

Einführung in Begriffe des Internets

Der Begriff Internet

Die Abkürzung Internet steht nicht, wie man vielleicht denkt, für "internationales Netzwerk". Tatsächlich wurde der Begriff Internet nämlich bereits benutzt, als das Internet noch eine rein nationale Angelegenheit der USA war. Um die Entstehung des Begriffes Internet zu ergründen, muss man die Situation in den frühen 70er Jahren betrachten, also jene Zeit, in der das Grundkonzept des Internet entstand.

Die typische Rechnerkonstellation, z. B. in einer Behörde oder einer Hochschule, bestand aus einem relativ leistungsfähigen Zentralrechner, an den zahlreiche Terminals angeschlossen waren, über die die einzelnen Mitarbeiter auf den Rechner zugreifen konnten. Bei dieser Vorgehensweise wurde rasch ein Zustand erreicht, in dem der Bedarf an Computerkapazität mit einer einzelnen Maschine, nicht mehr abgedeckt werden konnte. Es wurden somit mehrere Computer parallel betrieben. Es entstand hierdurch die Notwendigkeit eine Möglichkeit zu schaffen, dass einzelne Benutzer von ein und demselben Terminal aus auf mehrere dieser Rechner zugreifen konnten.

Die Konsequenz war die Erfindung lokaler Netzwerke (LAN = Local Area Network). Aus der Möglichkeit mit zahlreichen unterschiedlichen Computern kommunizieren zu können, entstand bei den Anwendern natürlich schnell der Wunsch, auch auf Computer anderer Einrichtungen zugreifen zu können. Die Technologie der LANs ist jedoch grundsätzlich beschränkt. Sowohl die Längen der Netzkabel als auch die Anzahl der angeschlossenen Einheiten dürfen aus technischen Gründen gewisse Maximalwerte nicht überschreiten. Es wurde also die Entwicklung einer weiteren Stufe der Vernetzung angestrebt, nämlich die Vernetzung der Netze.

Hierbei war von Anfang an klar, dass die Technologie dieser höheren Vernetzungsebene prinzipiell anders geartet sein musste als die der lokalen Netze. Dieses Ziel wurde durch die Entwicklung des Internet (also eines "Zwischennetzes" – inter lateinisch = zwischen) erstmals gelöst. **Das Internet ist also ein "Netz, das Netze vernetzt".**

Netzwerk Topologie

Die Realisierung eines Netzwerks zum Zwecke des Informationsaustausches kann auf vielfältige Weise geschehen. Grundsätzlich zu unterscheiden sind hierbei die leitungsorientierten und die paketorientierten Verfahren. Bei einem leitungsorientierten Netzwerk wird im Bedarfsfall ein Teilbereich des Netzwerks, zum Beispiel ein konkretes Kabel, den miteinander zu verbindenden Teilnehmern zugeordnet. Es wird also eine Leitung geschaltet, welche die Teilnehmer verbindet und die dann für andere Zwecke nicht mehr zur Verfügung steht, bis die Teilnehmer ihre Verbindung aktiv beenden (s. g. Peer-to-peer Netzwerk). Die Leitung ist im Sprachgebrauch des herkömmlichen Telefonnetzes besetzt. Das leitungsorientierte Netzwerkkonzept wird konkret in der Telefontechnik eingesetzt und ist somit jedermann vertraut. Computernetzwerke funktionieren jedoch meistens prinzipiell anders.

In einem Netzwerk existieren in der Regel zahlreiche Verbindungen scheinbar zur gleichen Zeit. Eine Zuordnung zwischen einem konkreten Kabel und einer konkreten Verbindung gibt es nicht. Es muss somit jedem Teilnehmer möglich sein, aus der Vielzahl der Informationen, die über das Netzkabel übertragen werden, die herauszugreifen, die auch für ihn bestimmt sind. Die Lösung dieses Problems wird durch die sogenannte **Paketorientierung** ermöglicht. Die Daten

werden hierbei nicht, wie im leitungsorientierten Fall, unformatiert gesendet, sondern unter Beachtung besonderer Vorschriften für die Versendung vorbereitet.

Diese Vorschriften werden als Übertragungsprotokolle bezeichnet. Die klassische Brief- oder Paketpost ist ein alltägliches Analogon für diese Vorgehensweise. Ein Kurier oder Spediteur, der täglich zwischen einer Firma und einem Kunden hin- und herfährt, wäre die Entsprechung des leitungsorientierten Dienstes. Das Fahrzeug wäre unmittelbar den beiden Adressaten zugeordnet und es wäre möglich, dass die zu transportierenden Waren mehr oder weniger ungeordnet einfach in den Laster hineingeworfen werden.

Ganz anders verhält es sich bei Benutzung der paketerorientierten Post. Hier hat der Benutzer keinen Einfluss darauf, in welchen und in wie vielen Fahrzeugen seine Briefe und Pakete tatsächlich transportiert werden. Um eine korrekte Zustellung unter diesen Umständen zu ermöglichen, muss ein Brief oder Paket gewisse Formvorschriften einhalten. Es muss insbesondere mit einer Empfängeradresse und möglichst auch mit einem Absender versehen sein. Ebenso verhält es sich bei den Datenpaketen in Computernetzwerken.

Funktionieren können derartige netztypübergreifende Systeme jedoch grundsätzlich nur, wenn insgesamt ein einheitliches Protokoll für den Informationsaustausch benutzt wird. Dieses Protokoll zu definieren, war die zentrale Aufgabe bei der Erarbeitung der Internet-Technologie.

Das ursprüngliche Forschungsvorhaben, das letztendlich die Grundlagen für das heutige Internet geschaffen hat, war militärisch motiviert. Im Jahre 1973 wurde dieses Vorhaben vom US-Verteidigungsministerium ins Leben gerufen, um das militärische Forschungsnetzwerk ARPAnet mit diversen anderen Netztechnologien zu verbinden.

Berücksichtigt werden musste bei der Entwicklung auch die Robustheit des gesamten Netzwerks gegenüber militärischen Schlägen. Das Routing beschreibt die eindeutige Adressierung eines Datenpakets (über welche Stationen das Paket geht) an den Empfänger. Dabei wird der für das Paket optimale Weg durch die einzelnen Gateways behandelt. Insbesondere stand damals die Frage im Raum, ob ein überregionales Computernetz auch nach einem nuklearen Angriff noch funktionieren könnte.

Klassische Netzwerke mit sternartiger Anordnung sind offensichtlich sehr empfindlich. Im Falle eines Ausfalls der zentralen Vermittlungsstelle wird jede Kommunikation unmöglich.

Eine höhere Zuverlässigkeit lässt sich dadurch erreichen, dass das Netzwerk Maschen aufweist. Ein derartiges Netzwerk wird als redundant bezeichnet. In einem redundanten Netzwerk gibt es jeweils mehrere Wege, auf denen ein Datenpaket seinen Empfänger erreichen kann. Fällt einer dieser Wege aus, so kann sofort auf einen anderen Weg umgeschaltet werden. Hierdurch wird allerdings den Routern, die ja die „Route“ eines Datenpaketes festlegen, zusätzliche Leistungsfähigkeit abverlangt. Erkennt beispielsweise ein Router, dass eines der mit ihm verbundenen Netzsegmente nicht funktionstüchtig ist, so muss er die Datenpakete umleiten können. Dies trifft auch für den Fall zu, daß ein Netzsegment erheblich überlastet ist.

Die Umleitung eines Paketes besteht dabei in der Weiterleitung an einen anderen Router. Insgesamt kann dieses komplexe Verhalten des gesamten Netzwerks zur Folge haben, dass mehrere Datenpakete, die zu ein und derselben Informationseinheit gehören, auf unterschiedlichen Wegen schließlich bei demselben Adressaten ankommen. Hierbei geht allerdings die ursprüngliche Reihenfolge der Datenpakete verloren. Um die übertragenen Informationen dennoch rekonstruieren zu können, sind entsprechend aufwendigere Protokolle erforderlich. Die Lösung der formulierten Forschungsaufgabe lag somit in der Realisierung sehr leistungsfähiger und robuster Datenübertragungs-Protokolle. Die wichtigsten dieser Protokolle

heißen IP (Internet Protocol) und TCP (Transmission Control Protocol). Sie werden meistens in einem Atemzug genannt. Der Name der Lösung der Aufgabe wird also mit TCP/IP angegeben.

Die Aufgaben von TCP und IP

IP definiert die grundsätzliche Gestalt eines Internet-konformen Datenpakets. Ein sogenanntes IP-Paket ist mit einer Absender- und einer Empfängeradresse versehen. Beide Adressen müssen in Form von sogenannten IP-Nummern angegeben sein. Die Maximalgröße eines IP-Pakets beträgt 1500 Zeichen. Diese Begrenzung wurde mit Rücksicht auf die Ethernet-Spezifikationen vorgesehen, da das Internet tatsächlich im wesentlichen zahlreiche lokale Netze auf Ethernet-Basis verbindet.

Größere Informationseinheiten, die die Begrenzung auf 1500 Zeichen überschreiten, können also nur transportiert werden, wenn Sie zuvor in kleinere Einheiten aufgespalten werden. Um ein späteres Zusammenfügen der Teile zum ursprünglichen Ganzen zu ermöglichen, müssen den Datenpaketen weitere Informationen beigelegt werden. Diese Informationen können beispielsweise aus einem logischen Namen, der allen zu einer Informationseinheit gehörenden Datenpaketen gemeinsam ist, sowie einer fortlaufenden Nummer bestehen. Das Protokoll TCP übernimmt diese Aufgabe. Durch die Numerierung der einzelnen TCP/IP-Pakete können Paketverluste sofort erkannt werden.

Die Neuansforderung eines verlorenen oder zerstörten Paketes erfolgt unter TCP automatisch. Die Leistungsfähigkeit des Protokolls TCP ist wesentliche Grundlage für die Zuverlässigkeit der meisten Internetanwendungen.

Alle klassischen Internetdienste wie http (World Wide Web), FTP (File Transfer Protokoll) oder SMTP (Mailversand) bedienen sich bei der Übertragung des TCP/IP-Protokolls

Adressierung im Internet

Innerhalb des Internet muss jeder einzelne Computer über eine eindeutige Adresse ansprechbar sein. Zu diesem Zweck wird dem Rechner, wie schon erwähnt, eine IP-Nummer zugewiesen. Eine IP-Nummer ist eine aus vier Byte bestehende Informationseinheit. Ausgeschrieben werden IP-Nummern als Folgen von vier Zahlen im Bereich zwischen 0 und 255, die jeweils durch Punkte voneinander getrennt werden. Die Systematik der IP-Nummern ist hierarchisch. Beispiel: 194.68.127.18.

Die IP-Nummern mehrerer Rechner, die an ein und demselben Netzwerkstrang angeschlossen sind, werden sich beispielsweise nur im Wert der letzten drei Ziffern unterscheiden.

Diese hierarchische Gliederung ist erforderlich, um das oben behandelte Routing zu ermöglichen. Es wäre nämlich völlig undurchführbar jeden Router im Internet über jeden einzelnen Rechner und dessen geographische Position zu informieren. Tatsächlich gliedert sich die IP-Nummer daher in zwei Teile, nämlich die Adresse des lokalen Netzwerks, zu dem der adressierte Computer gehört, und einen Adreßteil, der diesen Computer innerhalb des lokalen Netzwerks eindeutig bestimmt.

Welche der vier Zahlen zu welchem der beiden Adreßbestandteile gehören, ist von der Größe des jeweiligen lokalen Netzes abhängig. Bei den größten Netzwerken, den sogenannten Klasse-A-Netzwerken, wird nur die erste Zahl als Netzwerk-Adresse benutzt. Die Klasse-B-Netze mittlerer Größe benutzen die ersten beiden Zahlen als Netzwerk-Adresse und die kleinsten Klasse-C-Netze benutzen schließlich die ersten drei Zahlen als Netzwerk-Adresse. Bei

Klasse-C-Netzen steht somit nur die letzte Zahl zur Identifikation des einzelnen Rechners zur Verfügung. Klasse-C-Netze können also höchstens 256 einzelne Computer umfassen.

Das Domain-Name-System

Die Tatsache, dass die internen Internet Adressen aus Zahlengruppen bestehen, ist natürlich für den Anwender eher unbequem. Zusätzlich zu den IP-Adressen werden den Computern daher Namen gegeben. Auch bei der Identifizierung eines einzelnen Rechners über einen derartigen Namen muss natürlich Eindeutigkeit gewährleistet sein. Um dies zu erreichen, wurde ebenfalls ein hierarchisches System erdacht, das sich jedoch nicht so streng an die physikalischen Gegebenheiten der Netztopologie anlehnt wie das System der IP-Nummern.

Es ist vielmehr so, dass die Gesamtheit der per Internet verbundenen Rechner nach logischen Kriterien in Bereiche, die sogenannten Domains, eingeteilt werden. Das Benennungssystem heißt daher Domain-Name-System oder kurz DNS. Ähnlich wie bei den IP-Adressen sind auch bei den Domain-Namen die einzelnen Namensbestandteile durch Punkte voneinander zu trennen. Die Anordnung ist jedoch bzgl. der hierarchischen Bedeutung genau andersherum.

Der erste Namensbestandteil identifiziert einen einzelnen Rechner, der nächstfolgende einen Teilbereich zu dem der entsprechende Rechner zu zählen ist. Dies geht so weiter bis zum letzten Namensbestandteil, der eine allergrößte Bereichszuordnung darstellt. Derzeit ist es so, dass dieser letzte Bestandteil des Domain-Namens einen Hinweis auf die Nation, in welcher sich der Rechnerstandort befindet, darstellt, sofern der Rechner nicht in den USA beheimatet ist. Innerhalb der USA werden sechs Bereiche unterschieden.

Dies sind der Bereich *com* für kommerzielle Unternehmen, *edu* für Hochschulen und sonstige Lehrinrichtungen, *gov* für nichtmilitärische Einrichtungen des Staates, *mil* für militärische Einrichtungen, *org* für nicht profitorientierte Organisationen privater Art und schließlich *net* für spezielle Netzwerk-Ressourcen.

Aufbau eines Rechner-Namens

Eine typischer Rechnername innerhalb Deutschland sieht z. B. etwa so aus:

benhur.technikerschule-noerdlingen.de.

Diese Adresse gehört zum Computer, der den Namen "benhur" trägt und sich in der Technikerschule Nördlingen befindet. Der letzte Namensbestandteil *de* steht hierbei, wie gesagt, für die Nation, also für Deutschland.

Anwendungsprogramme werden vom Benutzer meist nur den Domain-Namen eines Computers, zu dem eine Verbindung aufgebaut werden soll, erfragen. Damit nun eine Verbindung auch wirklich hergestellt werden kann, muss dieser Name natürlich in die IP-Nummer, die denselben Computer identifiziert, übersetzt werden. Zu diesem Zweck werden spezielle Rechner, die sogenannten Name-Server, im Internet betrieben.