Zeitschrift: Reihe Kriminologie / Schweizerische Arbeitsgruppe für Kriminologie = Collection criminologie / Groupe suisse de travail de criminologie

Herausgeber: Schweizerische Arbeitsgruppe für Kriminologie

Band: 19 (2001)

Artikel: (Un-)Möglichkeiten der Inhaltskontrolle mit technischen Mitteln im Internet

Autor: Semken, Hartmut

DOI: https://doi.org/10.5169/seals-1051172

HARTMUT SEMKEN

(UN-)MÖGLICHKEITEN DER INHALTSKONTROLLE MIT TECHNISCHEN MITTELN IM INTERNET[1]

Sie benutzen offenbar das Internet. Schämen Sie sich dessen eigentlich? Wo wir doch aus den Medien wissen, dass das Internet eigentlich nur aus Pornos, Anleitungen zum Bombenbau, Volksverhetzung und so weiter besteht. Oder ist da noch was anderes? Gibt es vielleicht auch ein «sauberes» Internet?

1 Wahrnehmungsgrenzen

Offenbar haben wir hier eine Wahrnehmungsgrenze erfahren. Nicht nur der übliche Medien-Effekt, nur das Extreme, das besonders Böse darzustellen, trifft auf die Wahrnehmung des Internet zu. Auch ist die Erfahrung jedes einzelnen Internet-Teilnehmers ein kleiner Auszug aus dem ganzen Spektrum dessen, was Internet ausmacht. Wir wollen im Folgenden versuchen, uns diesem Phänomen – nur dieser Begriff scheint mir dem Internet gerecht zu werden – einmal ein wenig zu nähern. Dabei soll der Schwerpunkt auf jenen Aspekten liegen, die zu einer strafrechtlichen Dimension von Internet beitragen. Aber nicht nur um Straftaten solle es gehen sondern generell um das Verständnis davon, was Internet ist und wie es dazu kommt.

---

1 Dieser Text ist hervorgegangen aus einem Vortrag des Autors anlässlich der Jahrestagung 2001 der Schweizerischen Arbeitsgruppe für Kriminologie. Er steht nun auch unter www.hase.net bereit und wird in loser Folge erweitert und durch neue Gedanken ergänzt. Anregungen, Kritik und Fragen sind immer gern gesehen; bitte schreiben Sie an webmaster@hase.net.

2 Begriff Homepage

Betrachen wir zunächst einmal den Begriff der Homepage. Fragt man Internet-Nutzer, ob sie eine Homepage haben, dann hört man oft ein «nein». Was an sich verwundert. Die Homepage ist die Seite, die der Browser anzeigt, wenn man das «home», das «Häuschen» Icon anklickt. In der Regel wird diese auch beim Programmstart dargestellt. Einer der ersten Web-Browser, NCSA Mosaic, führt diesem Begriff von Homepage ein, er verwendet das Wort in dieser Bedeutung. Die Wortbedeutung hat sich jedoch aus für den Autor nicht ersichtlichen Gründen zu etwas anderem entwickelt.

Ein junger Mann – Student wohl zu der Zeit – stellte fest, dass diesen Internet-Dings ja ganz brauchbar ist. Man muss sich halt nur immer merken, welche Informationen man wo findet. Als schreibkundiger macht man sich da eine Liste und – wenn man sowieso einen Computer hat – gleich elektronisch. Legt man eine solche Liste von Sites, die man öfter besucht, als HTML-Dokument an, dann kann man nicht nur den Link (also einen Verweis auf die Site) sondern auch einen Kommentar dazu ablegen. Éine solche Liste wird schnell unübersichtlich, also beginnt man sie nach Kategorien zu ordnen. Da die Liste ohnehin HTML ist kann man sie auch leicht in einen Webserver stellen und so auch andern zur Verfügung stellen. Der Rest der Firmengeschichte von Yahoo soll hier der Kürze halber weggelassen sein. Diese erste Liste mit Links auf interessante Sites und zugehörigen Kommentaren war die Homepage (im Sinne von Mosaic) der Yahoo-Gründer: Diese Seite sollte beim Programmstart und beim Home-Icon erscheinen. Sie wurde auch schnell zur Homepage (im Sinne Mosaic) anderer Nutzer; aber irgendwie wurde diese neu entstandene Site als die «Homepage der Firma Yahoo», also die Website, die von Yahoo mit Informationen beschickt wird, verstanden. Dies dürfte dazu beigetragen haben, dass sich der Begriff hinter dem Wort «Homepage» von der Bedeutung wie bei Mosaic in ein Synonym zu «meine Website» gewandelt hat.

3 Begriff «Internet»

Der Begriff hinter dem Wort Internet ist einer Wandlung unterworfen, diese ist jedoch nicht so einfach darzustellen. Also will ich einfach mal versuchen darzustellen, wie sich für einen Techniker, der für einen ISP arbeitet, das Internet darstellt.

3.1 Transportnetz

Das Internet ist zunächst einmal ein Transportnetz, das Daten transportieren kann. Dabei kommen die Abläufe (die Beschreibung eines Ablaufes einer Kommunikation ist in der Informatik ein «Protokoll») gemäss dem «Internet Protocol» (IP) zum Einsatz.

Protokolle treten in der Informatik in der Regel als Familie auf: Mehrere verschiedene Protokolle regeln Abläufe, die alle Teilaspekte desselben Kommunikationsvorgangs darstellen. So bleiben die einzelnen Protokolle einfach und damit handhabbar für einen Implementierer. Das Internet ist also erst einmal ein Transportnetz für IP-Pakete, ganz so wie das ISDN ein Transportnetz für zeitsynchrone Datenströme ist.

Die Familie TCP/IP hat gegenüber älteren Protokollen gewisse Vor- und Nachteile. Da die Vorteile für viele Anwendungen überwiegen hat IP viele andere Techniken schon verdrängt. Dort wo Defizite sichtbar sind, entwickeln sich nach und nach Workarounds oder Erweiterungen des Protokolls, die diese ausgleichen. Dadurch nagt IP auch an den Domänen anderer Techniken immer stärker. Die Eigenschaften von IP-Netzen wollen wir nun kurz beleuchten.

3.1.1 Universalität

Die Universalität, die Möglichkeit, praktisch alle Daten transportieren zu können, ist eine wichtige Eigenschaft von IP-Netzen: IP ist neutral hinsichtlich der Daten, die es transportiert. Im Prinzip kann

man alles, was sich digital kodieren lässt, über IP-Netze transportieren; dabei kommen in der Regel weitere Protokolle der Familie zum Einsatz.

Das wohl am meisten verwendete andere Protokoll ist TCP, daher wird die ganze Familie auch als TCP/IP bezeichnet; ohne die Mitglieder UDP und ICMP wäre sie jedoch nicht komplett noch wirklich einsatzfähig.

Die Protokollfamilie stellt im wesentlichen zwei Transportmechanismen zur Verfügung, TCP und UDP. UDP (User Datagram Protocol) funktioniert im wesentlichen wie der Briefdienst der Post: Der Absender packt seine Daten in ein Paket (einen Briefumschlag), adressiert dieses Paket, bei IP wird dazu die Quelladresse wie auch die Zieladresse angegeben, beide sind einfach Zahlen, und übergibt es dem Postdienst (IP protocol layer). Das Datenpaket (der Briefumschlag) wird dann anhand der Zieladresse transportiert – aber eine Quittung über den korrekten Empfang erfolgt nicht. Es ist leicht einzusehen, dass ein Empfängerprogramm eine Quittung selbst versenden kann und dass der Programmierer so eine gesicherte Übertragung schaffen kann; daher kommt der Name User Datagram Protocol: Der Benutzer (Programmierer) kann darüber Datagrams versenden und selbst alle erforderlichen weiteren Massnahmen ergreifen.

Oft ist es aber erforderlich, dass der korrekte Empfang der gesendeten Daten sichergestellt wird, dass eine gesicherte Kommunikation erfolgt. TCP stellt diesen Dienst zur Verfügung. Ein TCP-Kanal wird einmal geöffnet und kann dann zwischen seinen beiden Endpunkten (Programme, die auf Computern laufen) Daten transportieren. Die TCP-Software sorgt dabei dafür, dass die Daten, die abgesendet wurden, komplett und in der gesendeten Reihenfolge empfangen werden – der Programmierer der Anwendung muss sich um die Korrektur von Übertragungsfehlern und verloren gegangene Pakete nicht kümmern. TCP stellt einen fehlerkorrigierten Stream von Ende zu Ende sicher.

Die Programme, die über TCP oder UDP miteinander kommunizieren, müssen nicht immer auf verschiedenen Computern laufen. So kann ein Datenbank-Client durchaus auf einen Datenbank-Server zugreifen, der auf demselben Computer läuft. Für den Programmierer stellen TCP und UDP ein universelles Mittel zur Kommunikation dar – innerhalb eines Computers oder über ein Datennetz.

3.1.2 Technische Anspruchslosigkeit

Eine andere wichtige Eigenschaft aus Sicht des Technikers ist die Anspruchslosigkeit gegenüber dem «darunter liegenden» Netz: Praktisch jede Technik, die Bits übertragen kann, kann für den Transport von IP verwendet werden. Das mag zunächst selbstverständlich erscheinen, aber in der Tat ist es ein Durchbruch in der Technik, der grösser kaum sein könnte. Zum ersten Mal war mit IP ein Protokoll verfügbar, das universell war, das sowohl einsetzbar war über eine Modemstrecke mit 9600 Bit/s als auch auf einem Ethernet mit 10 Megabit/s – wie auch auf zukünftigen Netzwerken (FastEthernet ist eine Entwicklung, die aus der Sicht von IP zukünftig ist). Diese Eigenschaft trägt zum grossen Erfolg des Internet bei, da sie die Anwendung auf vorhandenen wie zukünftig zu installierenden Infrastrukturen erlaubt. Für Unternehmensnetzwerke ist diese Eigenschaft besonders wichtig, da ohne einen Bruch im System, ohne eine Veränderung der Struktur jede neue Technik leicht eingegliedert werden kann. Diese Eigenschaft und die herstellerübergreifende Standardisierung haben IP zum Protokoll der Wahl werden lassen und haben IPX (Novell) Vines (Banyan), AppleTalk (Apple) und LAT (Digital Equipment) verdrängt. Nur SNA hält sich noch recht hartnäckig, da die Umstellung der IBM-Mainframes nicht immer einfach ist.

3.1.3 Überbuchung

Ein weiterer wichtiger Erfolgsfaktor ist die Fähigkeit zur Überbuchung von Leitungen, die IP bietet. Diese ist der Schlüssel zu niedrigeren Kommunikationspreisen als bei anderen Techniken.

Die Überbuchung beschreibt die Eigenschaft, dass sich mehrere Benutzer eine Leitung teilen. Dies ist auch bei anderen Netzen völlig üblich. Beim alten Telefonnetz (in Deutschland System 55V der Post) teilen sich z.B. 100 Teilnehmer eine «Wählergruppe» aus 7 Wählern. Es können damit maximal 7 Teilnehmer gleichzeitig eine Verbindung zu einem anderen Teilnehmer aufbauen. Der achte Teilnehmer würde beim Abheben des Hörers gleich ein Besetzzeichen erhalten. Die Wählergruppe ist also um den Faktor 14,4 «überbucht». Anders als bei solchen Netzes wirkt bei IP die Überbuchung jedoch nicht «serviceverhindernd» sondern «servicevermindernd». Hat ein Netzknoten, an den 100 Kunden mit je 2 MBit/s angeschlossen sind, eine Kapazität von 25 MBit/s zu anderen Netzknoten, dann ist das Backbone um den Faktor 8 überbucht. Sollten also alle Teilnehmer gleichzeitig Daten zu übertragen versuchen, dann erhält jeder nur 256 kBit/s an Durchsatz. Diese Form der Serviceeinschränkung ist für die Benutzer zum einen weniger sichtbar, denn viele Faktoren (von der Quelle bis zum Ziel) tragen zur tatsächlichen Übertragungsrate bei, zum anderen ist sie in der Regel viel erträglicher als ein vollständiger Ausfall der Leistung.

3.1.4 Garantien

Andere Eigenschaften enthalten das Fehlen jeglicher Servicegarantien – IP garantiert keinen Datendurchsatz, keine Maximalzeit für die Zustellung und auch keinen fehlerfreien Transport – also Eigenschaften, die man von der Post kennt. Meist geht alles ganz gut, aber eine Postkarte kann schon mal abhanden kommen oder statt der üblichen drei Tage auch drei Wochen unterwegs sein.

IP arbeitet ohne jegliche Garantien für die Übertragungszeit. Dies ist ein Resultat dessen, was oben als technische Anspruchslosigkeit dargestellt ist. Wenn das Protokoll auch auf Netzen funktionieren soll, die kein definiertes Zeitverhalten zeigen können (z.B. Modemverbindungen mit Kompression und Fehlerkorrektur oder Ethernet-Busnetze) dann kann es kein definiertes Zeitverhalten liefern.

Dieser Nachteil von IP gegenüber Protokollen wie ATM wiegt die Vorteile aber nicht auf. Zum einen, weil IP Netze schlicht so viel preisgünstiger als andere universelle Netze sind, zum anderen weil viele der «harten» Echtzeitanforderungen gar nicht so hart sind:

- Für die Darstellung von Audio- oder -Videostreams reicht es aus, wenn das Netz im Mittel über ca. 10 Sekunden die Datenrate des Streams liefern kann. Man überträgt zunächst 10 Sekunden Material in einen Puffer beim Empfänger. Wenn das Netz dann im Mittel die Daten in den Puffer nachliefern kann, reisst der Datenstrom beim darstellenden Programm niemals ab, der Film kann ohne Ruckeln oder Lücken dargestellt werden.
- Für Echtzeitanwendungen wie Telephonie reicht die Übertragungsqualität des Netzes in Strenge nicht aus. In den meisten Fällen reicht sie jedoch aus und die Qualitätseinbussen ertragen viele Anwender im Gegenzug für niedrige Preise.

IP und die auf IP basierenden Netze werden ständig weiterentwickelt, insbesondere an den Echtzeiteigenschaften wird ständig gefeilt. Der bereits verabschiedete Standard für das IP der Zukunft (IPv6) enthält diverse Mechanismen, mit denen dem Kunden eine Mindestleistung garantiert werden kann; die Einführung dieser neuen Anwendungen wird jedoch noch eine Weile dauern. Dennoch bietet IP mit solchen Erweiterungen wie sie teilweise schon durch die Verwendung von ATM als Medium unter IP realisiert werden das Potential dazu, im Spannungsfeld aus Qualität und Preis der Leistung für jeden Kunden das gewünschte zu bieten – kein «one size fits all».

3.2 Anwendungen

Wie im Abschnitt Transportnetz dargestellt beschreibt «Internet» zunächst einmal eine Technik, ein Verfahren für die Datenkommunikation. «Internet» beschreibt also nicht die Anwendungen wie WWW, Real Video Streaming, Dateikopie (ftp) oder store-and-forward

Transport kleiner Textnachrichten (smtp) oder Speicherung kleiner Textnachrichten zur späteren Abholung (POP3 eMail).

Zwar ist der Begriff dabei sich zu wandeln, ähnlich wie unter Homepage dargestellt. Das Wort Internet scheint sich mehr an den Begriff zu nähern, der die Gesamtheit aller über das kommerzielle IP-Netzwerk der Arin-, RIPE- und APNIC-Adressräume verfügbaren Dienste darstellt. Wenn man «Internet» jedoch in dieser Bedeutung verwendet, dann stellt dies zwei Fehler dar:

- ein Wort für «die technische Infrastruktur des globalen IP-Netzes» steht nicht mehr zur Verfügung
- alle Anwendungen werden über einen Kamm geschoren.

Gerade letzteres ist offenbar nicht gut: Persönliche eMail und eine Publikation auf einer hoch frequentierten Website sind offenbar verschieden – und müssen ggf. verschieden bewertet werden. Worin aber liegt der Unterschied zwischen den verschiedenen Anwendungen?

4 Medium?

Das Fragezeichen in der Überschrift stellt es schon dar: «Internet» ist kein Medium. Das wird deutlich, wenn man die Anwendung telnet (remote login für eine Mehrbenutzermaschine) ansieht. Diese stellt offenbar keine mediale Anwendung dar, denn es fehlt an der Kommunikation schlechthin.

Internet ist kein Medium genau wie Papier kein Medium ist. Erst die Kombination mit der Drucktechnik und die Anwendung als Zeitung oder Zeitschrift lässt ein Medium entstehen. Bei einer Zeitschrift kommt jedoch niemand auf die Idee, den Spediteur, den Drucker oder den Setzer (alle können externe Dienstleister für den Verlag sein, sind es oft oder in der Regel) für den Inhalt verantwortlich zu machen.

Beim Thema Internet aber versagen unsere bewährten Vorstellungen davon, was erlaubt und was verboten sein sollte ungewöhnlich oft und mit weit reichenden Folgen.

Das liegt zum einen daran, dass eben undifferenziert alles mögliche unter «Internet» subsumiert wird und dann mediale und nichtmediale Anwendungen über denselben Kamm geschoren werden. Zum anderen liegt es daran, dass sich das Internet so rasch weiterentwickelt, dass es durchaus nicht trivial ist, den Überblick zu behalten.

Ein weiterer Grund dürfte sein, dass der Einfluss des Internet auf die Entwicklung der Gesellschaft sehr gross werden kann. Schon jetzt deutet sich an, dass der Impact des Internet die Auswirkungen von Buchdruck mit beweglichen Lettern + Bibelübersetzung übertreffen könnte. Mit nur wenig Phantasie kann man durchaus sehen, dass auch das Fernsehen in den Schatten gestellt werden könnte. Solche Veränderungen machen Angst, werden als bedrohlich empfunden; dies wird – wie üblich – dadurch verstärkt, dass die Berichterstattung zum Thema immer die Extreme darstellt und so einen Schatten auf das Ganze wirft.

5 WWW

Die wohl prominenteste Anwendung des Internet ist heute das WWW. Für den Autor stellt sich diese Anwendung klar getrennt von E-Mail und filetransfer dar. Auch verschiedene Anwendungen, für die eine Webseite einen Katalog darstellt, sind nicht Bestandteil des WWW – so wie ein Auto kein Druckerzeugnis ist, nur weil sein Prospekt auf Papier gedruckt wurde. Das WWW ist eine massenkommunikative Anwendung fast klassischer Art: Die Kommunikation richtet sich von der (definierten) Quelle an eine zahlenmässig und von der Personengruppe her undefinierte Leserschaft. Man kann das WWW daher als Medium auffassen.

5.1 WWW ist anders

Anders als in klassischen Medien liegt jedoch die Eintrittsschwelle zur Teilnahme weit niedriger: Eine Website kostet heute weniger als der Betrieb und die Abschreibungen für den Computer, den man zum Aufbereiten und Einstellen des Inhaltes benötigt. Anders als bei Print- oder elektronischen Medien ist damit eine Publikation für jedermann zum ersten Mal möglich. Ein anderer Unterschied zu klassischen Medien ist die scheinbar geringe Persistenz. Der Inhalt bei einem Buch hat eine Lebensdauer von einigen Jahren bis einigen Jahrzehnten. Die Persistenz des Buches ist sogar noch eine bis zwei Grössenordnungen höher. Bitte betrachten Sie hier nicht die klaren Ausnahmefälle (der Inhalte der Bibel hat massive Bedeutung seit mehreren Menschenaltern, Klassiker wie Goethe oder Shakespeare erreichen ebenfalls eine aussergewöhnliche Lebensdauer) sondern die Regelfälle, das «Durchschnittsbuch». Für Fachbücher wie Romane sind die Zeiten der Bedeutsamkeit und der materiellen Existenz ähnlich. Andere Publikationen erreichen weit kürzere Bedeutungslebensdauern: Die Zeitung von gestern taugt nur noch zum Einwickeln der Handelsware auf dem Fischmarkt. Dennoch ist auch bei Zeitungen eine hohe Persistenz erreicht, denn professionelle Archive stellen auch heute noch Zeitungen aus der Weimarer Republik bereit und ermöglichen so z.B. Gero Ganderts Werk «Das Handbuch des deutschen Films in der Weimarer Republik».

Im Web sind Bedeutungslebensdauer und Persistenz (also Verfügbarkeitslebensdauer) der Inhalte sehr verschieden. Manche Site verschwindet schon nach einigen Tagen wieder auf Nimmerwiedersehen, andere werden scheinbar ewig archiviert. Das Web bietet sich auch in seiner Natur an für alle Arten der Publikation: Tages-, ja minutenaktuelle Nachrichten (Heise newsticker) sind genauso verfügbar wie Sammlungen von Datenblättern für schon vergessen geglaubte, steinalte (>4 Jahre) Computerhardware. Das Web ist also anders, aber was macht es in der Tat so neu? Das neue ist der Hyperlink.

5.2 Hyperlinks

Das Element, das http-Anwendungen von ftp-Anwendungen so stark unterscheidet, ist wohl der Hyperlink. Rasend neu ist die Idee nicht, katalysiert durch die kostengünstige und schnelle Kommunikationsplattform Internet ist die Wirkung aber bestechend: Das WWW stellt sich als eine Anwendung dar, bei der durch die Verzweigungen von einer Site auf eine andere so etwas wie ein «Informationsmeer» entsteht; die Wortentlehnung «surfen», also die Beschreibung des Stöberns in den Informationen als ein Wellenreiten, zeigt das ja schon sehr schön. Hyperlinks kann man zunächst einmal in zwei Klassen unterteilen, die sich grundlegend voneinander unterscheiden.

5.2.1 Manuell verfolgte Links

Einen Hyperlink dieser Form muss man manuell verfolgen. Der Computer nimmt dabei viel Arbeit ab, dennoch ist eine Aktion des Benutzers erforderlich, um die Verfolgung einzuleiten: Vor der Benutzeraktion (Mausklick bzw. Tastendruck) wird das Ziel des Link nicht abgerufen. Manche Maschinen verfolgen den Link schon auf Verdacht – so z.B. gewisse Proxies – um dann, wenn der Benutzer wirklich auf den Link klickt, schneller reagieren zu können. Auch ist es leicht ein Programm zu schreiben, das mit einem url versehen alle Links automatisch verfolgt und alle so erreichbaren Daten beschafft (z.B. wget oder WinHTTrack, http://www.httrack.com).

Insofern ist die Einstufung als «manuell verfolgter» Link eine, die auf den interaktiven Standardbrowser abstellt.

5.2.2 Automatisch verfolgte Links

Andere Links werden automatisch verfolgt. In diese Klasse gehören die Links, die ein Bild eingebettet in den Text einer Seite erscheinen lassen oder die Links, die in den Frames eines Frameset verschiedene HTML-Dateien aufrufen. Auch diese Website[2] macht zur Darstel-

2 Bei der Online-Version dieses Textes.

lung eines Menüs Gebrauch von dieser Technik. Die Links werden also vom Standardbrowser automatisch verfolgt, das Ziel des Link wird abgerufen und in die Darstellung der Seite einbezogen. Das Verhalten ist dabei zwischen den Browsern verschieden (Lynx verfolgt keine Links für Inline-Bilder) und teilweise auch vom Benutzer einstellbar (Bilder automatisch laden ist heute Defaulteinstellung, noch bei Netscape 3.x war dies anders).

Insofern stellt auch hier die Klassifizierung als «automatisch verfolgt» auf Bedingungen ab, die nur besonders häufig aber nicht notwendig sind.

5.2.3 *Tiefe* Links

Interessanterweise ist in jüngerer Zeit eine Diskussion um so genannte «deep» links entstanden. Für den Autor stellen diese Links keine eigene Klasse dar, ein Kriterium zur Einstufung als «deep» oder «nicht deep» fehlt grundsätzlich.

Im Web sind von der technischen Natur her alle Seiten erst einmal gleich. Jede Seite die einen eigenen persistenten url hat ist eigenständig – auch wenn sie zu einem Verband gehört der von demselben Autor stammt. So ist denn auch ein Link auf eine solche Seite ohne weiteres technisch möglich. Die Links http://www.hase.net und http://www.hase.net/hase/transportnetz.html#tranportnetz sind also technisch genau das gleiche.

Es gibt verschiedene Möglichkeiten mit der ein Inhalts-Ersteller dafür sorgen kann, dass seine Inhalte zwar mit der Technik des Web erreichbar sind, aber nicht im verlinkbaren url-Raum liegen.

Die üblichste Technik ist die Verwendung eines Programms, das auf dem Webserver abläuft und Daten liefert. Für diesen Zweck wurde das Common Gateway Interface entworfen: Über diese Schnittstelle kann eine Webseite innerhalb des Webservers eine Aktion auslösen, die ein Programm startet. Die Ausgabe dieses Programms wird dann

an den Webserver und über diesen an den Browser des Benutzers weitergegeben, der die Seite mit dem CGI-Aufruf abgerufen hat. Die Daten, die über CGI-Programme geliefert werden, können in der Regel nicht Ziel eines Links sein, befinden sich also nicht im Link-Zielraum.

Die Diskussion um «tiefe» links («deep» links) ist aber nur entstanden, weil sich Autoren vorbehalten wollten, einen Querverweis auf Inhalte, die sie in den Linkzielraum eingestellt hatten, zu untersagen. Zum Leiden des Web als einer interessanten Technik hat sich ein (eventuell unkundiger?) Richter einer solchen Argumentation schon einmal gebeugt.

6 Napster

Eine hochinteressante Anwendung des Internet ist die Abkehr von klassischen Client-Server Strukturen und die Wendung hin zu einer «alles ist Server und Client» Struktur. In der reinen Client-Server Lehre ist die Rolle der Maschinen C und S streng getrennt: Der Server erbringt einen Dienst, der Client nutzt ihn. Der Client kann mehrere Dienste nutzen und er kann auch selbst Dienste erbringen (also selber Server sein); wenn er Dienste erbringt, dann jedoch andere als die, die er von seinen Servern bezieht.

Zu abstrakt? Ein Beispiel. Ein Webserver erbringt einen Dienst, nämlich http. Um dies zu tun greift er auf einen Fileserver zu auf dem die Daten gespeichert sind, die er dem Nutzer überträgt. Dieser Computer ist also Webserver und NFS-Client. Es gilt als untaugliches Systemdesign würde der Webserver seinerseits Dienste eines Webservers in Anspruch nehmen. Der Browser des Benutzers sollte dann besser direkt auf den Zielwebserver zugreifen. Auch darf ein NFS-Server einen Speicher, den er selbst per NFS importiert nicht wieder Dritten bereitstellen. Auch hier sollten diese Clients direkt auf den Originalserver zugreifen.

Diese Auszeichnung bestimmter Maschinen, die die Unterscheidung in «Client» und «Server» darstellt und die in der Regel den Server zum «reinen Server» erklärt, ist jedoch nicht in allen Fällen das ideale Design. So ist die verteile Datenspeicherung im LAN schon mit AppleTalk (auf der LocalTalk Infrastruktur) in einem Netz aus gleichberechtigten Maschinen realisiert. Die Daten werden dann da gespeichert wo sie am häufigsten zugegriffen werden, stehen aber auch an anderer Stelle über das Netz zur Verfügung. Diese «Peer to Peer» Netze finden sich mit AppleTalk und SMB bisher nur im LAN und waren beschränkt auf Fileserving und Printersharing.

Mit Napster und seinen Vettern wurden aber auch im globalen Internet Peer To Peer Anwendung realisiert. Berühmtheit hat Napster dadurch erlangt, dass diese Anwendung je nach Rechtsauffassung legal oder illegal ist und möglicherweise eine Grauzone ausnutzt. Wir wollen uns mit Napster daher an dieser Stelle etwas eingehender befassen.

6.1 Funktionsweise

Die Funktionsweise von Napster ist vergleichsweise einfach. Der Napster-Benutzer installiert zunächst auf seinem Computer eine Software, die zur Teilnahme am Dienst befähigt. So ähnlich wie ein Webbrowser einen Client für http und für ftp darstellt stellt die Napster-Software einen Client für den Napster-Verzeichnisdienst und den Napster Sharing-Dienst dar. Die Napster-Software stellt aber zugleich einen Server dar, der den Napster Sharing-Dienst anbietet.

Der Verzeichnisdienst wird von Napster an zentraler Stelle erbracht. Hier steht das Verzeichnis aller über den Dienst erreichbaren Dateien (Napster ist qua definitionem auf Dateien mit der Endung .mp3 beschränkt, das Verfahren jedoch ist universeller). So wie eine Suchmaschine ein Verzeichnis diverser Webserver darstellt, stellt der Napster-Server ein Verzeichnis diverser Napster-Sharing-Server dar;

die Sharing-Server sind wie oben dargestellt einfach alle Computer, die

- eine Kopie der Napster-Software ablaufen lassen
- mit dem Internet erreichbar verbunden sind.

Die Übertragung der Dateien erfolgt nun «Peer to Peer» also von einem der (tausenden) Napster-Sharing-Server direkt zum Napster-Sharing-Client.

Die Software ist dabei so programmiert, dass alle Dateien in demselben Verzeichnis zu liegen kommen, mithin alle heruntergeladenen Dateien wiederum im Server bereitstehen. Der Napster-Verzeichnisserver nimmt an der Übertragung der Dateien nicht teil; man kann jedoch sagen dass er die Übertragung katalysiert.

Andere Dienste arbeiten für die Dateiübertragung ähnlich wie Napster: Die Daten werden direkt zwischen den Computern zweier Anwender übertragen ohne einen zentralen Server. Bei Gnutella gelingt es zudem, auch den Verzeichnisdienst komplett zu dezentralisieren. Hier gibt es gar keine zentrale Komponente mehr, die eine Kontrolle ausüben könnte. Man kann sagen, dass Gnutella so ähnlich strukturiert ist wie das Internet selbst – ohne jede Zentrale.

6.2 Rechtliche Einordnung

Wie ist Napster nun rechtlich einzuordnen? Ohne eine zu diesem Thema fällige Doktorarbeit eines Juristen vorwegzunehmen will ein Ingenieur hier einmal einen Ansatz wagen.

6.2.1 Verzeichnisdienst

Der Napster-Verzeichnisdienst ist schlicht ein Katalog erstellt über diverse Server. Er ist rechtlich an sich völlig unbedenklich. Ein Verzeichnis meiner CD-Sammlung oder sonstiger Werke können die Urheber der Werke nicht untersagen. Dennoch wurde in einem Ur-

teil gegen Napster der Betrieb des Servers verboten, da er nach amerikanischer Rechtsauffassung eine Verletzung der Rechte der Urheber der verzeichneten Musikstücke darstellt. Das ist zunächst nicht nachvollziehbar. Mit CDDB steht ja auch ein Verzeichnisdienst bereit, der die Namen von Musikstücken verzeichnet. CDDB ermöglicht bei Angabe des Media-Code (einer eindeutigen Nummer, die eine Musik-CD identifiziert) diverse Angaben zur CD zu erhalten, so z.B. Titel, Interpret, Namen aller Tracks auf der CD. CDDB ist daher extrem nützlich, wenn man seine CD-Sammlung in das MP3-Format überführen will. Die «Ripper» Software, die die CD digital ausliest und dann als MP3 kodiert, kann aus CDDB die Bezeichnungen für alle Tracks ermitteln und die kodierten Musikstücke auf diese Weise sinnvoll benennen.

Die Argumentation zu Napster als Urheberrechtsverletzung ist etwas verzwickt, kann aber wie im Folgenden dargestellt nachvollzogen werden.

6.2.2 Spezielle Software

Die Teilnahme am Napster-Dienst erfordert die Verwendung der Napster-Software. Ohne die Software, die exklusiv von Napster angeboten wird, ist eine Teilnahme an Dienst nicht möglich. Insofern kann man die Software von Napster und den Betrieb des Verzeichnisservers als ursächlich für alles ansehen, was mit Hilfe dieser Software dann geschieht. Diese Argumentation ist offenbar sehr zweischneidig: Die Hersteller gefährlicher Güter wie «Auto» oder «druckverflüssigtes Propangas» könnten mit derselben Argumentation schnell ein Problem bekommen. Bei Software scheint man aber noch immer davon ausgehen zu können, dass der Benutzer keinerlei Kontrolle über das hat, was der Computer mit der Software tut.

6.2.3 Software erzwingt Freigabe

Der entscheidende Schritt in der Argumentation dürfte aber sein, dass die Napster Software eine Kontrolle durch den Benutzer der Software gezielt verhindert. Jeder Anwender der Napster-Software

macht seinen Computer unweigerlich zum Server. Es ist möglich, z.B. durch Zuhilfenahme einer Firewall für die Internet-Anbindung, eine Teilnahme am Napster-Verzeichnis zu ermöglichen und auch Dateien herunterzuladen – nicht aber selbst Dateien freizugeben. Eine derartige Konfiguration ist mit der Napster-Software allein nicht möglich. Es wäre technisch ohne weiteres möglich, dem Benutzer der Software die volle Kontrolle darüber zu geben, welche Daten er für «jedermann» freigeben will und welche nicht. Die Napster-Software verhindert dies. Insofern kann man die Software als ursächlich für die Urheberrechtsverletzungen ansehen, die die Anwender mit Hilfe dieser Software begehen.

6.3 Inhaltskontrolle

Das Urteil gegen Napster verbietet nicht den Dienst selbst; das wäre wohl auch unzulässig. Das Urteil fordert nur, dass Napster als Betreiberin des Verzeichnisdienstes verhindert, dass Urheberrechte verletzt werden. Das Urheberrecht für Musikstücke – und nur um diese geht es bei Napster – ist keine einfache Materie: Kopien von legal erworbenen Vertriebsstücken eines Musikstückes sind unter verschiedenen Bedingungen erlaubt und völlig legal. Einige dieser Bedingungen schliessen sogar eine Weitergabe an Dritte ein. Die Rechtslage ist in Europa und den USA leicht unterschiedlich.

Aber nicht nur diese Rechtslage erschwert einem Verzeichnisdienst das «Ausfiltern» von «illegalen» Inhalten aus einem Dienst wie Napster. Das Verzeichnis kennt von einem Stück nur den Namen und gibt diesen Namen an die Teilnehmer weiter. Der Name allein reicht aber als Kriterium für die Beurteilung der Frage «ist die Weitergabe zulässig oder nicht» wohl nicht aus. Zumal dasselbe Stück durchaus unter verschiedenen Namen im Dienst vorhanden sein kann: Tippfehler und willentliche Abkürzungen sind ja nicht ausgeschlossen. Es ist praktisch unmöglich ein Filter zu konzipieren, das auch nur einen Bruchteil der inkriminierten Stücke ausblendet.

Damit ist nach amerikanischem Recht offenbar der ganze Dienst anzuschalten.

7 Filtern des Internet

Eine Frage die immer wieder gestellt wird ist, warum denn niemand einen Computer so programmiert, dass er aus den im Internet angebotenen Inhalten einfach alles Unerwünschte ausfiltert. Solche Filter könnten – so die Argumentation – von den Providern betrieben werden und alle Inhalte vor der Weitergabe an die Kunden auf unerwünschte Inhalte untersuchen; Auszufilterndes würde dann schlicht unterdrückt.

In der Tat werden bei diversen Providern derartige Filter ja schon angeboten: «familienfreundliches» Surfen soll so ermöglicht werden. Wir wollen und mit der Theorie des Filterns hier kurz befassen.

7.1 Filter und Suchmaschine

Wie kann ein Computer Inhalte filtern? Wenn wir das theoretisch betrachten, dann ist das ganz einfache Logik (oder Mengenlehre, das ist das gleiche). Im Internet gibt es die Information I, die sich anhand des Kriteriums K zerlegen lässt. Die Gesamtmenge aller Information zerfällt in zwei Mengen: K trifft zu (IK) oder K trifft nicht zu (Ik).

$$I = IK+Ik$$

Diese Summe zeigt auch, dass jedes Quäntchen Information (das kann z.B. eine Webseite sein) immer zu einer der beiden Teilmengen gehören muss: K trifft zu oder trifft nicht zu, *tertium non est.* An dieser Stelle ergeben sich für den geübten Juristen die ersten Probleme. Eine «juristische Standardantwort» lautet ja «es kommt darauf an».

In der Tat werden in der Rechtswissenschaft Kriterien für die Bewertung einer Tat oft etwas unscharf formuliert. Das finale Urteil fällt dann ein Richter anhand einer Fülle von Informationen über den Fall: Die näheren Umstände, die verschärfend oder mildernd sein können. Ein Computer kann ein solches Urteilsvermögen nicht haben – ein Kriterium K trifft entweder zu oder nicht. Ein Filter würde nun – geeignet programmiert – alle Inhalte IK unterdrücken für K = «ist Kinderpornographie» oder für K = «ist Aufruf zum Rassenhass». Die Teilmenge Ik, also jede für die das Kriterium nicht zutrifft, würde ausgegeben.

Ein solches Filter würde aber auch eine ideale Suchmaschine darstellen. Man könnte sein Kriterium eingeben (z.B. «gutes Restaurant in Berlin Wilmersdorf») und würde alle Daten erhalten, für die das Kriterium zutrifft. Alle anderen würden unterdrückt.

Ganz offenbar existiert eine solche Maschine nicht: Sie alle haben schon mit Suchmaschinen gearbeitet und sich gefragt, wie doof ein Computer denn eigentlich sein kann. In der Tat ist ein Computer extrem dumm – trotz 30 Jahren Forschung an «künstlicher Intelligenz».[3]

Computer agieren rein formal, niemals inhaltlich. Ein Computer kann einen Text nicht verstehen. Er kann Formalia erkennen wie die Anzahl der Wörter oder den Ort, wo der Text gespeichert ist. Text ist hier in dem Sinne verwendet, in dem die Informatik das Wort üblicherweise verwendet; alle Sorten von Daten sind darin enthalten, auch Bilder, Laufbilder und Töne – alles was irgendwie digital verarbeitet wird. Wenn aber ein Computer nur formal agieren kann, dann muss ein Filter eben an formalen Kriterien ansetzen statt an inhaltlichen. Das Right Protection System (RPS) versucht genau dies.

3 Der Autor hält echte Intelligenz der menschlichen Art für Computer der uns bekannten Bauarten übrigens mit DOUGLAS HOFSTADTER für unmöglich.

7.2 RPS

Wir wollen im Folgenden ein System beschreiben und untersuchen, das anhand von urls Inhalte aus dem Internet bewertet und filtert. Andere Systeme (z.B. der Filterproxy der Etisalat, der Telefongesellschaft der Vereinigten Arabischen Emirate) arbeiten auch mit schlagwortbasierten Filtern; diese wollen wir hier nicht untersuchen. RPS ist – soweit dem Autor bekannt – ein System das allein anhand von urls arbeitet. Tatsächlich hat der Autor noch keinen RPS-Filter eingehend untersucht; die Überlegungen an dieser Stelle sind daher rein theoretisch und eventuell auf RPS nicht zutreffend.

Ein «Text» (wieder in der allgemeinen Bedeutung) im Internet hat in der Regel einen uniform resource locator. Dieser stellt ein formales Kriterium dar, anhand dessen man einen Text filtern kann: Man erstellt eine Liste mit urls, die Blacklist, und diese werden dann gefiltert. Da jeder Zugriff auf einen url immer mit einem TCP-Verbindungsaufbau beginnt muss ein Filtercomputer lediglich diese Verbindungswünsche analysieren, den url extrahieren und wenn dieser auf der Blacklist steht, das Paket mit den Verbindungswunsch verwerfen. Der Server erhält somit niemals den Verbindungswunsch und eine Verbindung kommt nicht zustande. Damit ist ein effektiver Filtermechnismus gegeben.

Wir wollen hier nicht die Probleme mit der Bewältigung der schieren Anzahl an Verbindungswünschen je Sekunde in einem Providernetz oder die der Anzahl der zu inkriminierenden urls diskutieren. Sehen wir uns lieber hases RPS-Absurder an. Zunächst ist festzustellen, dass immer der gesamte url bewertet werden muss: Einfach auf die IP-Adresse zu verkürzen ginge zu weit, denn unter derselben IP können ja mehrere Server laufen und völlig legale, nicht zu filternde Inhalte verbreiten (so laufen z.B. alle bei PSINet gehosteten Webserver in Europa unter nur 3 Adressen: je eine für jedes Hostingcenter. Hunderte Kunden teilen sich also eine IP.).

Wenn man nun den url einer resource, die man vor dem Filter schützen will, regelmässig ändert – sagen wir alle 15 Minuten – dann wird ein url-Filter absurd: Schon nach 15 Minuten wäre die Blacklist veraltet. Eine Änderung des url ist aber trivial möglich. An den Dateinamen hängt man einfach die aktuelle Uhrzeit an. Alle 15 Minuten lässt man ein Programm laufen, das diese Veränderung des Dateinamens vornimmt. Eine andere Variante ist, die Anzahl der Zugriffe auf die Datei, an den Dateinamen anzuhängen. Damit veränderte sich der url bei jedem Zugriff (!). Dennoch wäre für die Nutzer der «Piratensite» der Zugriff noch immer auf die gleiche Weise möglich. Eine Verzeichnisseite enthielte die (permanent variierenden) Links auf die inkriminierten Dateien.

Praktisch alle Filter sind auf diese Weise trivial auszuhebeln. So sind einige free-hosting Provider dazu übergegangen, Dateien mit der Endung «.mp3» nicht mehr zuzulassen. Ergo benennen unsere Hobbypiraten ihre Dateien in «.doc» und liefern eine Anleitung zum Umbenennen mit. Filter sind unwirksam. Eine Vorschrift, die Filter fordert, ist daher unverhältnismässig. In Deutschland wäre ein Filtergesetz damit verfassungswidrig.

Die These von der Unwirksamkeit wird eindrucksvoll bestätigt von den hilflos anmutenden Versuchen der Etisalat (des Telefon- und Internet-Monopolisten der Vereinigten Arabischen Emirate), ihren Internetservice frei von Inhalten pornographischer oder unislamischer Natur zu halten. In den UAE ist Meinungsfreiheit kein Verfassungsgut, ein Filter nicht nur erlaubt sondern vorgeschrieben. Es verpufft aber praktisch wirkungslos: Schon ein primitiver Web-Tunnel wie www.safeweb.com unterläuft einen Filterproxy.

7.3 Filterkontrolle

Besonders sensibel ist bei der Filterdiskussion das Thema, wer die Kriterien für die Filter festlegt. Angenommen die perfekte Suchma-

schine wie in Filter und Suchmaschine beschrieben existiere. Das Filter habe sogar Textverständnis und kann daher inhaltlich filtern. Wer würde nun die Filterkriterien für das Filter festlegen?

7.3.1 Zensur

Das Filter ist offenbar eine Maschine, keine Person. Die Maschine kann keine Verantwortung für ihr «Handeln» tragen. Die Verantwortung kommt immer der Person zu, die die Maschine betreibt. Im Fall des Filters wäre die Verantwortung also bei demjenigen, der die Kriterien des Filters festlegt. In Deutschland kann dies nicht der Staat sein, denn «eine Zensur findet nicht statt» sagt das Grundgesetz.

Eine Konstruktion über eine Indirektion ist denkbar und schon versucht worden. Die Internet Service Provider könnten zur Verantwortung gezogen werden für alle Inhalte, die sie transportieren. Sie könnten ja das perfekte Filter einsetzen um den Transport illegaler Inhalte zu verhindern. Dennoch wäre auch hier wieder der Staat die Institution die festlegte, welche Inhalte mindestens zu filtern wären; der Provider könnte darüber nur hinaus gehen. Auch dies wäre nach Ansicht des Autors klar eine staatliche Zensur.

7.3.2 Monopole

Wenn Filter nicht vorgeschrieben sind aber dennoch beim Provider implementiert werden – sei es als «vorauseilender Gehorsam», sei es als «Dienstleistung» für den Kunden – dann enthält das die Gefahr eines Missbrauchs. Ein Warnzeichen ist die Fusion von AOL – einem der grössten ISPs für Consumer – mit Time Warner – einem der grössten Contentanbieter. Ein «Kinderschutzfilter» kann hier leicht eingesetzt werden, um «gewaltverherrlichende» Comics der Konkurrenz in der Verbreitung zu behindern – rein zum Schutz der Kinder natürlich.

Einen interessanten Fall hat es schon gegeben: Die Website der Firma Disney – eines bekanntlich gewaltverherrlichenden und kin-

derschändenden Konzerns – hat es schon auf die Blacklist einer der selbsternannten Wächterorganisationen geschafft. Der Grund: Der Kinofilm «Arielle die Meerjungfrau» enthält angeblich ein einzelnes Bild von Arielle ohne den Muschel-BH. Derartige pornographische Nacktheit einer gezeichneten Meerjungfrau ist natürlich nicht zu dulden. Schliesslich sehen sich alle Kinder den Kinofilm nur einzelbildweise an. Ob die Videofassung das inkriminierte Einzelbild enthält ist dem Autor unbekannt.

Die Grundfrage ist auch hier: Wer legt die Kriterien fest für die Filter, die dann bestimme was ich zu sehen bekomme oder nicht? Auch wenn die perfekte Filtermaschine existiert ist das Problem der unerwünschten Inhalte im Internet nicht gelöst – es ist nur auf ein neues abgebildet.

8 Anhang – Technik des Internet

An dieser Stelle wollen wir verschiedene Techniken, die im Internet zur Anwendung kommen, kurz darstellen. Diese Abschnitte dienen als Referenz und an anderer Stelle wird ggf. darauf verwiesen.

8.1 URL

Der uniform resource locator url ist ein wichtiges Strukturelement des WWW: Ein url gibt eine Fundstelle für eine «resource» an. Eine resource kann ein Text/Bild/Ton sein oder ein Programm, das eine Ausgabe erzeugt. Die Form des url ist standardisiert (uniform) um ein universelles Adressierungsmittel zu schaffen. Der url ist wie folgt aufgebaut:

protocol**://**IP-Adresse/resource
Beispiel: http**://**www.hase.net**/**hase/vortrag/datei.html

Die fetten Zeichen sind dabei vorgeschriebene Trenner. Beachten Sie bitte, dass der erste slash (/) ein vorgeschriebener Trenner ist. Danach folgende slashes sind möglich und Bestandteil von «resource». Die IP-Adresse wird in der Regel als Name angegeben, der mit dem DNS aufgelöst wird – nach der DNS-Auflösung steht hier die übliche 32-Bit-Zahl. Der protocol-Teil erlaubt, verschiedene Protokolle zum Laden der Daten zu verwenden und in einer Webseite z.B. auf eine ftp-resource zu verweisen. Die Protokollspezifikation kann auch eingesetzt werden um dem Browser die Anweisung zu geben, ein externes Programm zu starten, z.B. einen eMail client (mailto: als protocol), einen telnet-client (telnet: als protocol) oder andere Anwendungen.

8.2 Usenet News

Befasst man sich mit dem Phänomen «Internet» – dieses stellt ja nicht nur eine Technik dar, seine stürmische Entwicklung lässt nur die Klassifikation als Phänomen zu – dann taucht hier und da auch der Begriff «Usenet» auf. An sich sind beide Begriffe Antipoden: Usenet bezeichnet einen Satz von Techniken (Protokolle, Applikationen), die zunächst unabhängig von den Internet-Techniken wie TCP, IP und so fort entwickelt wurden. Die Usenet-Protokolle und -Applikationen wurden erdacht als spezielle Kommunikationsplattform: sie waren auf einen bestimmten Zweck, bestimmte Trägermedien und Anwendungen optimiert. Diese Optimierungen machen news ungeeignet für bestimmte Anwendungen.

8.2.1 eMail

News wie wir sie heute kennen haben eine längere Entwicklungsgeschichte hinter sich. News sind eine Weiterentwicklung der eMail und verwenden daher an vielen Stellen ähnliche oder gleiche Techniken. Anfangs verwendeten eMail und news auch den gleichen Transportmechanismus. Wir wollen daher hier auch kurz die eMail be-

leuchten, da dieser Cousin zum Verständnis der Natur von news beiträgt.

8.2.1.1 Innerhalb eines Systems

In der Tat wurde eMail zunächst nicht zwischen Benutzern verschiedener Computer ausgetauscht sondern zwischen den Benutzern eines einzigen Systems. Vor der massenhaften Verbreitung der PCs waren ja Mehrbenutzersysteme der Klassen Mainframe und Minicomputer (Systeme wie IBM 360, IBM 380, Siemens BS2000, Nixdorf, Digital Equipment PDP11 und Vax und später zunehmend Unix-Systeme verschiedener Anbieter) der Stand der Technik und die regelmässig anzutreffende EDV-Ausstattung. Diese Systeme bedienten mehrere Benutzer gleichzeitig. Mit der Einführung des interaktiven Betriebes (Terminal, Arbeit direkt mit dem System, nicht über offline angelegte Jobs) und der Verbreitung der CRT-Terminals (Bildschirm statt Drucker) kam auch der Wunsch auf, dass die Benutzer des Systems untereinander Daten austauschen konnten. Hier wurde eMail geboren. eMail innerhalb eines Systems ist in der Regel recht einfach gestrickt: da alle Benutzer des Systeme über eindeutige Namen (Benutzernamen) verfügen können diese für die eindeutige Adressierung verwendet werden. Das System muss nun nur Mechanismen bereitstellen um über die Sicherheitsschranken zwischen den Benutzerbereichen hinweg Daten senden zu können – eine recht einfache Aufgabe.

8.2.1.2 Zwischen Systemen

Gerade die Hersteller der Minicomputer vom Range einer Vax stellten bei der Einführung von Vernetzung auch die Verteilung von eMail bereit. Diese Vernetzungen waren aber in der Regel beschränkt auf lokale Bereiche (LANs) und die Techniken zum eMail-Austausch gebunden an die spezielle Netzwerkarchitektur des Herstellers.

8.2.1.2.1 UUCP

Unix stellte im Gegensatz zu anderen Minicomputersystemen keine Mechanismen bereit, um eMail zwischen Computern auszutauschen.

Diesem Manko wurde – Unix-typisch – auf einfache Art abgeholfen. Da bei Unix die eMail innerhalb eines Systems schlicht Dateien (im heute noch anzutreffenden mbox-Format) innerhalb des Filesystems war reduzierte sich die Aufgabe der Übertragung darauf, einen Mechanismus zum Kopieren einer Datei zwischen zwei Systemen herzustellen. Dieser wurde dann unter dem Namen uucp – Abkürzung für Unix to Unix copy – hergestellt. Uucp verwendet zur Verbindung zweier Computer serielle Schnittstellen aus zwei Gründen: diese sind bei jedem Unix-System vorhanden (serielle sind die typischen Schnittstellen für Terminals an Unix-Maschinen) und sie eignen sich sowohl für die lokale Kommunikation über ein Kabel wie auch für die Datenfernübertragung via Modem. Netzwerkhardware wird nicht benötigt, was das Verfahren preiswert machte.

Uucp stellt nun nur den Übertragungsweg für Dateien bereit. Den Rest erledigt andere Software, die sich in das bei Unix schon vorhandene System für eMail einklinkt: mail mit dem Ziel in einer Mailbox auf einem entfernten System wird statt direkt in die mailbox des Benutzers in eine temporäre Datei geschrieben die in einem Verzeichnis lag, dessen Inhalt zyklisch an ein Nachbarsystem übertragen wird. Die Adressierung der eMail erfolgt dabei unter Angabe des Pfades der message in der Form Computer1!computer2!computer3!computer3!benutzer Diese sog. Bang-Pfade («bang» wird gern als Namen für das Ausrufezeichen verwendet, die «Adresse» gibt tatsächlich den Pfad zum Ziel an) können dabei gelegentlich sehr lang werden. Auch hat das System den Nachteil, dass die Namen aller teilnehmenden Computer eindeutig sein müssen, kein Name doppelt vergeben werden kann. Die Benutzernamen müssen nur innerhalb eines Systems eindeutig sein. Es ist leicht einzusehen, dass diese Form der Adressierung direkt vom Transportmechanismus hergeleitet ist: ein Computer, der eine message empfängt (sie liegt als Datei in seinem «Eingang vom Nachbar1» Verzeichnis) entfernt den Namen seines Nachbarn2, der jetzt am Anfang des Pfades steht und legt die Datei in das Verzeichnis «Ausgang für Nachbar2». Uucp

basiert darauf, dass die Inhalte dieser Verzeichnisse zyklisch an die Nachbarn übertragen werden.

Wie viele Protokolle stellt auch uucp eine Sicherung der Datenübertragung gegen Fehler bereit: gerade die Anwendung auf Modemstrecken macht dies erforderlich. Modems der 300 und 1200 Bit/s Klassen boten seinerzeit keine Fehlerkorrektur: sie stellten in der Tat nur die Umwandlung von Bits in Töne und zurück bereit.

8.2.1.2.2 Netz
Bald bildete sich ein Netz auf der Basis von uucp: Organisationen verschiedener Art tauschten eMail und news über uucp-Links aus, in der Regel über Telefon-Wählverbindungen und Modems. Der Bedarf an dieser Technik wurde so stark, dass Modemhersteller wie Telebit ihre fehlerkorrigierenden Modems mit speziellen uucp-features ausstatteten und so gegenüber dem Wettbewerb einen mächtigen Marktvorteil erzielten. Es bildeten sich auch verschiedene non-profit-Organisationen, die nur für die Verteilung von eMail und news betrieben wurden. Aus einer der bekanntesten dieser Organisationen namens UUNet ist später einer der grössten (kommerziellen) ISPs der Welt hervorgegangen.

8.2.2 Mailinglists
Email in der dargestellten Form stellt eine Kommunikation one-to-one bereit, ganz analog ihrem Vorbild der Post. eMail liegt jedoch als digitale Daten vor, kann also recht leicht kopiert werden (in der Tat ist der uucp-Transportmechanismus ein Kopiermechanismus). Es ist also ein leichtes ein Programm zu entwerfen, das alle eMail, die es empfängt, einfach neu versendet und zwar an alle Empfänger ,die auf einer Liste verzeichnet sind. Solche Programme wurden auch schon bald entwickelt und im uucp-Netz eingesetzt. Sie stellen eine Erweiterung des eMail-Konzeptes hin zu einer many-to-many Kommunikation dar. Die mailinglist Automaten stellen insofern etwas neues bereit: die one-to-many Variante der eMail kann man einfach als ein mehrfaches Versenden darstellen; in der Tat ist die einfachste

Implementation von Kopienversand (CC: header field) auch einfach, dass die message in mehrere mailboxen (bzw. mehrere Ausgangsverzeichnisse) kopiert wird. Mailinglists erweitern dies im Sinne einer vollwertigen many-to-many Kommunikation.

8.2.3 Bulletin board systems (BBS)

Recht schnell nach Einführung von Modems in den privaten Gebrauch wurden auch die Bulletin Board Systeme (BBS) entwickelt. Der Oberbegriff des BBS umfasst diverse verschiedene Systeme die sich im Funktions- und Leistungsumfang stark unterscheiden. Typisch für die deutsche Mailboxszene waren lange Zeit die «selbstgestrickten» Programme, of auf Maschinen der Klasse Apple][, C64 oder später Atari ST und MS-DOC PC programmiert und über genau ein Modem zugänglich. Spätere Entwicklungen unterstützten dann auch zwei oder mehr Modems und ermöglichten damit onlinechat neben den anderen Funktionen.

8.2.3.1 Funktionsumfang

Der typische Funktionsumfang eines BBS besteht aus drei Teilen:

- eMail
- bulletin boards
- file area

Die eMail ist dabei auf das System beschränkt. Man kann Nachrichten von einem Benutzer an einen anderen senden. Andere Benutzer können diese Nachrichten nicht lesen, der «SysOp», der Administrator – und oft auch Eigner des gesamten Systems – aber sehr wohl. Die bulletin boards stellen so etwas wie elektronische Varianten einer Pinwand dar: hier von einem Benutzer eingestellte Nachrichten können von jedem anderen Benutzer gelesen werden. Die file area stellt Dateien zum download bereit. In der Regel handelt es sich dabei um Raubkopien meist von Spielen für die aktuelle Gerätegeneration. Genau darin liegt denn auch oft die Motivation der privaten Betreiber einer «Box»: er erhält Zugang zu allen eingestellten Dateien ohne den Aufwand, selbst im Dschungel der Boxen suchen zu

müssen. Oft sind für den Download «tit-for-tat» Regeln implementiert: herunterladen darf nur, wer zuvor eine gewisse Datenmenge in das System gestellt hat. Für den Zugang zum BBS und zur eMail war in der Regel nur ein Terminal oder Computer mit Terminalprogramm einfachster Art erforderlich[4]. Andere Systeme verwenden eine spezielle Software auf der Client-Seite und stellen darüber zusätzliche Funktionen wie Menüsteuerung oder einfaches Multitasking (gleichzeitig Dateien übertragen und eMail/boards lesen) bereit. Auch Funktionen wie offline-lesen und offline-posten wurden über Systeme mit speziellen Clients realisiert.

8.2.3.2 Wirkungsweise

Bulletin boards und file area bedürfen der besonderen Betrachtung. Der SysOp liest in der Regel alle Einträge in den BBs mit und nimmt eine Überwachungsfunktion wahr. Durch Steuerung der Benutzer («User» im Jargon) mittels Ermahnungen (per persönlicher eMail innerhalb des Systems), Zensur (Entfernen von Beiträgen) bis hin zum Ausschluss von der Teilnahme (Usersperre) bestimmt er, was in seinem Reich geschieht. Für jede «Box» stellt sich so eine Nutzerschar ein, die was Meinung und Verhalten angeht ein recht geschlossenes Bild abgibt. Wie schon angesprochen stellt die file area oft die Motivation für den Betrieb einer «Box» – immerhin ist eine signifikante Menge an Hardware erforderlich, um eine für die User interessante Box zu betreiben – dar, da so der Betreiber Zugang zu allen Daten erhält. In diversen Fällen[5] ist auch das Gefühl der Meinungsherrschaft für die boards die Triebfeder, teilweise ist es Altruismus, in der Mehrzahl wohl jedoch der Zugang zu den eingestellten Raubkopien.

---

4 Ein Terminal (deutsch oft «Datensichtgerät» verarbeitet keine Daten sondern stellt diese nur dar: Tastendrücke werden direkt über die Schnittstelle gesendet, eingehende Daten unmittelbar dargestellt. Frühe Terminals stellen nur Text dar, Grafikterminals galten lange als «Luxus» und wurden noch vor einer weiten Verbreitung von den PCs ersetzt. Terminals verfügen nicht über einen Massenspeicher.
Ein Terminalprogramm verwandelt den Computer in ein solches «dummes» Gerät, jedoch mit dem Zusatznutzen, Dateien übertragen zu können zwischen dem lokalen und dem entfernten Massenspeicher.

5 Typisch ist hier der Teenager, der mit einem irgendwo abgeschwatztem Mailbox-programm und einem C64 mit Floppylaufwerk agiert.

Interessant sind die Boxen, bei denen der SysOp per seiner Vorherrschaft die Verbreitung von Raubkopien verbietet, die file area also nur für die Verbreitung von Public Domain Software dient. Die Motivation dieser Betreiber ist schwer zu erfassen. Die Beschränkung der Ursysteme auf ein Modem und die zur Hochzeit der Mailboxszene übliche Modembandbreite von 2400 Bit je Sekunde sorgt für eine starke Beschränkung der Nutzerzahl (da bei der Anwahl oft besetzt ist) und damit für einen sehr privaten Kreis. Man kennt sich, man kann miteinander oder man sucht sich eben eine andere Box. Typische Boxen hatte Nutzerzahlen weit unter Hundert, davon in der Regel kaum mehr als zwei Dutzend wirklich «aktiv», also regelmässig online.

8.2.3.3 Illegalität

Neben der grenzenlosen Kommunikation – alle User sind gleich, jeder stellt nur dar was er äussert, keinen Geruch, keinen Anzug, keinen Rang[6], keine Stellung[7] – begründen wohl hauptsächlich zwei Faktoren die Faszination für diese Spielzeuge (der Nutzeffekt der Mailboxen dürfte eher gering einzuschätzen sein). Der eine ist die neue Form der Kommunikation zwischen Schrift und gesprochener Sprache. Der Autor nennt dies gern den eMail Stil, auch wenn eMail sich mit der Mauser zum Kommunikationssystem des business sehr der formalen Schriftform genähert hat, so ist dennoch etwas von verblieben davon, dass eMail eben nicht ein formeller Brief ist. In Deutschland und auch anderen Ländern Europas kam einer weiterer wichtiger Punkt hinzu: fast alle Modems zum Betrieb von Mailboxen galten als illegal, da sie keine Zulassung der Postbehörden besassen. Das Gefühl, an etwas subversivem zu partizipieren – das doch ganz offensichtlich völlig harmlos war – darf als Motivation für diverse Teenager-Aktivitäten in der Mailboxszene gelten. Die relativ starke Verbreitung von Raubkopien unter der Teilnehmerschaft der Mailboxszene (teils per DFÜ, teils direkt über per Mailbox anbe-

6 Einige Systeme implementieren so etwas wie eine Rangordnung. Stellung erlaubt da oft längere Online-Zeiten oder grössere Downloads.

7 Ausnahme ist hier der «gottgleiche» SysOp, der in seinem Reich allein herrscht.

raumter Kontakte) hängt vermutlich mit der durch das Eigenschaftspaar (illegal aber harmlos) der Boxen zusammen, das eine Hemmschwelle und ein Unrechtsbewusstsein senkt. Möglicherweise ist dies aber auch nur Spekulation und gar nicht wahr.

8.2.3.4 Fido Net

Aus der Mailboxszene geboren ist das Fido Net als ein Netz von BBS entstanden. Die Idee ist dabei, dass die BBS die Inhalte der Bulletin Boards mit den Nachbarn synchronisieren, letztlich also ein BBS bilden, das mit mehreren räumlich verteilten Zugangspunkten arbeitet. Jedes am Fido[8] Netz angeschlossene System («Point») steht dabei in separater Eigentümerschaft und Verwaltung und tauscht mit einer begrenzten Anzahl Nachbarn Daten aus. Für den User stellt sich ein Fido Point wie eine klassisches BBS dar, nur eben mit grösserem Angebot an Nachrichten in den boards. Die Fido Software wurde für MS-DOS entwickelt und lief nur dort. Das Protokoll zum Datenaustausch mit den Nachbarn war integraler Bestandteil der Software. Zur Teilnahme an einem Fido-System war – anders al bei vielen Mailboxen – ebenfalls eine spezielle Software erforderlich. Diese erlaubte insbesondere das telefonkostensparende offline-Lesen von BBS-Inhalten und das offline-Posten; zu diesem Zweck werden die Daten dann während eine Verbindung besteht mit maximaler Datenrate übertragen (nicht mit der Lesegeschwindigkeit des Benutzers). In der Heimat des Fido-Netzes, den USA, machte dies die Mailboxnutzung über Fernverbindungen finanziell erst möglich. Wie bei vielen BBS-Systemen hat ein Fido-Point typisch genau ein Modem, das für die Synchronisation mit allen Nachbarn und für die Datenübertragung zu allen Usern verwendet wurde. Spätere Versionen der Software unterstützen auch bis zu vier Modems, so viele wie MS-DOS Schnittstellen unterstützt. Die Anzahl der Modems und ihre Bandbreite stellt wiederum eine Beschränkung des Systems dar, sodass die Tragweite nicht sehr gross ist. Durch die Grösse des Netzes

8 Benannt ist das System nach dem Hund des Entwicklers der Fido-Software.

stellt sich jedoch schon eine starke Anonymität der Benutzer voneinander ein.

8.2.3.4.1 Organisation
Die Points des Fido Netzes stehen in separater Eigentümerschaft und Verwaltung. Jeder Betreiber legt im Prinzip selbst fest, wie er mit seinem System umgeht. Der Entwickler des Fido-Systems machte jedoch die Unterwerfung in bestimmte zunächst von ihm, dann von der «Gemeinschaft der Sysops» aufgestellte Regeln zur Lizenzbedingung für die Benutzung seiner Software – und damit zur Standesregel im Fido-Netz

8.2.4 Notes
Das Verfahren mit Mailinglisten ist nicht besonders effizient: mehrere Kopien derselben message, die denselben Weg haben, belegen in der Tat mehrmals die Bandbreite der Leitung. Ausgehend von Modemverbindungen mit 1200 bis 9600 Bit/s Übertragungskapazität und Ferngesprächsgebühren kann das schnell eine teurer Spass werden[9]. Daher wurde ein anderes Verfahren ausgedacht, das den many-to-many Kommunikationsansatz (in der relativ geschlossenen Gruppe der Listenteilnehmer) durch eine Variante des klassischen Broadcast ersetzt. Das erste dem Autor bekannte Verfahren trug den Namen notes, wurde aber relativ bald von news abgelöst. Da es sich in seinen Eigenschaften von news nur geringfügig unterscheidet, soll es bei dieser Erwähnung bleiben.

8.2.5 News
8.2.5.1 Basis
Die mangelnde Effizienz und der relativ hohe Verwaltungsaufwand für Mailinglisten – zunächst wurden die Teilnehmerlisten in der Regel von Hand gepflegt – liess einen neuen Ansatz aufkommen. Die üblicherweise themenbezogene Mailingliste wird abgebildet auf eine

9 Der Autor hat zu gegebener Zeit massive Probleme mit der Bewältigung seiner Telefonrechnung gehabt. In der Tat konnte uucp schnell das Budget eines Schülers sprengen, aber auch das von Firmen, die solche Spielerei betrieben.

Datenbank zum Thema. Die Datenbank enthält dabei messages (postings). Zu jedem posting werden einige Daten erfasst, die nicht zum Inhalt gehören wie Verfasser (zunächst in der Form des Bang-Pfades), ein Zeitstempel und eine message-ID, die eindeutig ist. Mit uucp steht ja ein (universeller) Mechanismus zum Kopieren von Daten zwischen Systemen bereit, der auch news zugrunde liegt. Eine message, die auf einem System S1 erzeugt wird, wird in die lokale Datenbank[10] eingestellt und wird zusätzlich in alle «Ausgang für Nachbar x» Verzeichnisse kopiert. Damit ist – aus Sicht der Software – die message an die Nachbarn verteilt, selbst wenn sie noch lokal gespeichert ist. Beim folgenden zyklischen Datenabgleich erfragt nun der Nachbar, welche messages im «Ausgangskorb» für ihn anstehen und ruft diejenigen ab, die er noch nicht auf anderen Wegen empfangen hat. Dann löscht er alle für ihn gespoolten Daten, leert also den Ausgangskorb. Alle so eingehenden messages legt das System wiederum in alle lokalen Ausgangskörbe. In einem typischen Szenario verteilt sich eine message also in einem Schneeballsystem, da jeder Empfänger die message an alle Nachbarn weiter verteilt.

Die besondere Effizienz des Systems liegt darin, dass so eine message, die mehrere Nutzer des System S2 lesen, nur einmal an S2 übertragen wird: selbst wenn dieselbe message auf verschiedenen Wegen hereinkommen hätte können (da mehrere Nachbarn sie gespoolt, also in den «Ausgangskorb» gelegt hatten) wird sie nur einmal übertragen: vor der Übertragung erfolgt ein Abgleich der anstehenden message-IDs, der Empfänger kann also schon empfangene messages von der Übertragung ausnehmen.

Die Grundzüge des Systems sind damit beschrieben: im Schneeballsystem werden alle auf einem System S1 lokal erzeugten oder von einem Nachbarn empfangenen messages an alle Nachbarn verteilt. Die Lokalität der Nachbarbeziehungen und Verfahren, jede message nur einmal zu übertragen, sorgen dabei für eine gegenüber Mailing-

10 In der Regel ist das einfach ein Verzeichnisbaum im Dateisystem, da die Daten keine Relationen haben.

lists stark verbesserte Effizienz der Bandbreitennutzung. Das System news stellt also eine Fortschreibung dessen dar, was Mailboxsysteme schon darstellen: elektronische Pinwände. Nur werden die Inhalte der «Pinwände» (bulletin boards) nun zwischen verschiedenen Systemen verteilt – ganz ähnlich wie bei Fido Netz und anderen, ähnlichen Entwicklungen.[11]

News erreichen aber eine sehr viel stärkere Verbreitung als Fido oder vergleichbare Systeme. Der Grund liegt wohl darin, dass news als System auf Unix-Maschinen entstanden sind (wie auch uucp) und diese klassische Mehrbenutzersysteme darstellen. News war also ein verteiltes «BBS» für Maschinen, auf die in der Regel viele Benutzer Zugriff hatten, auch gleichzeitig. Damit war effektiv die Bandbreitengrenze der Fido points (Anzahl Modems) aufgehoben. Diese Genealogie aus der eMail heraus ist in der Tat eine Vereinfachung. Das news-System ist eine parallele Entwicklung zu Mailbox-Systemen wie Fido Net, (geschlossenen, kommerziellen) Online-Diensten wie Compuserve und später AOL und Entwicklungen in der Netztechnik. Dennoch hilft diese Darstellung vielleicht zum Verständnis: news stellen ein Schneeballsystem dar. Dabei führen in der Regel mehrere Wege zum gleichen Knoten (ein Weg ist eine Folge von Knoten und Kanten des Netzes). Diese führt zu einer starken Fehlertoleranz (ausgefallene Kanten oder Knoten stellen das System nicht in Frage) aber auch dazu, dass der einzelne Administrator eines news-Knotens praktisch keinen Einfluss auf das System als solches hat. Sein Einfluss ist beschränkt auf die Nutzer seines Systems.

8.2.5.2 Details

Bei den Varianten B-news und C-news kommen nun noch einige Features zu diesem Grundkonzept hinzu. Wie schon erwähnt werden die messages nach Themenbezug in die lokale Datenstruktur eingestellt. Diesen Themenbezug stellt der Benutzer her, indem er die message mit einem bestimmten Thema assoziiert. Der tatsächliche

11 In Deutschland haben sich diverse Mailboxverbunde gebildet, darunter Znetz und das Maus Netz.

Themenbezug wird damit nicht ausgedrückt. Zur Ordnung der Datenbank wird eine hierarchische Klassifikation von Themen verwendet. Die oberste Klassenebene stellt dabei Grundkonzepte wie

- Computer – comp
- Recreation – rec
- Social – soc
- Miscelaneous – misc

dar. In jeder dieser Klasen findet auf der zweiten Ebene eine weitere Unterteilung statt, z.B. in

- Betriebssysteme – os
- Systeme – sys
- Peripherie – periph

etc statt, die dann weiter unterteilt werden. So ergeben sich für die «Bretter» (engl. Boards von der Pinwand-Analogie) dann Namen wir comp.sys.atari.st, comp.os.mac, rec.humor.funny, soc.law.rape oder alt.fan.swedish-chef.dork.dork.dork. Die Newshierarchie unter «alt.» bildete sich aus einer Revoluzzer-Haltung zu einer Zeit, da die allmächtigen Administratoren des Systems ziemlich elitär entschieden, welche Gruppen im System geführt werden sollten und welche nicht. Irgendwie wurde die «news-Subkultur» unter alt. dann doch schnell Teil der news-Gesamtkultur.

8.2.5.2.1 Eingang ins Internet

News ist im Prinzip unabhängig vom Transportmedium. Uucp stellt nur einen Mechanismus für das Kopieren von Dateien bereit und mehr benötigt news nicht. Mit der Verbreitung des Internet und einer Anwendung http (WWW) wurde zunächst angenommen, dass news stark in den Hintergrund treten und praktisch ersetzt würden. Das Gegenteil war der Fall – aus Gründen die dem Autor nicht einleuchten wollen. News sind im Prinzip ein archaisches Medium der one to many Kommunikation, da sie auf Text beschränkt sind[12]. Die schiere Masse von news scheint dem Autor auch den Nutzwert zu

12 Andere Datentypen werden zunächst in eine Textform umkodiert, dieser Text wird dann übertragen. Ein news-Beitrag kann nur einen Datentyp enthalten, eine Mischung wie bei HTML üblich ist nicht möglich.

erdrücken: einem Thread (Diskussionsfaden) kann man kaum folgen, will man nicht sehr viel Zeit investieren. Doch scheint es ausreichend viele Benutzer zu geben. Was den Transportmechanismus angeht ist news wie oben dargelegt nicht sehr anspruchsvoll: jede Technik, die es erlaubt, Dateien zwischen zwei Maschinen zu kopieren ist im Prinzip geeignet. Die ersten Portierungen auf die neue Kommunikationsinfrastruktur TCP (statt Modemleitung) verwenden die TCP-Verbindung denn auch um darüber mit uucp Daten zu kopieren. Spätere Weiterentwicklungen effektivieren die Übertragung indem sie die Fehlerkorrekturverfahren von uucp nicht mehr verwenden; der TCP-Kanal ist bereits fehlerbereinigt und zuverlässig.

8.2.5.2.2 Auswahl

Ein Administrator eines Systems, das news für Benutzer bereitstellt[13] kann festlegen, welche Gruppen/Boards auf seinem System geführt werden und welche nicht. Im Prinzip stellen B- und C-news einen Mechanismus bereit, neue Gruppen automatisch anzulegen: spezielle messages in der speziellen Gruppe «control» (die jedes System führen muss) legen ein neues Board an. Es steht dem Administrator jedoch frei, diese Messages manuell auszuwerten und so die Entscheidung zu treffen. Ist einmal eine Gruppe geführt, dann greift wieder der o.a. Mechanismus: alle anstehenden (neuen) Messages werden beim Polling-Zyklus übertragen. Abgewiesen werden nur solche messages, die schon übertragen wurden.

8.2.5.2.3 Speicherzeit

Bisher haben wir nur beleuchtet, wie eine message in die Datenbank eines news-servers gelangt, nicht wie sie wieder entfernt wird. Im wesentlichen entfernen zwei Mechanismen eine message vom Server: «expire» und «cancel». Cancel beschreibt eine message, die in der speziellen Gruppe control gesendet wird und die eine message-ID enthält. Die so angegebene message wird dann aus dem Server,

13 Ein System, das einen Dienst bereitstellt, wird gern als «Server» bezeichnet. Das Wort Server hat im Zusammenhang mit IP-Netzen jedoch auch die Bedeutung «Prozess, der einen Dienst bereitstellt». Beide sind begrifflich nahe verwandt aber dennoch zu trennen.

in dessen control die cancel auftaucht, entfernt und als «bekannt und bereits entfernt» in der Datenbank verzeichnet (sie wird also auch in der Zukunft nicht übertragen). Das Recht zu einem cancel wird in der Regel nur den Autoren der entsprechenden message eingeräumt: wer sich eines besseren besinnt, der kann so versuchen, den Schaden durch seine Äusserungen zu minimieren.

Expire ist ein zyklisch laufender Prozess, der alle messages älter als eine gewisse Schwelle löscht. Die Datenbank verzeichnet dann immer noch die ID der message als «bekannt», sie wird also nicht neu übertragen sondern über den o.a. Abweisungsmechanismus permanent von der lokalen Datenbank ferngehalten. Expire ist an sich einfach nur ein Programm, das zyklisch angestossen wird. Einige Implementationen erlauben nur die Angabe einer globalen Speicherzeit, die dann für alle messages in allen Gruppen gilt. In der Regel ist eine Einstellung der Speicherzeit für jede Gruppe individuell vorgesehen. Für neue Gruppen wird dann in der Regel die Einstellung einer höheren Hierarchieebene übernommen. Andere expire-Versionen erlauben auch, den Speicherbedarf einer Gruppe zu begrenzen: immer wenn die Speicher-Quota für so eine Gruppe überschritten ist, dann fliegen die ältesten Artikel raus. Typische Speicherzeiten messen sich in Wochen für die internen newsgroups eines Providers (wie snafu. technik, snafu.announce des Berliner Providers Interactive Networx[14] oder Minuten wie für «da läuft doch eh nur immer derselbe Müll durch» Gruppen wie alt.binaries.pictures.

8.2.5.2.4 Moderation

Bisher haben wir nur den vollautomatischen Teil von news beschrieben, der die Masse der messages enthält. News stellt seit den Anfängen auch einen Mechanismus bereit, der die automatische Verteilung durch eine menschliche Kontrolle ergänzt: die Moderation. Erhält eine Gruppe den Status «moderated», dann hat nur eine Benutzer das Recht, messages in diese Gruppe zu stellen. Versucht ein

14 Interactive Networx ist heute im deutschen Zweig von Inter.net aufgegangen.

anderer Benutzer eine message einzugeben (zu «posten»), dann wird diese automatisch in eine eMail an den Moderator umgewandelt. Der Moderator entscheidet dann, welche messages zur Veröffentlichung gelangen. Es gilt dabei als guter Moderationsstil, nicht als Redakteur (Editor) aufzutreten sondern eine message entweder unmodifiziert einzustellen oder gar nicht. Der Moderator ist jedoch im Prinzip frei, selbst zu entscheiden; Moderatoren mit einer unglücklichen Hand werden bald ihre «Kunden» verlieren, zumal die Aufmerksamkeitsspanne bei news besonders gering ist.

8.2.5.2.5 Form und Inhalt

Ein Benutzer von news gibt beim Verfassen an, in welchen Gruppen sein Beitrag erscheinen soll. Es ist möglich, eine message in mehreren Gruppen erscheinen zu lassen. Moderne Implementationen von news speichern die message dann nur einmal und entfernen sie, wenn der längste Speicherzeit (expire) überschritten ist. Solche crossspostings werden aber in der Regel nicht gern gesehen. Bei moderierten Gruppen kann der Moderator noch entscheiden, ob die message ins Thema der Gruppe passt. Wenn nicht, verwirft er sie. Bei den üblichen (unmoderierten) Gruppen wird eine solche Entscheidung nicht getroffen: eine Message, die nur aus dem Satz «diese message hat gar nichts mit Berlin zu tun» besteht, kann also ohne weiteres in der Gruppe «rec.restaurants.berlin» auftauchen. Formale Kriterien der message wie message-ID, Gruppe, Zeitstempel (Uhrzeit beim Posten) können nicht für eine Bewertung des Inhalten herangezogen werden.

8.2.5.3 D-news

Eine Weiterentwicklung der Version C des news-systems verabschiedet sich von der historischen Wurzel in einem Wählleitungs-Pollingsystem. Solange ein newsfeed nicht mehr als ein paar Megabytes am Tag liefert, ist das Verfahren von C-news durchaus brauchbar. Der Nachteil von C-News tritt aber gerade mit der Steigerung des newsvolumens (ein typischer newsfeed liegt heute bei über 500 Gigabytes im Monat, also über 2 Megabit je Sekunde im Monatsmittel) zu

Tage: da der server sämtliche verfügbaren Daten bei jedem poll abholt muss er auch all das speichern, das niemals gelesen wird.

D-news adressiert diesen Nachteil, indem nur diejenigen Daten beschafft werden, die auch tatsächlich abgerufen (gelesen) werden. Im Prinzip funktioniert ein server mit D-news ähnlich einem cache, allerdings mit einem Unterschied: der server beschafft sich von seinem feed (jenem Nachbarn, der Daten bereitstellt; in der Regel sind das mehrere) zunächst das gesamte Inhaltsverzeichnis. Er überträgt also die Liste aller verfügbaren messages an der Stelle, wo C-news die gesamten Artikel holen würde. Das so erstellte Inhaltsverzeichnis wird dem Nutzer dargeboten, ganz so wie beim älteren Verfahren. Ruft nun ein Nutzer einen Artikel ab, dann wird zunächst geprüft, ob der Artikel im lokalen Speicher verfügbar ist; falls nicht, wird er on-the-fly beschafft. In beiden Fällen wird er dann dem Nutzer übertragen. Dermassen einmal beschaffte Artikel werden lokal gespeichert ganz wie bei den älteren Varianten. Auch Artikel, die Benutzer des lokalen Servers verfassen und posten werden lokal gespeichert und den Nachbarn bereitgestellt. Der Expire-Prozess wird wie bei C-news regelmässig angestossen um veraltete Artikel zu entfernen. Dabei werden auch Artikel «entfernt», die lokal gar nicht gespeichert sind sondern nur im Inhalt verzeichnet. Aus Sicht eines news-client (Programm) ist zwischen C-news und D-news server kein Unterschied, beide verhalten sich gleich. Für den Betreiber ist der Unterschied, dass er nur solche Artikel übertragen muss und speichert, die wirklich abgerufen werden (oder lokal verfasst wurden). Diese gewaltige Ersparnis von Bandbreite und Speicherkapazität bei gleicher Funktionalität hat für eine rasche Verbreitung dieses Konzeptes gesorgt, die ältere Variante ist aber auch noch anzutreffen.

8.3 Providerhaftung

Betrachtet man die deutsche Gesetzgebung, dann fällt im TKG der «Proxy-Paragraph» auf. Der Gesetzgeber stellt darin einen Netzbe-

treiber frei von der Verantwortung für Inhalte, die er auf Nutzeranforderung durchleitet und für diesen Zweck ggf. kurzzeitig speichert. Die anwendbare Definition von «kurzzeitig» muss sicher noch die Rechtsprechung klären. Vergleichen wir C-news und D-news, dann fällt auf, dass bei ersteren Daten schon vor einer Nutzeranforderung übertragen werden, ja ohne dass eine Anforderung eines Nutzers des lokalen Servers überhaupt je vorliegt: Artikel können beschafft, gespeichert und wieder gelöscht werden ohne dass ein einziger Nutzer sie je abgerufen hätte. Das Verfahren bei D-news dagegen beschafft Artikel nur aufgrund einer Nutzeranforderung: auch das Einstellen (posten) durch einen lokalen Nutzer müssen wir als Nutzeranforderung betrachten. Damit scheint nach deutscher Rechtslage der Fall klar: der Provider beschafft Daten auch ohne Anforderung eines Nutzers, die Freistellung für Proxy-Inhalte greift also nicht; damit wäre eine Verantwortung des Providers für alle Inhalte im Rahmen der geltenden Gesetze gegeben.

Eine solche Konstruktion könnte aber fatal sein: die Daten werden unabhängig von ihrem Inhalt rein anhand formaler Parameter (Zugehörigkeit zu einer geführten Gruppe) beschafft durch einen automatischen Prozess. Der Name einer Gruppe und ihr Inhalt haben in der Regel einen losen Zusammenhang. Dieser besteht jedoch nicht in Strenge: verabreden sich z.B. zwei Individuen (z.B. mittels persönlicher eMail) einen Datenaustausch über eine bestimmte news Gruppe zu führen, dann können sie jede Gruppe für ihre Zwecke missbrauchen[15]: so könnten Raubkopien von Software (z.B. Musikstücken) in der Gruppe rec. humor.funny gepostet werden, dem Computer ist alles gleich. Einem ISP ist es in der Regel nicht möglich festzustellen, wer einen Artikel gepostet hat. Für Artikel, die auf dem lokalen Server originieren ist das oft noch möglich (auch wenn teilweise Datenschutzgesetze Einschränkungen für die Sammlung von Nutzungsdaten auferlegen), für einen Artikel der auf einem fremden server gepostet wurde ist es praktisch unmöglich. Konstru-

15 Derartiges kann sinnvoll sein, wenn z.B. die Grösse der Mailbox oder bestimmte Tarifmodelle für eMail eine direkte Übertragung nicht erlauben oder wenn eine Anonymisierung gewünscht ist.

iert man nun eine Verantwortung der Provider analog der Produkthaftungs- und Gewährleistungsverantwortung von Importeuren konstruieren, dann kommt das weitgehend einem Verbot bestimmter Kommunikationsformen gleich. Die deutsche Gesetzgebung ist hier bereits in einem Spagat: einerseits wird die Möglichkeit zur anonymen Nutzung der Kommunikationsangebote festgelegt, gleichzeitig eine Ausweispflicht gefordert. Der Spagat ist bei klassischen Medien nicht sichtbar: die Schwelle zur Publikation lieg bei Büchern, Zeitschriften etc. doch recht hoch. Die dadurch relativ kleine Zahl von Autoren wird dann noch einmal auf eine kleinere Zahl von Quellen (Verlage, Sender, Agenturen) abgebildet und hier setzen die klassischen Mechanismen der Kommunikationskontrolle (Sendeeinschränkungen/Altersbegrenzungen, Verbote) an. Mit news ist aber ein Medium entstanden das – anders als das WWW! – Leser und Autoren in gleicher Grössenordnung enthält und gleichzeitig eine sehr geringe Zugangsschwelle hat. Das führt unter anderem auch zu den berüchtigten flame-wars und zu Gruppenkasper-Erscheinungen, aber insgesamt sind news aufgrund der chaotisch-unkanalisierten Struktur ihrer Quellen, Transportwege und Senken sehr frei. Sie sind daher kaum «ein wenig» zu zügeln; eine Providerhaftung für die Inhalte würde sie schlicht ganz verbieten. Das dürfte – news Benutzer haben einen sehr jungen Altersquerschnitt – schnell zu Trotzreaktionen und ausweichen auf andere, ebenfalls nicht kontrollierbare System führen, das Phänomen dürfte es nicht stoppen. Nebenwirkungen eines (de facto oder de lege) Verbotes wären aber immer die Sündenbockhaftung sowie eine insgesamte Senkung von Unrechtsbewusstsein bei der Kommunikation. Letztere ist sogar an prominenter Stelle zu bewundern: die schillernde Leitfigur des grössten unabhängigen ISP PSINet WILLIAM SCHRADER weigert sich militant irgendwie an einer Kontrolle von Kommunikation oder Verfolgung von Straftätern mitzuwirken. Dies folgt der amerikanischen Rechtstradition wonach das krasse Gegenteil einer als «schlecht» erkannten Handlungsweise offenbar uneingeschränkt gut sein muss.

So radikal wird man nicht sein müssen oder dürfen. News sind sehr unter Beschuss geraten aus verschiedenen Gründen. Sie geben jedoch auch immer wieder Anlass zu Optimismus: immerhin hat sich dieses System über mehr als ein und ein halbes Jahrzehnt nun quasi selbst verwaltet und hat sich eigene, sehr akzeptable Standesregeln gegeben (die «netiquette»). Die news-Gemeinde hat sich in dieser Zeit als resistent gegen diverse parasitäre Erscheinungen wie Flamer, Lamer und Trottel gezeigt – an sich eine sehr erstaunliche Leistung wenn man nur an das Schlechte im Menschen wirklich glaubt. Diese erhaltenswerten Tugenden lassen also Augenmass bei der Gesetzgebung und bei der richterlichen Beurteilung von news und Usenet geboten erscheinen.

8.4 IP Transportnetz

Basis des Internet ist das weltweite Netz das sich ergibt, wenn man die (in der Regel kommerziell betriebenen) Netze der verschiedenen Provider zum «Internet» zusammenfasst. Interessanterweise sind die Provider alle gemeinsam am Entstehen des Internet beteiligt: jeder transportiert auf seinem Netz Daten und übergibt diese an definierten Schnittstellen (meist «Peerings» bezeichnet) an andere Provider. Der Begriff Netz ist hier etwas mehrdeutig. Das Internet Protocol wurde aber ausdrücklich entworfen zur Kommunikation zwischen Netzen. Dieser technische Sprachgebrauch zieht sich daher überall durch die Diskussion. Ein «Netz» im Sinne von IP (besser des Classless Interdomain Routing) ist ein «Block» von Adressen. IP-Adressen sind bekanntlich schlichte Ganzzahlen (integer numbers) von 32 Bit Länge. Ein Block ist dann ein Ausschnitt aus dieser Menge der keine Lücken aufweist: eine fortlaufende Reihe von Nummern. Dieser Block muss noch einer weiteren Anforderung genügen: Beginn und Ende müssen bestimmten Zweierpotenzen entsprechen, aber das ist eine technische Feinheit.

IP Adressen sind eindeutig – so sind sie definiert. Die interessante Ausnahme stellen die in RFC1918 ausgewiesenen Adressbereiche dar, die hier definierten Adressen werden mehrdeutig verwendet. Um die Eindeutigkeit der Adressen zu gewährleisten wurden drei Organisationen geschaffen, die über die Adressvergabe wachen. Für Europa, Afrika und den Nahen Osten ist RIPE zuständig, Arin heisst das Gegenstück für die Amerikas. IP-Adressen sind ein knappes Gut, da ihre Zahl begrenzt ist. von den 2^32 Adressen sind diverse Blöcke für spezielle Aufgaben vergeben sodass nur ein Teil – in etwa die Hälfte – wirklich zur Adressierung zur Verfügung steht. Diese Adressen wurden in der Zeit, da IP noch ein Forschungsgegenstand/ Spielzeug von Akademikern war sehr grosszügig vergeben und sind heute praktisch nicht nutzbar. Die Knappheit ist daher real und auch schon spürbar: der grösste Teil des Adressraumes, den RIPE zur Verfügung hat, ist bereits an Provider delegiert. Bei der aktuellen Rate, mit der Adressen vergeben werden, ist ein Ende schon in wenigen Jahren absehbar; wenn bis dahin keine praktikablen Methoden zur Erweiterung des Adressraumes gefunden werden, dann können keine neuen Teilnehmer zum Internet hinzugefügt werden. Diese Knappheit steht ganz im Gegensatz zum Domain Name System wie im Abschnitt DNS beschrieben. Das DNS ist praktisch beliebig um neue Domains erweiterbar.

8.4.1 regionale Verteilung

Die regionale Verteilung wurde recht einfach vorgenommen. Von den möglichen Adressen wurden 50% an Arin vergeben und je 25% an RIPE und das für Asien zuständige Gegenstück. Diese Verteilung zeigt offenbar die damaligen Abschätzungen davon wie sich die Adressnutzung in der regionalen Verteilung darstellen würde. Sie korreliert offenbar nicht mit demoskopischen Daten: das Internet war zu jener Zeit nicht als Massenphänomen und Endkundenprodukt erkennbar.

8.4.1.1 Arin

Wegen der recht grossen Zuteilung an Arin schlummern dort noch recht grosse Reserven. Dies führt dazu, dass mehr und mehr «amerikanische» Adressen auch in Europa verwendet werden. Dieser Effekt verstärkt sich dadurch, dass die meisten global agierenden Provider aus den USA stammen (und ihre Adressen dort erhalten) und diese dann frei in ihrem ganzen Netz einsetzen.

8.4.1.2 Ripe

RIPE hat als erste begonnen, ein strenges Adressregime einzuführen. Zunächst einmal muss ein Provider seine Adressen bei RIPE beziehen – sonst kann er mit keinem anderen Provider Pakete austauschen – und ist aus dem Geschäft bevor es begonnen hat. Es ist interessant festzustellen dass diese Organisation komplett ausserhalb staatlicher Kontrolle agiert und vor allem auf der Basis von Konventionen der Provider funktioniert. Das erinnert wieder ein Wenig an Seerecht. RIPE vergibt Adressen nach Bedarf an die Provider: diese erhalten einen Block (in der Regel erst einmal einen kleineren) auf Antrag und wenn der Block «verbraucht» ist, einen weiteren. Verbraucht werden Adressen dadurch, dass sie einem Kunden des Providers überlassen werden. Der Provider ist gehalten, über die Verwendung der Adressen Buch zu führen und die Verwendung in einer Datenbank bei RIPE zu verzeichnen. Er wird dazu nicht gezwungen: den Block den er hat kann er frei verwenden und auch verschwenden. Beantragt ein Provider einen weiteren Block, dann prüft RIPE ob die schon im Besitz des Providers befindlichen Blöcke sinnvoll und sparsam genutzt sind; die Quelle für diese Daten ist wiederum die RIPE-Datenbank. Sind die Blöcke nicht verbraucht (oder der Verbrauch nicht korrekt dokumentiert, das ist dasselbe), dann gibt es keine neuen Adressen. Dieses Verfahren führt bei den Providern in Europa in der Regel zu einem sehr strengen Adressregiment – schon aus kommerziellen Interessen (auch morgen noch IP-Adressen zu haben). Die RIPE-Datenbank stellt offenbar eine gute Datenquelle auch für unseren Zweck dar: habe

ich eine Adresse, dann kann ich bei RIPE erfahren, wer der Inhaber dieser Adresse ist.

8.4.2 Funktionsweise

IP-Netze transportieren alle Daten in Form von Paketen variabler Länge. Eine Mindestlänge existiert nicht, die sinnvolle Untergrenze ist jedoch ca. 60 Bytes Nutzdaten, da in jedem Paket auch Adress- und Verwaltungsinformation übertragen wird, die über 64 Bytes hinausgeht. Typische Paketgrössen sind ca. 800 bis 1500 Bytes.

8.4.2.1 Routing

Kommunikation findet immer zwischen Quelle und Senke statt; die Wegfindung für die Daten von der Quelle zur Senke ist daher ein grundlegendes Problem für alle Datennetze. In vielen Netzen werden dazu Wege simuliert, die denen des Telefonnetzes entsprechen, bei dem ja ein direkter Weg für die Dauer der Verbindung hergestellt wird.

IP verwendet einen interessanten neuen Ansatz. Eine Datenquelle sendet ein Paket aus und dieses gelangt zum ersten Router. Der Router trifft nun die Entscheidung, über welches seiner Interfaces er es weiterleitet. Diese Entscheidung ist lokal und betrifft nur den nächsten Abschnitt des Weges. Das Paket kommt also Schritt für Schritt dem Ziel näher – oder es läuft im Kreis. Es gibt keine Möglichkeit für einen Router, eine Entscheidung über mehr als einen Wegabschnitt, einen «Hop» bis zum nächsten Router zu treffen. Auch «routing loops» also Kreise sind nicht erkennbar. Der Ansatz ist deshalb so interessant, weil er eine Art der Imperfektion anderer Netzprotokolle – nämlich unter Umständen nicht die optimale Route zu wählen – zum Prinzip erklärt. Pakete können sich gegenseitig überholen, können auf suboptimalen Routen transportiert werden – alles ist erlaubt. Die IP Software in jedem Router versucht nur, das Paket näher an das Ziel zu bekommen. IP ist ein «best effort» Protokoll. Die Entscheidung, welchen Weg das Paket nehmen soll, hängt im Prinzip nur von der Zieladresse ab. Alle andere Para-

meter des Paketes (Grösse, Quelladresse etc.) können eventuell berücksichtigt werden, in der Regel werden sie aber ignoriert. Nur ein Parameter wird in der Regel noch verwendet: ein Paket wird nicht über das Interface gesendet, über das es empfangen wurde: an dieser Stelle befindet sich ein Router, der das Paket schon einmal hierher gesendet hat, er wird es höchstwahrscheinlich wieder tun.

8.4.2.2 Mythen

Einer der verbreitesten Mythen ist, dass ein IP Netz «unzerstörbar» ist. Es war in der Tat das Ziel, ein Protokoll zu schaffen, das in der Lage ist, ein Paket immer noch zu befördern, selbst wenn Teile des Netzes ausgefallen sind. Da die Routing-Entscheidungen lokal getroffen werden und nur bis zum nächsten Nachbarn reichen, ist dies möglich: ist der Nachbar nicht mehr erreichbar, dann werden folgende Pakete für dasselbe Ziel auf anderen Wegen gesendet. Solange noch ein Weg zum Ziel besteht wird das Paket ankommen. Das kommerzielle Internet hat mit dieser Unzerstörbarkeit jedoch weitgehend aufgeräumt. Da hier im Gegensatz zu einem Forschungsnetzwerk oder einem militärischen Kommunikationssystem für jedes Paket die Kosten eine grosse Rolle spielen ist es erforderlich, das Netz weitgehend zu minimieren: für ein Ziel existiert daher in vielen Fällen nur eine einzige mögliche Route. Gibt es auf dieser Route eine Störung in einem Router oder Übertragungssystem, dann ist das Paket unzustellbar. Es gilt als schlechte Netzqualität, keine Redundanz vorzusehen. Es kommt aber immer wieder vor.

8.4.2.3 Kommerzielles Internet

Das kommerzielle Internet hat die verwendeten Protokolle sehr stark verändert. Nicht nur, dass weniger Redundanzen vorgesehen werden, weniger ungenutzte Bandbreite (die dann in Störfällen als Reserveweg genutzt wird) installiert ist, auch die schiere Grösse des System hat gegenüber dem ursprünglichen Gedanken einiges verändert. Der vielleicht wichtigste neue Gedanke ist beinahe banal: die Nutzung des Netzes muss bezahlt werden. Dies war im ursprünglichen Entwurf der TCP/IP Protokolle nicht vorgesehen: nur die

Layer 3 und 4 des OSI Schichtenmodells werden implementiert, der session layer – der im Schichtenmodell die Funktionalität für die Erfassung von Nutzungsdaten enthält – bleibt undefiniert. Er ist in der Regel dann auch «leer implementiert», realisiert also keine Funktion.

8.4.3 Adressierung

Die regionale Verteilung der Adressen haben wir bereits angesprochen, hier wollen wir noch einmal grundlegend ansehen, wie die hosts des Internet adressiert werden.

8.4.3.1 Notation

Die am weitesten verbreitete Notation von IP-Adresse ist die Form des «dotted quad» in der Art 192.168.10.42. Die IP-Adresse ist an sich eine Dualzahl mit 32 Bits. In der dotted quad notation wird diese Zahl nun zunächst in ihre vier Bytes unterteilt und dann für jedes Byte der Zahlwert angegeben. Die Angabe der Werte erfolgt in der Regel dezimal wie oben dargestellt, seltener sedezimal (also 0A.EF. 42.B5). Wie unter CDR dargestellt teilen sich die 32 Bits der Zahl in einen Netz- und einen Host-Teil auf; die Aufteilung wird in der Regel als Netzmaske dargestellt. Die Netzmaske ist wiederum eine Dualzahl mit 32 Bits; sie enthält an all jenen Stellen, an denen die IP-Adresse eine Netzadresse enthält, eine 1, an den Stellen wo die IP-Adresse die Hostadresse enthält eine 0. Eine typische Netzmaske beginnt daher immer mit 255, da die ersten 8 Bit der IP-Adresse fast immer Netzadresse sind, die Maske also erst einmal 8 Eins-Bits enthalten muss; eine typische Netzmaske ist dann 255.255.128.0 (17 Bit Netz, 15 Bit host) oder 255.255.255.240 (29 Bit Netz, 3 Bit Host). Die Netzmaske dient also nur der Beschreibung, wie die IP-Adresse logisch aufgeteilt ist, sie hat keine Funktion zur Adressierung.

8.4.3.2 Permanente Adressierung

Der ursprüngliche Adresstyp ist die permanente Adresse: ein host – genauer ein Interface eines host – erhält eine Adresse zugewiesen und behält diese permanent. In der Regel werden die Adressen

blockweise vergeben: ein LAN erhält beispielsweise einen Adressvorrat, der ausreicht alle potentiellen hosts zu versorgen. Ein solcher Adressvorrat wird als ein «Netz» bezeichnet. Ein Netz stellt das grundlegende Strukturelement des Inter-Net-Protocol dar. Ein Router stellt eine Verbindung zwischen Netzen her. In jedem Netz wird eine Adresse verwendet um das Netz selbst (als Einheit) zu adressieren; in der Regel hat diese Adresse keine Bedeutung. Eine Adresse dient als Broadcast-Adresse, also um alle hosts in dem Netz zu adressieren. Diese kann nicht für einen host verwendet werden. Die Zuteilung von Adressen an einen Provider und – im Einflussbereich von RIPE – die Zuordnung zu einem Providerkunden ist in öffentlich zugänglichen Datenbanken dokumentiert.

8.4.3.2.1 Class based addressing
Zunächst wurden Netze in Klassen eingeteilt, die 2^{24} (16777216, Class A), 2^{16} (65536, Class B) oder 2^{8} (256, Class C) Adressen umfassen. Da dieses Verfahren zur Verschwendung von Adressen beiträgt – so muss ein LAN mit 1000 hosts schon zweckmässig ein «Class-B-Netz» erhalten und verwendet so 65536 Adressen – wurde Mitte der 90er Jahre das classless interdomain routing spezifiziert und eingeführt. Die Netze werden nun sehr viel granularer vergeben.

8.4.3.2.2 Classless interdomain routing (CDR)
Durch die Aufgabe der «Klassen» und die Betrachtung aller Adressen als einer einheitlichen Menge ergeben sich Vorteile für die Adressierung: die Grösse des verwendeten Netzes kann der Aufgabe angepasst werden. Ein Netz enthält immer eine Zweierpotenz an Adressen, das kleinste Netz ist also jenes, das 2^{2} (4) Adressen umfasst. Das nächstgrössere ist das Netz, das 2^{3}, also 8 Adressen umfasst und so fort. Die Schreibweise 192.168.2.0/23 (oder auch 192.168.2/23, Nullen am Ende werden weggelassen) deutet dabei an, dass die ersten 23 Bit der Adresse für das Netz stehen, die verbleibenden 9 Bit für den Host innerhalb dieses Netzes. Dies lässt sich auch als Netzmaske mit 23 Bit darstellen (255.255.250.0); da 9 Bits für die Hostnummern bereitstehen hat so ein Netz Platz für 512 Hostadressen.

8.4.3.2.3 Reservierte Adressen
Einige Adressen sind reserviert für besondere Zwecke. Die Adresse 127.0.0.1 steht für «dieser host». Sie wird nicht geroutet und steht nur für die Kommunikation innerhalb eines Computers zur Verfügung. Das mag auf den ersten Blick unsinnig erscheinen, ist aber genau die Methode mit der z.B. ein Webbrowser mit einem Webserver kommunizieren kann, der auf derselben Maschine läuft – selbst im Desktopbereich sind «persönliche» Webserver keine Seltenheit mehr, seit bei MacOS und später Windows diese Funktion Bestandteil des Betriebssystems geworden ist. Die Adresse 0 jedes Netzes – also jene, bei der die Bits des Host-Teiles der IP-Adresse alle 0 sind – steht für das Netz selbst. Sie hat in der Regel keine besondere Verwendung ausser bei bestimmten Spezialanwendungen (so können z.B. bestimmte Ethernet-Switches, die ja das «Netz» bilden, über Pakete an diese Adresse gesteuert werden). Die letzte Adresse – jene bei der der Host-Teil der IP-Adresse nur 1-Bits enthält – ist die broadcast-Adresse: Sendungen an diese Adresse werden von allen hosts innerhalb dieses Netzes empfangen und ausgewertet. Für ein /24-Netz ist die Broadcast-Adresse die 255, für ein /23-Netz die 511 und so fort.

8.4.3.3 Temporäre Adressen
Eine nachträglich eingeführte Idee ist die der temporären Adresse. Diese macht das kommerzielle Internet erst möglich. Die wichtigste Anwendung ist der Internet-Zugang über das Telefonnetz. Hier ist die Adresse dem Modem beim Provider zugeordnet und der Kunde erhält die Adresse, die seinem Modem entspricht. Diese Zuordnung ist also völlig zufällig. Die Zuordnung einer IP-Adresse zu einem Kunden ist eventuell rekonstruierbar, da Provider Protokolle über die Einwahl der Kunden führen (für Abrechnungszwecke). Diese Protokolle erlauben oft, nachträglich die Zuordnung (Zeit, IP, Kunde) herzustellen; solche Protokolle (in der Regel logs genannt, schon um die Verwechslung mit «protocol» als Ablauffestlegung zu vermeiden) sind datenschutzrechtlich bedenklich, für die Fehlerdiagnose im Netz aber oft unvermeidlich. Sie stellen eine interessante Daten-

quelle dar, will man jemanden ermitteln, dessen IP-Adresse man kennt.

8.4.3.4 Neue Adressen: IPv6

IP (die aktuell eingesetzte Version des Protokolls ist v4) bietet nicht sehr viele Features. Dies ist ganz im Sinne seiner Väter: nur einem einfachen Protokoll traute man die notwendige Robustheit in der Anwendung zu. Das kommerzielle Internet verlangt aber nach Eigenschaften, die IPv4 nicht bieten kann. Daher wird schon seit einiger Zeit an einer Fortentwicklung gearbeitet. Eine dieser Entwicklungen, IPv6, ist nunmehr weitgehend standardisiert und stellt interessante neue Eigenschaften bereit. IPv6 wird noch nicht auf breiter Front angewendet, die Technik ist noch zu jung, der Umstellungsaufwand selbst für kleine Provider sehr gross. Auch ist bisher nicht abschliessend geklärt, wie IPv6 Routing-Informationen zwischen verschiedenen Providern ausgetauscht werden können – das ist aber eine Voraussetzung für die Funktion der Peerings zwischen Providern. Daher wird IPv6 derzeit mehr in kleinen und mittelgrossen Netzen eingesetzt, in denen die neuen Features wie Bandbreitengarantien echte Vorteile bringen. IPv6 ist keine voll kompatible Fortentwicklung von IPv4 sondern in weiten Teilen ein grundlegend neuer Ansatz. Die Adressen sind in einer Richtung kompatibel: jede IPv4 Adresse ist auch eine IPv6 Adresse, der v4 Adressraum ist also ein Ausschnitt aus dem v6 Adressraum. Um IPv6 Pakete über Netze zu transportieren, die mit IPv4 arbeiten, werden IP-Tunnel verwendet.

8.4.4 Tunneling

Eine Technik, die viel zur Verwirrung des Adressraumes beiträgt, ist das Tunneling. Verwirrung insofern als die Tunnel einen Host an einen Ort erscheinen lassen, der gar nicht sein Ort in der Netztopologie ist. Da zusätzlich die Netztopologie geographische Gegebenheiten (und politische Grenzen) weitgehend ignoriert stellt diese Technik ein zusätzliches Problem dar, wenn man einen IP-Benutzer lokalisieren will. Ein Tunnel funktioniert eigentlich ganz einfach. Er besteht aus zwei Geräten, die die Tunnelendpunkte bilde. Der Tunnel-

eingang TE nimmt ein IP-Paket entgegen das an ein Ziel hinter dem Ausgang adressiert ist, sagen wir an Adresse Z. Aus Sicht des Host, der das Paket sendet, ist der Tunneleingang also ein Router wie jeder andere. Der Tunneleingang leitet das Paket jedoch nicht unmodifiziert weiter (wie ein Router es täte) sondern packt es wiederum in ein Paket (in der Regel wieder ein IP Paket, dann spricht man von einem IPIP Tunnel) ein, das er an den Tunnelausgang TA adressiert. Diese neue Paket schickt er dann wie gewöhnlich auf die Reise. Aus Sicht des Netzes (also aller Router auf dem Weg) ist die Zieladresse des Paketes der Tunnelausgang. Das ursprüngliche Paket liegt dabei dann in der Nutzlast des umgebenden Paketes. Der Tunnelausgang muss nur das Paket auspacken und das so rekonstruierte Ursprungspaket wieder auf die Reise schicken. Besonderen Nutzen entfaltet diese Technik um Pakete andere Protokolle (z.B. das früher sehr beliebte IPX oder aber AppleTalk) über IP Netze zu transportieren. Doch auch für Sicherheitszwecke werden Tunnel sehr gern verwendet. Der Tunneleingang verschlüsselt dabei in der Regel das Paket bevor er es an den Ausgang absendet. Solche secure VPN Applikationen sind keine Spezialfälle mehr; die entsprechende Software gehört seit Windows 2000 und Windows ME zur Standardausstattung jedes PC dieser Baureihen. Auch das beliebte Paket PGP desktop security enthält eine Komponente zur Verschlüsselung des Netzverkehrs.

8.5 DNS

Das Domain Name System (DNS) stellt eine kritische Komponente des kommerziellen Internet dar: fällt es aus, dann wird dies in der Regel von den Benutzern als Ausfall des Internet an sich empfunden. Tatsächlich ist das DNS ein eigenständiges System, das dem Internet als «Hilfssystem» dient – ähnlich wie die Mathematik der Physik als Hilfsmittel für die Darstellung von Zusammenhängen dient. Das DNS ist nicht Bestandteil des ursprünglichen Entwurfes, es ist eine spätere Überlegung. Das IP-Transportnetz funktioniert

daher auch völlig ohne ein DNS. Das DNS stellt das Objekt «Domain» bereit. Die Natur dieses (gedanklichen, es ist Software) Gebildes ist oft Gegenstand der Diskussion gerade unter Juristen: keine der bekannten Kategorien «Recht», «Zeichen», «Name», «Adresse» scheint für eine Domain zu passen.

8.5.1 Domain ist Verzeichnis

Aus Sicht des Technikers ist eine Domain nicht mehr als ein Verzeichnis Dieses stellt sich in der Regel dargestellt als eine einfache Datei dar, die auf einem Datenträger gespeichert ist. Für den praktischen Gebrauch werden diese Dateien dann in Datenbanken überführt, die eine schnelle Suche zulassen, im Prinzip bleibt es aber die einfache Datei. Um das Verzeichnis eindeutig ansprechen zu können erhält es einen Namen. Der Name ist dabei einfach ein Symbol für die Adresse (IP-Adresse) des Hosts, der das Verzeichnis enthält und der Abfragen des Verzeichnisinhalt ermöglicht: des name server. Der Name der Domain ist eindeutig: jede Domain kann nur einmal existieren innerhalb ihres Kontext. Der Kontext bildet sich aus einer Hierarchie der Domains; genau das war der Durchbruch dieses Systems gegenüber dem zuvor im Usenet genutzten Namenssystems. Im Usenet gab es einfach nur Hosts. Jeder Host hatte einen Namen und jeder Name musste eindeutig sein: er war gleichzeitig Adresse des Host. Das DNS beseitigt diesen Misstand. Es errichtet erst einmal in der «Welt» die Top-Level-Domains. Im ersten Schritt ist dabei «Welt» die Gesamtheit aller Hosts des Internet zu verstehen. Da diese «Welt» zunächst auf die USA beschränkt waren wurden die TLDs für diese definiert. Mehrere TLDs wurden errichtet, jede mit einer dedizierten Funktion. Geographische oder andere physische Gesichtspunkte spielen bei der Definition von .com, .gov, .edu, .org, .net und .mil keine Rolle. Im zweiten Schritt wurde das DNS dann auf die (geographische) Welt ausgedehnt. Dabei wurden TLDs errichtet, die politischen/geographischen Gegebenheiten folgen: für jeden Staat wurde eine TLD errichtet, die als Namen das ISO-Landeskürzel dieses Staates trägt: .de für Deutschland, .ch für die confederatio helvetica, .ag für Antigua und .tm für Turkmenistan sind hier

die Beispiele. In den Staaten konstituierten sich nun NICs, die die Verwaltung der Domains übernahmen; diese erliessen Regeln für die Domainvergabe, die durchaus sehr verschieden sind. So wählten Grossbritanien und Taiwan zum Beispiel eine funktionale Unterteilung der Domain .gb dem Muster der USA folgend in .co.gb, .co.tw, .gov.gb und so fort. Andere NICs wählten explizit den Ausschluss dieser Kürzel: in Deutschland sind edu.de, gov.de, com.de zum Beispiel ausdrücklich von der Domainvergabe ausgeschlossen.

8.5.2 Frei wählbar, beliebig

Innerhalb der Regeln, die das NIC erlassen hat, ist der Name einer Domain frei wählbar. Jede beliebige Zeichenkette kann verwendet werden. Dies schliesst auch Zeichenketten ein, die Wortmarken oder Personennamen identisch sind oder an diese erinnern. Die NICs sind in der Regel Organisationen, die sich aus der Gesamtheit der Internetprovider in einem Land konstruieren; sie sind also in der Regel nichtstaatlich und konstruieren sich als Verband von Wettbewerbern. Das allein führt oft auf interessante Konstellationen. Die Freiheit der Wahl, die relative Unbekanntheit des Internet noch vor einem Jahrzehnt und die Eindeutigkeit des Domainnamens stellen dabei eine explosive Mischung dar: das «Domain Grabbing» griff um sich: professionelle Domaingrabber registrierten diverse Domains, die in ihrem Namen identisch zu Marken oder Firmennamen waren. Damit wurde den Markeninhabern der Zugang zu diesen Domains dann verwehrt, der Domaingrabber versucht dann «seine» Domain an einen anderen, der Interesse daran äussert, zu «verkaufen». Dieses Verhalten ist später geächtet worden; legal ist es jedoch oft, da es einen Straftatbestand «Domaingrabbing» nicht gibt und das Analogieverbot und der Bestimmtheitsrundsatz des Strafrechts einer Verurteilung aufgrund von Ähnlichkeiten zu strafbewehrtem Verhalten verbieten.