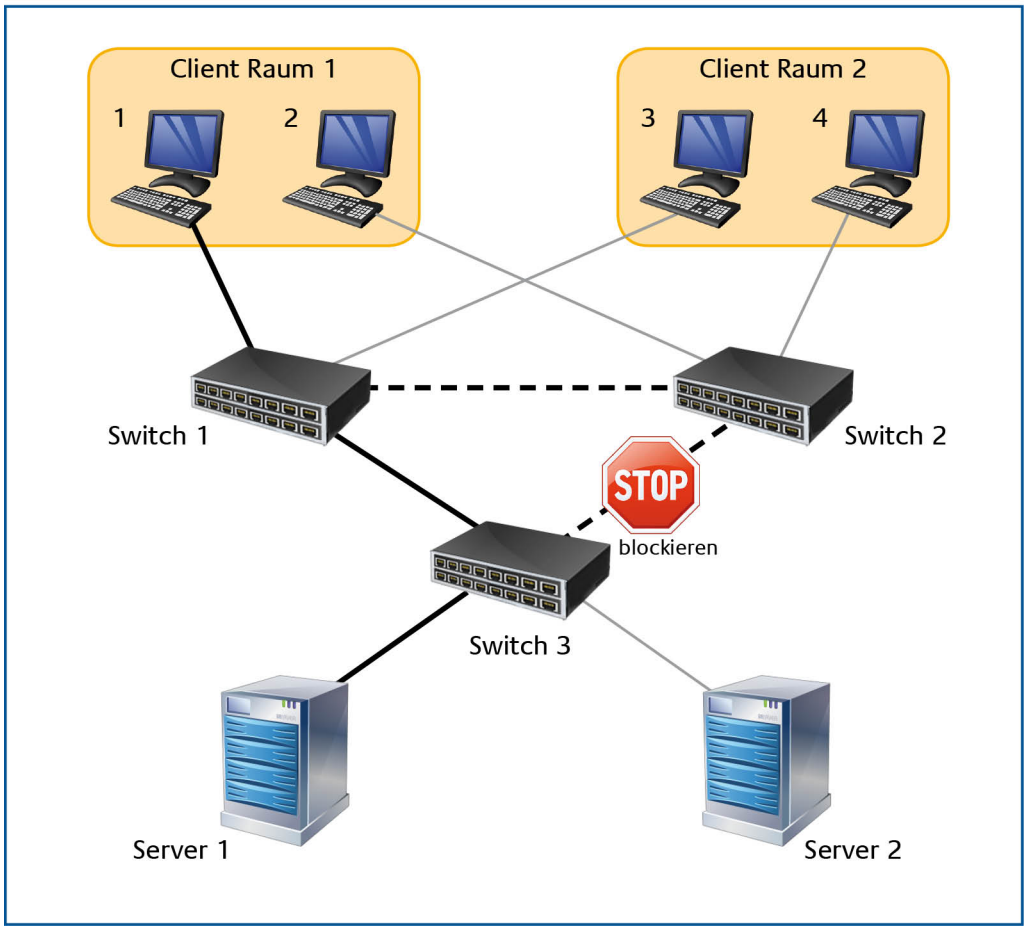


Abbildung 3: Vermaschte Netztopologie

Um dieses Problem zu lösen, wurde das Spanning-Tree-Protocol (STP) in den 80er Jahren entworfen, um solche Schleifen zu erkennen und Verbindungen, die zu ihnen führen, nicht zu nutzen (STP ist in IEEE Std 802.1D-2004 definiert). Weiterhin ist ein Heartbeat-Mechanismus über sogenannte HELLO-Pakete Teil des Spanning Tree Protocols. Werden solche Pakete von einer sie

erwartenden Bridge oder einem Switch nicht mehr empfangen, so nutzt dieses Gerät Alternativpfade. Obwohl das Prinzip durch seine Einfachheit und Funktionalität besticht, wird STP heute aufgrund seiner langsamen Konvergenz und Anfälligkeit gegenüber Manipulationen in HV-Umgebungen nur noch selten eingesetzt. Die Fail-Over-Zeit beträgt bei STP bis zu 30 Sekunden.

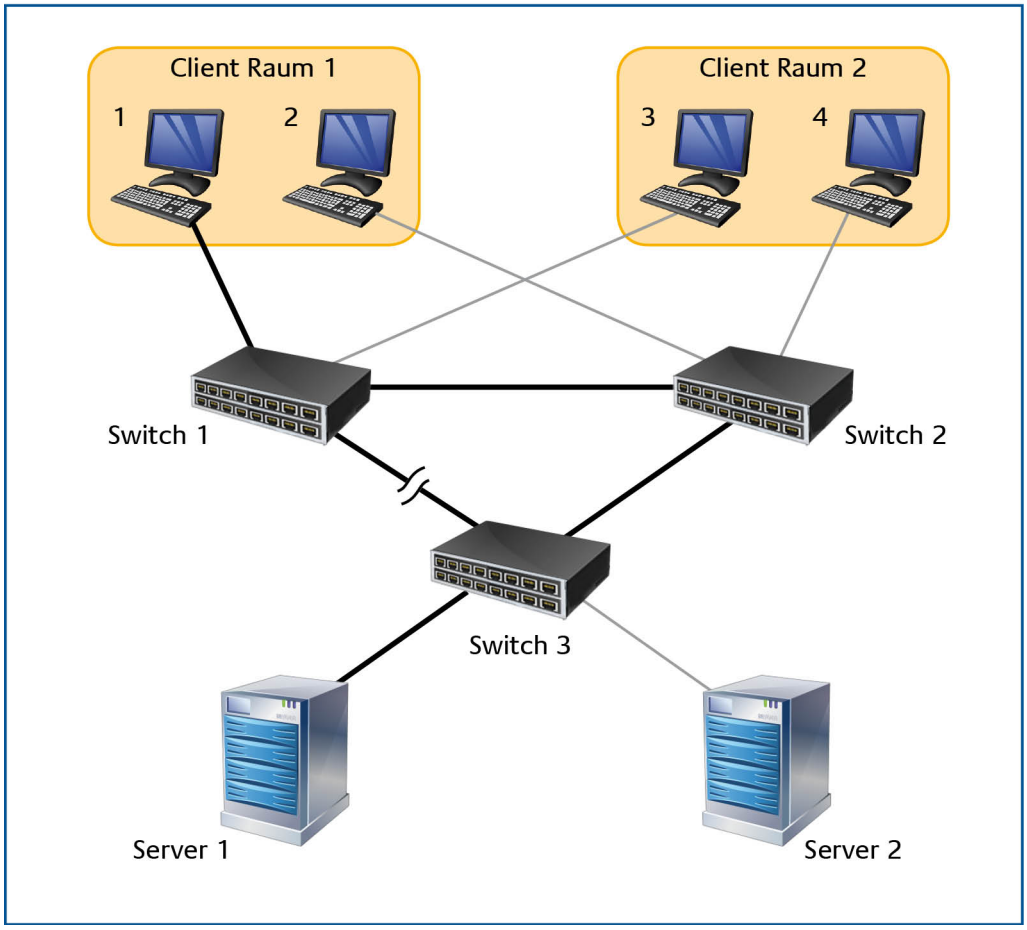


Abbildung 4: Realisierung des Spanning-Tree-Protokolls

STP war von Anfang an nicht als Automatismus für die Aktivierung von Redundanzen konzipiert. Der Einsatz von STP kann in Störungsfällen zu ungeplanten Umkonfigurationen führen, die Fehlersuche ist in diesem Fall in komplexen Netzwerken schwierig und eventuell auch zeitaufwendig.

Von verschiedenen Herstellern wurden ähnliche, verbesserte Alternativen vorgeschlagen, um diese Probleme deutlich zu reduzieren. So wurden mit RSTP (IEEE 802.1w) und MSTP (IEEE 802.1s, besonders im Umfeld des Einsatzes von VLANs) entsprechend leistungsfähigere Nachfolger entworfen, die eine durchschnittliche Fail-Over-Zeit von circa einer Sekunde versprechen. Abhängig von den Anforderungen kann auch hier das Prinzip der Link-Aggregation (s. o.) genutzt werden.

1.3.3 Schicht 3: Vermittlungsebene

Da auf Ebene 3 des OSI-Schichtenmodells (Vermittlungsschicht) den beteiligten Komponenten wesentlich mehr Informationen über Quelle und Ziel eines Datenpaketes zur Verfügung stehen, sind die Komponenten dieser Schicht am besten darauf vorbereitet, die komplexen Fragen der Wegefindung zu beantworten. Solche Geräte sind in der Regel Router, die Fragestellung trifft jedoch

auch auf Firewalls zu. Sie kennen nicht nur die sie unmittelbar umgebende Netztopologie, sondern können auch (etwa über Routingprotokolle wie RIP, OSPF oder BGP) Informationen über weiter entfernte Teilbereiche des Netzes berücksichtigen, sofern diese vorliegen. Ist dies der Fall, so kann ein Router auch flexibel auf Unterbrechungen in bestimmten Teilbereichen des Netzes reagieren und die Daten über einen alternativen Pfad umlenken. Dieses Verfahren wird dynamisches Routing genannt.

Durch das ursprüngliche Designziel des Internets (hohe Fehlertoleranz) ist dieses auch heute noch erstaunlich funktional robust gegen verschiedene Arten von Ausfällen, was jedoch nicht darüber hinwegtäuschen sollte, dass dies mehr aus einer ganzheitlichen Sicht gilt. Bestimmte Qualitätsmerkmale wie Durchsatz, Übertragungsgeschwindigkeit und Latenz können im Allgemeinen im Internet nicht garantiert werden. Diese sind andererseits bei der Spezifikation von Qualitätsmerkmalen einer Verbindung auch und besonders für HV-Architekturen von besonderer Bedeutung.

Eine der wichtigsten Aufgaben der Netzschicht ist die Wegewahl, das Routing. Mit mehreren Millionen zusammengeschlossenen Systemen, die zu einem nicht unerheblichen Grade dezentral verwaltet werden, kann diese Aufgabe der Wegewahl nur in mehreren Hierarchiestufen gelöst werden und dies hat daher besonders hohe Auswirkungen auf die Verfügbarkeit.

Innerhalb einer Organisation, teilweise auch nur in Unterabteilungen, kann die Netztopologie von den IT-Verantwortlichen im Allgemeinen überschaut werden. Innerhalb eines solchen Autonomen Systems (AS) sind die Routing-Regeln oftmals explizit festgelegt oder können in relativ einfachen Tabellen verwaltet werden. Ein AS ist im Kontext des TCP/IP-Netzstacks ein feststehender Begriff, der in RFC 1930 definiert wird (nicht zu verwechseln mit der Begriffsdefinition von autonomen Systemen im Beitrag „Prinzipien der Verfügbarkeit“ des HV-Kompendiums).

Die Wegewahl zwischen Übergabepunkten von Autonomen Systemen wird über die Protokolle aus der Klasse der Exterior Gateway Protocole (EGP) bzw. seiner fast ausschließlich anzutreffenden Realisierung durch das Border Gateway Protocol (BGP) realisiert (s. RFC 1654 (BGP-4)). Durch die Verwendung der Protokolle ist es auch möglich, sich ohne Änderung der IP-Adressen über verschiedene Service Provider Datenpakete zusenden zu lassen (Dynamisches Routing). Fällt ein ISP aus, so wird den anderen Autonomen Systemen via BGP mitgeteilt, dass eine Adresse nun über einen anderen Pfad erreichbar ist. Vorteil dieses Verfahrens ist die hohe Transparenz für die Endgeräte, da keinerlei Adressen oder Dienste neu konfiguriert werden müssen. Nachteil ist die mitunter schwer abschätzbare Fail-Over-Zeit, die je nach Kommunikationspartner im höheren Minutenbereich liegen kann.

Maßnahmen zur Verringerung von Ausfall- und Fail-Over-Zeiten lassen sich zwangsläufig am besten in den Bereichen des Netzes durchführen, auf die noch administrativer Zugriff besteht. Im Idealfall besteht für die wichtigsten Standortverbindungen ein eigenes physikalisches Netz unter eigener Administration. Ansonsten besteht grundsätzlich die Abhängigkeit von Providern und deren Einhaltung von SLAs. So gibt es inzwischen häufig die Möglichkeit, physikalische Glasfaserleitungen (Dark-fibre) direkt anzumieten und zur Übertragung zu nutzen. Im Bereich der eigenen Netzinfrastruktur und teilweise auch noch im Bereich des unmittelbar beauftragten Service Providers können mittels Fail-Over-Protokollen Fail-Over-Zeiten von deutlich unter einer Sekunde pro Systemgruppe erreicht werden. Diese Eigenschaften können natürlich nicht von allen durchlaufenen Teilsystemen im Internet erwartet werden; dies trifft umso mehr zu, wenn der entfernte Kommunikationspartner oder seine Positionierung im Netz nicht a priori bekannt sind, wie dies beispielsweise bei Zugriffen auf ein Web-Portal oder bei VPN-Nutzern im sogenannten Road-Warrior-Betrieb der Fall ist.

Die am häufigsten eingesetzten Protokolle in diesem Bereich sind VRRP (Virtual Router Redundancy Protocol), HSRP (Hot Stand-By Routing Protocol) oder in begrenztem Umfang CARP (Common Address Redundancy Protocol). Das Virtual Router Redundancy Protocol (VRRP) wurde von einem Firmenkonsortium entwickelt und ist heute für viele Netzkomponenten erhältlich (VRRP wird in RFC 3768 beschrieben). Da auf Teile von VRRP Patente bestehen, wurde von einer unabhängigen Gruppe das Common Address Redundancy Protocol (CARP) entworfen, das einige Designschwächen von VRRP umgeht, aber nicht so verbreitet ist (siehe carp(4) im OpenBSD-Manual). Das Hot Stand-By Router Protocol (HSRP) ist eine Entwicklung von Cisco, die aufgrund der hohen Markt-Durchdringung von Systemen dieses Herstellers eine nicht unerhebliche Verbreitung gefunden hat (HSRP wird in RFC 2281 beschrieben). Funktional sind alle Protokolle grundsätzlich erst einmal vergleichbar. Einerseits hat jeder Router in einem durch die Protokolle verwalteten Cluster eine eigene physikalische IP-Adresse, andererseits wird dem Cluster eine weitere virtuelle Adresse zugewiesen. Die Knoten in dem Cluster tauschen untereinander Statusinformationen aus. Damit wird u. a. bestimmt, welcher Knoten der aktive Router ist. Dieser konfiguriert zusätzlich zu seiner physikalischen Adresse noch eine virtuelle Adresse. Nach außen wird allen anderen Netzteilnehmern nur diese virtuelle Adresse bekannt gegeben. Fällt ein aktiver Knoten aus, so erfahren dies die Stand-By-Systeme und ermitteln ein neues Master-System. Somit stellt diese Form des Router-Clusterings eine active-passive Hot-Stand-By-Kopplung dar.

Die drei genannten Fail-Over-Protokolle verfolgen das Ziel, einen funktionalen Router durch Clustering mehrerer redundanter Einzelgeräte hochverfügbar gegenüber einem beliebigen Ausfall eines oder mehrerer dieser Geräte zu machen. Dazu werden zwei oder je nach Verfügbarkeits-anforderungen mehrere funktional gleichartige Router mit jeweils eigenen IP-Adressen in einem Subnetz aufgebaut. Zusätzlich wird eine virtuelle IP-Adresse vergeben, die funktional den gesamten Cluster abbildet, unabhängig davon, welcher Cluster-Knoten gegenwärtig aktiv ist. Die einzelnen Knoten stehen über einem Heartbeat-Protokoll in dauerhaftem Kontakt. Fällt der aktive Knoten aus, so übernimmt ein bis dahin passiver Knoten die IP-Adresse des Clusters. Die Protokolle sind alle einfach einzurichten und robust, da beim Routing vergleichsweise wenig Zustandsinformationen vorgehalten werden müssen bzw. diese von den Routern vergleichsweise einfach wieder anzulernen sind. Auf diese Weise ist eine komplexe Zustandsübertragung zumeist nicht notwendig. Zur Übermittlung des Heartbeat kann eines der eigentlichen Zugangssegmente verwendet werden. Alternativ besteht die Möglichkeit, ein separiertes Heartbeat-Segment zu verwenden. Die VRRP/RSVP/CARP-Lösungen sind jedoch so leichtgewichtig, dass das Zugangssegment zur Synchronisation in den meisten Fällen durchaus verwendet werden kann. Normalerweise koppelt ein solcher Aufbau nämlich ein internes LAN (typischerweise 100 MBit/s oder 1 GBit/s) an das WAN (typischerweise im Bereich von einigen MBit/s bis zu einigen Hundert MBit/s im Maximalfalle) an, sodass erhebliche Geschwindigkeitsunterschiede zwischen der Außenanbindung und der verwendeten lokalen Technik bestehen.

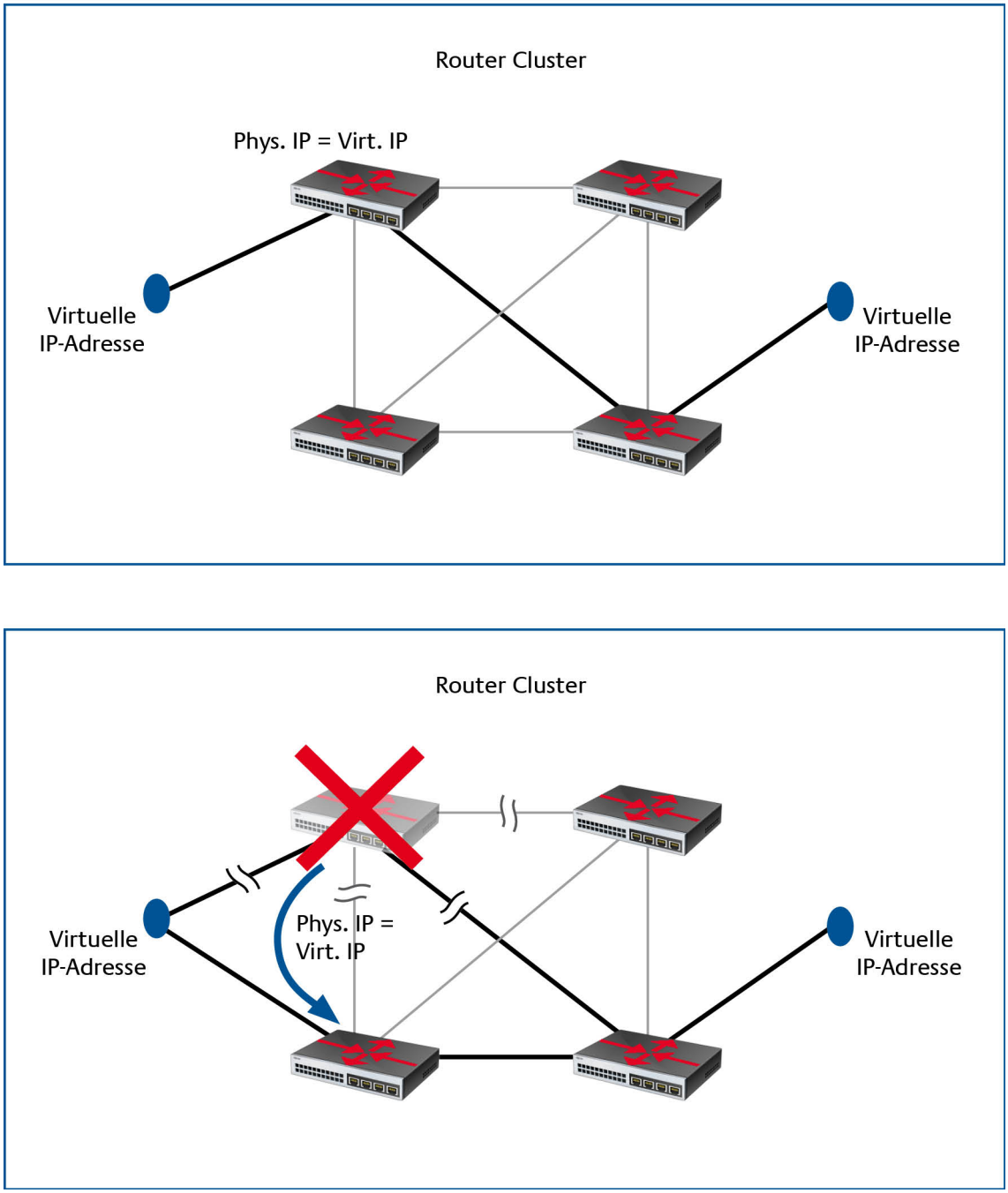


Abbildung 5: Fail-Over im Router Cluster

Ein im Effekt zu den äußeren (EGP/BGP) und inneren (VRRP/HSRP/CARP) Redundanzprotokollen ähnlicher, in der Umsetzung jedoch komplett anderen Ansatz ist Round Robin DNS. Wenn ein Zielsystem regelmäßig über seinen DNS-Namen statt direkt über seine Adresse von vielen verschiedenen Stellen angesprochen wird, so können im Domain Name Service unter einem Namen mehrere IP-Adressen abgelegt werden. Diese werden theoretisch alle in undeterministischer Weise bei Abfragen zurückgeliefert. Dieses Verfahren bietet sich effektiv eigentlich nur bei häufigen Web-Anfragen mittels HTTP (Hypertext Transfer Protocol) an und selbst dort ist mitunter durch den Einsatz von DNS-Caches, die schwer zu kontrollieren sind, nur

schlecht abzuschätzen, ob und inwiefern ein effektiver Fail-Over auf andere Systeme im Fehlerfall stattfindet. Vorteil des Verfahrens ist seine sehr einfache Umsetzung. Ein Nachteil ist, dass wirkliche Ausfälle eines der Server erst nach einer längeren Fail-Over-Zeit bei den Clients Wirkung zeigen. In diesem Fall muss eine Monitoring-Komponente das ausgefallene Gerät umgehend aus dem DNS-Server austragen. Daher wird dieses Verfahren nur bei Systemen mit sehr hohen Anfrageraten verwendet, bei denen die kurzzeitige Nichtverfügbarkeit einzelner HTTP-Anfragen vertretbar ist.

Für besonders wichtige Geschäftsprozesse sollten grundsätzlich mehrere alternative Übertragungswege in HV-Umgebungen vorgesehen werden. Dabei ist es oft angebracht, verschiedene Technologien zu verwenden. So kann eine Backup-Strecke unter Umständen durch eine Telefon- oder ISDN-Wählverbindung geschaffen werden, wenn als eigentliche Datenleitung ein DSL- oder Leased-Line-Verbindung bestand. Wird jedoch zu solchen Maßnahmen zur Verbesserung der allgemeinen Verfügbarkeit gegriffen, ist vorher sorgfältig zu prüfen, ob die Leitungsparameter der Lösung im Fehlerfalle für das Aufrechterhalten der Geschäftsprozesse ausreichen oder ob in diesem Falle das im Beitrag „Prinzipien der Verfügbarkeit“ gestellte Prinzip der Priorisierung der Datenkommunikation angewendet werden muss. Auch an dieser Stelle sei auf die Studie „Alternative Kommunikationswege für kritische Geschäftsprozesse“ verwiesen [ALKO08]. Eine andere, ergänzende Möglichkeit ist es, auf mehrere alternative Dienstanbieter zurückzugreifen. Hierbei ist darauf zu achten, dass die Zuleitung tatsächlich vollständig unabhängig ist, da auch unterschiedliche Dienstanbieter häufig auf bestimmten Ebenen miteinander kooperieren und gegenseitig Leistungen im Auftrag des anderen erbringen, sodass sich in diesem Fall faktisch doch keine echte Redundanz in der Zuleitung ergibt. Eine weitere Ausprägung dieser Redundanz ist die Wahl nicht-leitungsgebundener Übertragungswege (z. B. Übertragung via Satellit) für die alternative Strecke. Manche Hersteller bieten für diesen Fall inzwischen sogenannte Multi-Protokoll-Router an, die bei Ausfall einer physikalischen Verbindung in die Außenwelt automatisch auf eine andere physikalische Wegstrecke umschalten, also z. B. von einer Standleitung auf eine Satelliten-Übertragungsleitung.

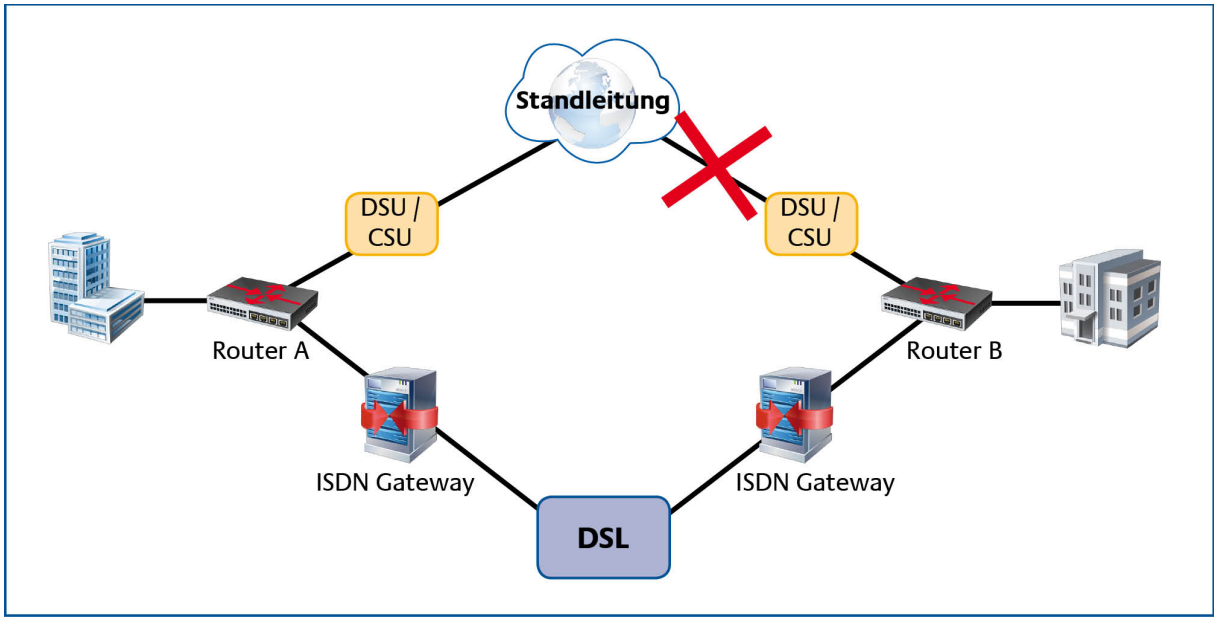


Abbildung 6: Alternative Weitverkehrsstrecken

1.4 Spezielle Netzkomponenten und Dienste

In diesem Abschnitt werden spezielle Netzkomponenten und Dienste betrachtet, deren Funktionalität über die Weiterleitung von Daten hinausgeht. Sie werden abschnittsweise anhand ihrer Funktionalität behandelt.

1.4.1 Sicherheitskomponenten des Netzes

Aus sicherheitstechnischer Sicht stellt eine Firewall den einzigen Übergang zwischen zwei Netzbereichen mit unterschiedlichem Sicherheitsniveau dar. In einem HV-Szenario ist diese Situation jedoch exakt die Definition eines Single Point of Failure. Aus diesem Grunde muss das Design der Firewall in einem HV-Szenario besonders sorgfältig konzipiert werden.

Wie in Abschnitt 1.3.3 ausgeführt wurde, lässt sich die Verfügbarkeit von Routing als Netzdienst der OSI-Schicht 3 gut durch Cluster erhöhen, da nur wenige Zustandsinformationen bei einem Fail-Over transferiert werden müssen. Ein Sonderfall von Netzkomponenten auf der Schicht 3 bilden hier Firewalls, die als statische oder dynamische Paketfilter oder nach dem Prinzip der Stateful Inspection realisiert sind, da sie funktional betrachtet ebenfalls als Router angesehen werden können. In diese Klasse gehören auch NAT-Gateways (von NAT, Network Address Translation), die private Adressbereiche auf einen festen Satz von öffentlichen Adressen und Ports abbilden.

All diesen Komponenten ist gemein, dass sie umfangreiche Statusinformationen sammeln, die sie für ihre Funktionsweise benötigen. Kann der Monitoring- und eigentliche Fail-Over-Mechanismus prinzipiell noch ähnlich wie bei VRRP, HSRP oder CARP ausgeführt werden, so besteht ein großer Unterschied in der Zustandsmigration. Hier werden unterschiedliche Ansätze verfolgt, die oftmals spezifisch für ein einzelnes Produkt oder zumindest für eine Produktgruppe sind. In den meisten Fällen werden dabei komplette Tabellen übertragen oder aktualisiert, die jeweils Einträge für alle virtuellen Verbindungen und die damit verknüpften Sicherheitspolicies (etwa blockieren, passieren

lassen oder protokollieren) enthalten. Zusätzlich werden noch Status- und statistische Informationen übermittelt.

Eine Stand-By-Komponente kann auch den kompletten Datenverkehr mitlesen und verarbeiten. Dies bedeutet, dass das System passiv mittels des Promiscous Mode alle Datenpakete intern verarbeitet und z. B. auch Zustandstabellen pflegt. In dieser Hinsicht ist eine solche Architektur dem active/active-Ansatz sehr ähnlich. Einzig das Weiterreichen von Datenpaketen an weitere Komponenten wird unterdrückt, bis ein Ausnahmezustand eintritt. Auf diese Weise bilden sich die gleichen Zustandsinformationen und im Fail-Over-Fall muss nur noch die Ausgabe der Daten aktiviert werden. Die vollständige Synchronität dieses Ansatzes ist durch die Komplexität des TCP/IP-Stacks auf dieser Ebene keine einfache Herausforderung. Solch aufwendige Lösungen ermöglichen die Sitzungstransparenz der bestehenden Datenverbindungen bei Ausfall einer Komponente.

Bei anderen Verfahren übermitteln die aktiven Komponenten jeweils zeitnah die relevanten Zustandsinformationen an die passiven Knoten. Es hängt von der Granularität dieser Übertragungen ab, in welchem Maße ein Fail-Over aus Benutzersicht völlig transparent abläuft.

Eine Lösung auf der Basis von Lastverteilung ist auch bei Firewalls möglich. In diesem Fall übernimmt die Lastverteilung die Aufteilung der einzelnen Sitzungen und muss, sofern es gefordert ist, beim Ausfall von Komponenten sicherstellen, dass die Sitzungen erhalten bleiben.

1.4.2 Verschlüsselungskomponenten

Auf sehr unterschiedlichen Ebenen zwischen Verbindungsschicht bis hin zur Transport- oder gar Anwendungsschicht werden Verschlüsslungsmechanismen zum Schutz der Vertraulichkeit eingesetzt. Die Endpunkte solcher Tunnel (etwa VPN-Gateways oder SSL-Server) sind neuralgische Komponenten. Ihnen muss besonderes Augenmerk geschenkt werden, da die Vorgänge zur Ver- und Entschlüsselung und der Verwaltung des Schlüsselmaterials komplex sind. Insbesondere die nicht-triviale Synchronisation von Schlüsselmaterial stellt besondere Herausforderungen an die Systeme. Für den Transport der verschlüsselten Datenpakete gelten hingegen die gleichen Bedingungen wie für alle anderen Pakete.

Da bei allen VPN-Lösungen Schlüssel eingesetzt werden, die mindestens in Teilen vertraulich und lokal begrenzt verbleiben müssen, ergibt sich ein gewisser Widerspruch beim Fail-Over, da spätestens dann Schlüsselmaterial auf einen neuen Knoten übertragen werden muss. Am einfachsten lässt sich dieses Problem bei statischen Schlüsseln (Preshared Keys, PSK) lösen, die aber wegen ihrer geringen Flexibilität nur selten eingesetzt werden. In diesem Fall muss der PSK optimalerweise schon a priori auf die Stand-By-Systeme aufgebracht werden, damit diese die Aufgabe bei einem Fail-Over unmittelbar übernehmen können.

Bei komplexeren Schlüsselaustausch-Protokollen wie dem beim im IPSec verwendeten IKE (Internet Key Exchange) müssen besondere Einheiten für einen kontinuierlichen Abgleich des Schlüsselmaterials sorgen. In diesem Fall ergeben sich für ein verbindendes Synchronisationssegment zwischen den einzelnen Knoten deutlich höhere Anforderungen, da im Gegensatz zu VRRP und Ähnlichen, wesentlich sensiblere Daten übermittelt werden. Es gibt verschiedene proprietäre Lösungen unterschiedlicher Anbieter auf dem Markt. Die Session-Transparenz verschlüsselter Verbindungen bei über SSL/TLS abgewickelten Protokollen (http, imap, ldap) kann durch Zusatzkomponenten (z. B. SSL-Proxy) sowie Re-Initialisierung erzielt werden.

1.4.3 Proxys

Proxys stellen eine weitere Zwischenstufe in der Betrachtung der Ende-zu-Ende-Datenkommunikation dar. Im weitesten Sinne sind Proxys aus Sicht des ISO/OSI-Schichtenmodells zumeist Server.

An dieser Stelle bleibt festzuhalten, dass zur Sicherung der Hochverfügbarkeit von Proxys häufig Loadbalancer eingesetzt werden, die von sich aus den Netzverkehr auf eine je nach Bedarfslage geeignete Instanz des Proxys weiterreichen. Der Vorteil dieses Ansatzes ist, dass in einigen Fällen keine expliziten Änderungen am Proxy selbst vorgenommen werden müssen. Die Transparenz-eigenschaften dieses Ansatzes hängen sehr stark von der Funktionsweise des Loadbalancers sowie auch vom verwendeten Protokoll ab, sodass dazu wenig allgemeine Aussagen getroffen werden können.

In eine ähnliche Kategorie fallen sogenannte Layer-4-Switches. Abgesehen von dem unglücklich gewählten Namen sind diese Systeme letztlich nicht viel mehr als eine Erweiterung eines Routers, der Anteile seiner Wegewahl auf Grundlage des Datenpayloads trifft, auf die er anderweitig nach strikter Auslegung des Schichtenmodells gar keinen Zugriff hätte. Insofern ist hier eine funktionale Verwandtschaft des Layer-4-Switches zu einem Loadbalancer, der Zustandsinformationen mitführt, festzustellen. Aus Sicht der reinen Ausfallsicherheit spielen diese Systeme daher eine untergeordnete Rolle. Als zusätzliche präventive Maßnahme zur transparenten Lastverteilung können sie sinnvoll sein, da sie eine Ende-zu-Ende-Betrachtung von Sitzungen gewährleisten können. Sie sind allerdings häufig nur für fest definierte Protokolle, häufig HTTP verfügbar.

1.4.4 DNS-Server

Das Protokoll des Domain Name Service (DNS) bietet bereits implizite Verfügbarkeitsfunktionalität. Diese ist für eine einfache Hochverfügbarkeit (VK-2) ausreichend. Für höchste Verfügbarkeit (VK-3) müssen zu den DNS-eigenen HV-Eigenschaften generische HV-Lösungen ergänzend hinzugenommen werden.

Eine DNS-Infrastruktur besteht immer aus zwei unterschiedlichen Teilen. Auf der einen Seite existiert ein Dienst (Server), der direkt oder indirekt autorisiert ist, für eine betreffende DNS-Zone die Namensauflösung durchzuführen. Auf der anderen Seite arbeitet ein sogenannter Resolver, der im Auftrag einer Applikation eine Namensauflösung durch Befragung des betreffenden Servers durchführt. Anfragen von Resolvern werden typischerweise über Proxy-Server geleitet.

Die genaue Funktionsweise von DNS kann dem Buch [AlLi01] entnommen werden.

1.4.4.1 DNS-Resolver

Der Resolver teilt sich auf in die Resolver-Bibliothek des jeweiligen Endgerätes sowie in die im Umfeld der jeweiligen Client-Gruppe platzierten Proxys. Jede Resolver-Bibliothek verfügt über die Möglichkeit mehrere DNS-Server anzugeben. Wenn ein Server nicht erreichbar ist, wird automatisch der alternative Server verwendet. Es handelt sich bei diesem Mechanismus um eine Client-basierte Arbitration (siehe Beitrag „Cluster-Architekturen“ im HV-Kompendium), da das Client-System über die Verteilung der Anfragen an die Proxy-Server entscheidet. Für den Proxy-Server sind für diese Funktionalität keine weiteren Maßnahmen erforderlich. Es kann im Fehlerfall aber zu durchaus bemerkbaren Verzögerungen im Bereich von Minutenbruchteilen kommen.

1.4.4.2 DNS-Proxy

Will man den Proxy-Server redundant auslegen, bieten sich hier „aktiv-passiv“ Lösungen (siehe Beitrag „Cluster-Architekturen“ im HV-Kompendium) auf Basis des VRRP (Virtual Router Redundancy Protocol, RFC 3768) an. Die Last auf einem DNS-Server ist typischerweise sehr gering, weswegen der Einsatz einer Lastverteilung in den seltensten Fällen notwendig ist. Sollte eine Lastverteilung trotzdem erforderlich sein, können nur Geräte eingesetzt werden, die in der Lage sind, UDP-Kommunikation zu verteilen. Dies stellt eine besondere Anforderung dar, weil UDP verbindungslos arbeitet und es auf Protokollebene keine Sessions (Sitzungen) gibt.

1.4.4.3 Versteckter Primärserver

Bei DNS-Servern wird häufig das Konzept des „Versteckten Primärservers“ (Hidden primary) verwendet. Dabei wird auf einem Primär-Server die Konfigurationsarbeit durchgeführt. Auf diesen greift aber kein DNS-Client zu. Von dem Primär-Server erfolgt ein Zonentransfer auf alle Sekundär-Server, die so alle auf dem gleichen Stand gehalten werden und die Client-Anfragen bearbeiten. Es handelt sich bei dieser HV-Architektur also um einen Baum-Cluster, bei dem jedoch auf den Wurzel-Knoten nicht zugegriffen wird. Der Client übernimmt die Funktion des Dispatchers. Alle dafür notwendigen Funktionalitäten gehören zum DNS und erfordern keine besonderen Extramaßnahmen.

Anhang: Verzeichnisse

Abkürzungsverzeichnis

Ein komplettes Verzeichnis hierzu findet sich in Band AH, Kapitel 5

Glossar

Ein komplettes Verzeichnis hierzu findet sich in Band AH, Kapitel 5

Literaturverzeichnis

Ein komplettes Verzeichnis hierzu findet sich in Band AH, Kapitel 5