in: Hans-Werner Moritz, Thomas Dreier (Hg.): Rechts-Handbuch zum E-Commerce, Verlag Dr. Otto Schmidt KG, Köln 2002, 1-17.

A. Einführung in die Grundlagen des elektronischen Geschäftsverkehrs

1. Technische Grundlagen des Internet

1.1 Aufbau des Internet

Das weltweite Rechnernetz „Internet“ wird gebildet durch den Zusammenschluß vieler Millionen Rechner und vieler Tausend Rechnernetze. Zu Beginn waren dies größtenteils Rechner und Rechnernetze wissenschaftlicher und technischer Einrichtungen. Mehr und mehr kommen jedoch Rechner aus allen Bereichen des Lebens (Wirtschaft, öffentliche Einrichtungen, private Haushalte) hinzu. Die verteilte Netzstruktur des Internet besteht aus Rechnern vieler verschiedener Hersteller mit sehr unterschiedlicher Hardware- und Softwareausstattung. Damit die daraus resultierende Vielfalt kein Hindernis bei der weltweiten Kommunikation ist, wurden technische und organisatorische Kommunikationsvereinbarungen getroffen, an die sich alle Rechner des Internet halten müssen.

Die Vielfalt an Benutzern und Betreibern hat weiterhin die Konsequenz, daß man nicht davon ausgehen kann, daß sich alle Akteure im Internet kooperativ verhalten. Es existiert zwar eine sog. Netiquette, aber niemand ist gezwungen, sich daran zu halten. Nicht kooperatives Verhalten wird durch das Internet größtenteils noch nicht verhindert. Anders herum gesagt: Es existieren derzeit nur sehr wenige Sicherheitsfunktionen, die Betreiber und Benutzer vor Angriffen auf die Verfügbarkeit, Integrität, Zurechenbarkeit und Vertraulichkeit von Diensten und Daten schützen. Dieses Defizit muß für die ernsthaft geschäftsmäßige Anwendung des Internet, also für E-Business, beseitigt werden, sonst leidet auf lange Sicht die Vertrauenswürdigkeit eines „im Netz“ agierenden Unternehmens.

Technisch gesehen ist das Internet ein stark vermaschter Graph von einzelnen kleineren Rechnernetzen. Das Internet ist international und organisationsübergreifend, nicht nur bezüglich der abrufbaren Daten, sondern ebenfalls bezüglich der Kommunikationsprotokolle, die weltweit standardisiert sind (siehe I.2.2). Im Internet sind eine Vielzahl von Diensten realisierbar; die wichtigsten (z.B. E-Mail, World Wide Web, Telnet) sind ebenfalls standardisiert.

Transport in Datenpaketen

Das zentrale Protokoll für den Transport von Daten im Internet ist das **Internet Protocol (IP)**, das dem Netz auch seinen Namen gab. Alle transportierten Daten – egal ob Text, Sprache, Video, Bilder oder binäre Dateien – werden bei der Übertragung in einzelne Datenpakete zergliedert. Diese sog. IP-Pakete besitzen einen Nachrichtenkopf (Header) und den eigentlichen Nachrichteninhalt (Payload), siehe Abbildung 1.

Abbildung 1. Alle Inhalte werden in Pakete verpackt

Die wichtigsten Datenfelder des Headers sind:

- Versionsnummer des IP-Protokolls (meist 4, zukünftig 6)
- Größe des gesamten IP-Paketes,

- Protokoll-Feld,
- IP-Adresse des Absenders,
- IP-Adresse des Empfängers.

Das Protokoll-Feld legt fest, wie die im Payload enthaltenen Daten interpretiert und transportiert werden. Die beiden wichtigsten Transportprotokolle sind das **Transmission Control Protocol (TCP)** und das **User Datagram Protocol (UDP)**. Nahezu alle derzeit wichtigen Dienste benutzen entweder TCP oder UDP. Deren Eigenschaften werden in Abschnitt I.1.2 genannt.

IP-Adressen

Jeder an das Internet angeschlossene Rechner besitzt eine IP-Adresse. Eine IP-Adresse besteht aus 4 Zahlen (Bytes) zwischen 0 und 255. Die IP-Adresse (z.B. 141.76.75.101) dient der „Wegewahl“ (Routing) vom Absender zum Empfänger. Der Empfänger eines IP-Pakets antwortet dem Absender, indem er die Anwort an die IP-Nummer im Absenderfeld sendet und seine IP-Adresse als Absender einträgt.

Es gibt im wesentlichen **Client-, Server- und Vermittlungsrechner**. Der Client (z.B. der Browser auf dem Rechner) nutzt einen vom Server angebotenen Dienst, während die Vermittlungsrechner an den Knotenpunkten der Kommunikationswege anhand der IP-Adressen lediglich Datenpakete durch das Netz transportieren, ohne deren eigentlichen Inhalt näher zu untersuchen (siehe Abbildung 2).

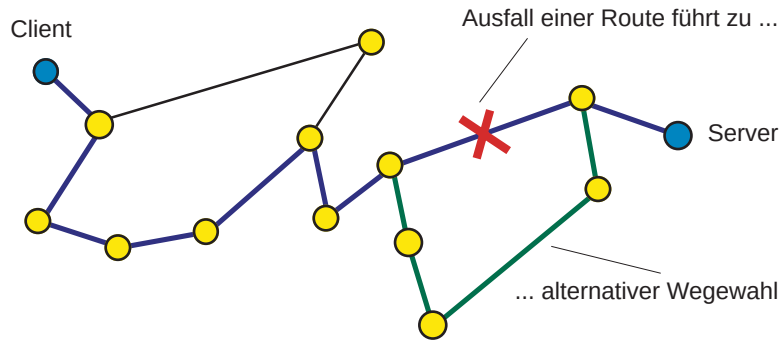

Abbildung 2. Ausfall einer Route führt zu alternativer Wegewahl

Verteiltes System

Das Internet ist ein verteiltes System. Das bedeutet, daß die Ressourcen (z.B. Speicher, Rechenleistung) nicht zentral angeordnet sind, sondern über das Netz verteilt sein können. Beispielsweise kann der Aufruf einer Webseite dazu führen, daß viele, weltweit verteilte Kommunikationsverbindungen vom Client (hier also dem Browserprogramm auf dem PC) angefordert werden. Am Ende entsteht die Webseite als Ganzes, obwohl die auf ihr dargestellten Objekte (z.B. Bilder oder Texte) im Internet verteilt gespeichert waren.

Ursprünglich wurde das Internet für **Hochverfügbarkeit** entwickelt. Das bedeutet, daß ein unterbrochener Kommunikationsweg nicht zwangsläufig zur Unterbrechung der Kommunikation führt, da die Vermittlungsrechner in der Lage sind, alternative Kommunikationswege zu finden (siehe Abbildung 2).

Ein Rechner kann gleichzeitig Client- und Serverfunktionen übernehmen und auch Kommunikationspakete vermitteln. Man muß also jeweils für jeden Dienst unterscheiden, welche Rolle ein Rechner im Kommunikationsgeschehen einnimmt.

Zugang zum Internet

Zur Kopplung eines einzelnen Rechners oder eines organisationsinternen **Local Area Network (LAN)** an das Internet bieten **Internet Service Provider (ISP)** entsprechende Zugänge an. Die

Kopplung erfolgt meist über ein Modem, das die physikalische Anpassung an die jeweilige Übertragungsstrecke vornimmt. Als Internet-Zugänge werden vorrangig eingesetzt:

- das analoge Telefonnetz mit einer maximalen Übertragungsrate von 55 kbit/s,
- Integrated Services Digital Network (ISDN) mit maximal 128 kbit/s,
- Digital Subscriber Line (DSL) mit etwa 0,7-2 Mbit/s, meist als Asymmetric DSL, wobei die Raten für eingehende und ausgehende Nachrichten unterschiedlich sind.

Das gebräuchlichste Übertragungsprotokoll, mit dem die IP-Pakete über die jeweilige physische Übertragungsstrecke transportiert werden, ist das **Point-to-Point Protokoll (PPP)**. Der Internet Service Provider betreibt einen PPP-Server, über den sich die Kunden mit dem Internet verbinden. Der auf der Kundenseite nötige PPP-Client ist heutzutage Bestandteil der Standardbetriebssysteme. Damit Internet-Zugänge nicht durch Unberechtigte mißbraucht werden können, muß sich der Client (bzw. sein Benutzer) gegenüber dem ISP authentisieren. Gebräuchlich sind hier Passwörter, sog. Callback-Verfahren[1], aber auch proprietäre, d.h. nicht standardisierte Verfahren.

Um mehrere Rechner eines LANs an das Internet zu koppeln, setzt der Betreiber des LANs einen Router ein, der für alle Rechner des LANs den Zugang zum Internet bereitstellt. Für die organisationsinterne Kommunikation wird meist ebenfalls IP verwendet. Man spricht dann von einem **Intranet**. Somit ist potentiell jeder Rechner des LAN auch aus dem Internet erreichbar, da er über eine IP-Adresse verfügt. Ohne zusätzliche Schutzmechanismen birgt dies erhebliche Risiken für die Datensicherheit und den Schutz firmeninterner und personenbezogener Daten, da es meist sehr aufwendig und teilweise auch unmöglich ist, jeden einzelnen Rechner des Intranet gegen unbefugten Zugriff von außen zu schützen. Deshalb wird das Intranet meist mit einer **Firewall** gegen unbefugten Zugriff von außen gesichert.

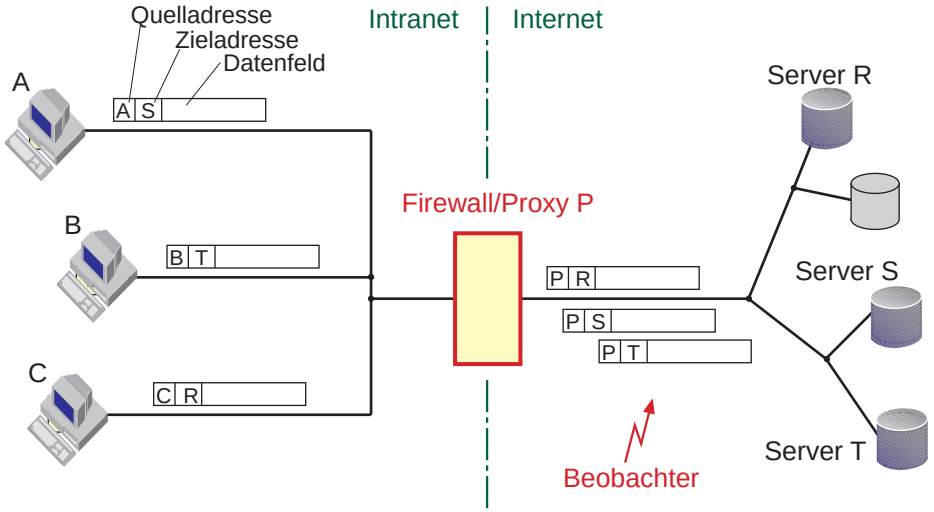

Abbildung 3. Proxies als Schutz vor Beobachtern im Internet

Meist kommt die Firewall zusammen mit einem **Proxy** (Stellvertreter) zum Einsatz. Der Proxy verbirgt bei Zugriffen, die aus dem Intranet stammen und in das Internet gehen, welcher Rechner den Request abgesendet hat. Hierzu bauen die Rechner des Intranet zunächst eine Verbindung zu einem Proxy-Server auf, der seinerseits (stellvertretend) die vom Benutzer gewünschte Verbindung zum Zielrechner (z.B. einem WWW-Server) aufbaut (siehe Abbildung 3). Beobachtern im Internet bleiben damit die (firmeninternen) Netzstrukturen und Adressen verborgen. Dies erschwert Hackern das Eindringen in das Intranet. Darüber hinaus verhindern Proxies die Beobachtung einzelner Nutzer, da die Ursprungsadresse einer Verbindung vor dem Internet verborgen wird.

[1] Das Modem des Servers ruft den Client zurück, aber nur dann, wenn der Zugriffswunsch von einer dem Server bekannten und berechtigten ISDN-Rufnummer stammt. Das Callback-Verfahren kommt z.B. bei firmeneigenen Modem-Zugängen zum Einsatz.

Proxies und Firewalls schützen nicht vor einem Beobachter oder Saboteur, der im Intranet verbreitet ist, d.h. berechtigten Zugang zum Intranet hat (z.B. ein Mitarbeiter). Ebenso schützen sie nicht, wenn der Proxy/Firewall selber der Beobachter/Saboteur ist.

Ein Verfahren, das selbst gegen Beobachtung durch den Betreiber des Proxys sicher ist, ist das Mix-Netz[2]. Ein praktisch nutzbares System zum anonymen Web-Surfen ist der Java Anon Proxy[3].

1.2 Die Internet-Protokoll-Familie

Bevor im Abschnitt I.2. näher auf die wichigsten Anwendungen im Internet eingegangen wird, sollen einige Grundbegriffe der Übertragungsprotokolle im Internet eingeführt werden.

Die Nutzerdaten werden im Internet mit Hilfe von zwei **Übertragungsprotokollen** transportiert, dem Transmission Control Protocol (TCP) und dem User Datagram Protocol (UDP). Jeder Dienst eines Servers wird auf einer sog. Portnummer angeboten. Die Zuordnung von IP-Adressen zu den üblichen Rechnernamen erfolgt durch das Domain Name System (DNS).

Die angebotene Sicherheitfunktionalität der IP-Familie ist derzeit noch sehr gering, was sich aber mit künftigen IP-Versionen ändern wird.

Transmission Control Protocol (TCP)

Das Transmission Control Protocol (TCP) wird bei **Punkt-zu-Punkt-Verbindungen** zwischen zwei Endpunkten, z.B. einem Browser und einem Webserver eingesetzt. Bei TCP wird darauf geachtet, daß alle vom einen Endpunkt gesendeten Bits auch tatsächlich beim anderen Endpunkt ankommen und auch deren Reihenfolge nicht durcheinander kommt. Falls Daten beim Transport verloren gehen, werden sie erneut gesendet (*Retransmission*). Dieses Transportprotokoll wird z.B. beim Transport von Webseiten, E-Mails, Dateien etc. angewendet, da man sicher gehen möchte, daß die Daten auch wirklich beim Empfänger ankommen.

Sollen mit Hilfe von TCP-Verbindungen viele Nutzer mit dem gleichen Inhalt von einem Server versorgt werden, muß jeder Nutzer eine eigene Verbindung zum Server aufbauen. Der Bandbreitebedarf wächst dadurch linear mit der Teilnehmerzahl, da der Server jeweils eine Verbindung pro Client und Request unterhält.

User Datagram Protocol (UDP)

Beim User Datagram Protocol (UDP) sendet der Sender Datenpakete aus, die in Abhängigkeit von der Auslastung des Netzes den Empfänger rechtzeitig, zu spät (*delayed*) oder auch gar nicht (*dropped*) erreichen. UDP wird hauptsächlich für Datenströme verwendet, bei denen eine Retransmission nicht möglich ist. Beispielsweise bei Audio- und Videoströmen, die synchron gesendet und konsumiert werden, ist es nicht sinnvoll, verlorengegangene Datenpakete erneut zu senden, da der fehlende „Abschnitt“ des Datenstroms zeitlich hinter dem aktuell gesendeten liegt. UDP-Pakete werden beispielsweise vom Real Player[4] verarbeitet. Der Verlust von Datenpaketen macht sich je nach Kodierung der Medienströme durch Qualitätsverschlechterung oder Aussetzer bemerkbar.

Ein typischer UDP-Dienst ist der Domain Name Service (DNS, siehe weiter unten).

Mit UDP lassen sich sowohl Punkt-zu-Punkt-Übertragungen, d.h. vom Sender zu einem einzigen bestimmten Empfänger, als auch Punkt-zu-Mehrpunkt-Übertragungen (Multicast, Broadcast) realisieren. **Multicast-Verkehr** wird zukünftig den Bereich des Webcasting abdecken. Dabei verbindet sich ein

---

[2] David Chaum: Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. Communications of the ACM 24/2 (1981) 84-88.

[3] Der Java Anon Proxy (JAP, http://anon.inf.tu-dresden.de/) basiert auf dem Mix-Netz und leitet jeden Zugriff über mehrere unabhängige, von unterschiedlichen Betreibern betriebene Mix-Proxies. Solange mindestens ein beteiligter Mix-Proxy korrekt arbeitet, ist der Client unbeobachtbar.

[4] http://www.real.com/

Benutzer z.B. mit einem Videodatenstrom über eine sog. Multicast-Adresse (*join*). Dies wird durch das sog. Internet Group Management Protocol (IGMP) realisiert. An den Vermittlungsstellen wird der Datenstrom erst dann verzweigt, d.h. in mehrere verschiedene Richtungen weitergesendet, wenn es keine günstigere Route zur Versorgung mehrerer Teilnehmer der Gruppe gibt. Dieses Vorgehen spart gegenüber der Punkt-zu-Punkt-Übertragung erheblich Bandbreite und macht die Übertragung von Bewegtbildern in sehr hoher Qualität an viele Empfänger erst möglich.

Derzeit wird massiv an der Zusicherung sog. **Quality-of-Service-Merkmale (QoS)** gearbeitet, um die auftretenden Verzögerungen und Datenverluste derart vorhersagen bzw. vermeiden zu können, daß dem Endbenutzer eine gleichbleibend hohe Qualität der Übertragung zugesichert werden kann. Die bisher entwickelten Protokolle, die alle Sonderformen von UDP sind, tragen Namen wie Real-Time Transport Protocol (RTP), Real-Time Control Protocol (RTCP) und Real-Time Streaming Protocol (RTSP). Im Zusammenhang mit QoS spielt das Resource Reservation Protocol (RSVP) noch eine Rolle. In der gegenwärtigen Distributionspraxis im Internet spielen die genannten Protokolle noch keine große Rolle, was sich aber mit steigenden Übertragungskapazitäten ändern wird.[5]

Portnummern

Jeder Dienst, egal ob er mit TCP oder UDP realisiert wird, ist eindeutig gekennzeichnet durch die IP-Adresse des Servers und die Portnummer, auf der der Dienst angeboten wird. Es können pro IP-Adresse insgesamt jeweils 65535 Dienste für UDP und die gleiche Anzahl für TCP angeboten werden. Einigen Portnummern sind bestimmte Dienste fest zugeordnet. Sie werden „well-known ports“ genannt. Beispielsweise wird das im Word Wide Web wichtigste Protokoll HTTP (Hypertext Transfer Protocol) von einem Webserver meist auf TCP-Port 80 angeboten.

Domain Name System (DNS)

Das Domain Name System (DNS) sorgt für die **Zuordnung von IP-Nummern zu Rechnernamen** und umgekehrt. Damit sich die Benutzer anstelle der IP-Nummer einen Rechnernamen merken können, werden den im DNS registrierten Rechnern ein oder mehrere Namen zugewiesen. Beispielsweise merkt man sich www.datenschutz.de viel besser als 195.244.234.206.

Das Registrieren und Einrichten einer Domain im Internet war in den Anfangszeiten des Internet ein rein technischer Vorgang. Wer als Erster die Einrichtung einer Domain veranlaßte, bekam die Domain. Dies führte zum „**Domain-Grabbing**“ durch gewiefte Geschäftemacher: Es wurden „vorsorglich“ Domain-Namen registriert, die Firmen-, Marken- oder Produktnamen entsprechen. Als die betroffenen Firmen das Internet entdeckten, wurden die Domain-Namen teilweise teuer zurückgekauft.

Die **Namen** sind hierarchisch aufgebaut. Beispielsweise gibt der Rechnername ikt.inf.tu-dresden.de an, daß der Rechner innerhalb der deutschen Domain (de), von der TU Dresden (tu-dresden), Fakultät Informatik (inf) betrieben wird, und zwar am Lehrstuhl Informations- und Kodierungstheorie (ikt). Die Landesbezeichnung in der Domain (z.B. de für Deutschland oder tv für Tuvalu, einem kleinen Inselstaat im Pazifischen Ozean) entspricht nicht notwendigerweise dem geographischen Standort des Rechners, der unter dieser Domain registriert ist.

Bei der Eingabe eines Rechnernamens, etwa in einen Browser, fragt der zunächst das Domain Name System nach der 4 Byte großen IP-Adresse, die nötig ist, um schließlich die Verbindung zum Server herzustellen.

Das DNS ist *kein* Suchsystem. Zwar können Rechnernamen ausprobiert werden, was zu einer erfolglosen DNS-Anfrage führt, aber zum gezielten Suchen von Rechnernamen oder Organisationen sollten Suchmaschinen eingesetzt werden, deren Technologie aber nichts mit dem DNS gemeinsam hat.

Die wirtschaftliche und politische Bedeutung eines Domain-Namens und des DNS wurde von den Entwicklern des Internet lange Zeit unterschätzt. DNS sollte das Merken von Adressen vereinfachen. Das Verhalten und die Erwartungen der Benutzer sind aber anders: Wer www.ibm.com eingibt, möchte gewöhnlich auch auf der Webseite des Rechnerherstellers IBM landen.

---

[5] An technischer Einführungsliteratur kann empfohlen werden: D. Kosiur, IP Multicasting, Wiley, 1998.

Sicherheit von IP

Die Internet Protokolle besitzen nahezu keine Funktionen zur Gewährleistung von Vertraulichkeit und Integrität der übertragenen Daten. Der Schwerpunkt wurde von den Entwicklern auf die Zuverlässigkeit der Übertragung (hier: Fehlertoleranz) gegen unbeabsichtigte Fehler (z.B. Ausfall eines Rechners oder einer Kommunikationsverbindung) gelegt. Das Internet bietet keinerlei Garantie, daß ein IP-Paket tatsächlich und unverändert beim Empfänger ankommt. Ebenso kann man sich nicht darauf verlassen, daß die in einem IP-Paket enthaltene Absenderadresse authentisch ist. Ein Hacker kann ohne große Mühe eine beliebige Absenderadresse in das IP-Paket einfügen. Dieses sog. IP-Spoofing wird von Hackern z.B. bei Denial-of-Service Angriffen (DoS) angewendet, um die Rückverfolgung zum Hacker zu verhindern (siehe auch IV.1.9).

Ein Schutz gegen böswillige Angriffe auf die Vertraulichkeit, Integrität und Verfügbarkeit von Daten und Diensten existierte in den Anfangszeiten des Internet nicht einmal in Ansätzen. Besonders deutlich wird das durch die erfolgreichen Angriffe auf Firmen, die ihre Dienstleistungen nur noch im Internet anbieten, womit erfolgreiche Angriffe existenzbedrohend sind.[6]

IP ohne Berücksichtigung von Security ist keine gute Voraussetzung für den Durchbruch von E-Commerce im Internet. Alle Sicherheitsfunktionen, beispielsweise Secure Socket Layer (SSL) oder Pretty Good Privacy (PGP), müssen derzeit als Add-On, d.h. aufsetzend auf IP, TCP und UDP, eingesetzt werden. Besser wäre es, wenn es einen Grundschütz gäbe. Mit der künftigen Generation des Internet Protokolls, IP Version 6 (siehe folgender Abschnitt), wird die Situation wahrscheinlich besser.

IP Version 6

In Anbetracht der starken Verbreitung und Nutzung des Internet mußte die derzeit meistgenutzte Version der Internet Protocols (IP Version 4) erweitert werden, um zukünftigen Anforderungen gerecht werden zu können. Dabei bot sich die Gelegenheit, die auf Netzebene unterstützten Sicherheitsfunktionen (IPSec) mit IP Version 6 (IPv6) grundlegend mit zu erweitern. IPv6/IPSec[7] besitzt folgende Funktionen, mit denen direkt **kryptographische Funktionen** genutzt werden können:

- Mit einem **Authentication Header (AH)** ist es möglich, den Sender eines Paketes zu authentisieren sowie die Nutzdaten (Payload) vor unerkannter Verfälschung zu schützen.
- In einem **Encapsulating Security Payload (ESP)** ist es möglich, den Inhalt eines Paketes verschlüsselt zu übertragen. Dies bietet einerseits ideale Voraussetzungen für das sog. Tunneling, bei dem beliebige Datenpakete (auch anderer Übertragungsprotokolle) zwischen den Endpunkten des verschlüsselten Übertragungskanals („Tunnels“) ausgetauscht werden. Andererseits können im ESP auch Daten der übergeordneten Protokolle (z.B. TCP) verschlüsselt werden.

Obwohl es natürlich bereits heute mit entsprechenden Erweiterungen möglich ist, über das Internet vertraulich und authentisch zu kommunizieren, bietet das Internet mit IPv6/IPsec erstmals eine Sicherheitsbasis, auf die sich alle Dienste, besonders auch solche, die im Bereich E-Business Anwendung finden, stützen können.

Die zu verwendenden kryptographischen Verfahren sind nicht näher spezifiziert. Aus Kompatibilitätsgründen wird bei Authentication Headern zumindest der MD5-Hash-Algorithmus vorgeschrieben und bei Encapsulating Security Payloads der DES (Data Encryption Standard) in der Betriebsart CBC (Cipher Block Chaining). Ebenso ist das Schlüsselmanagement nicht Bestandteil der Spezifikationen.

---

[6] 1. Im Frühjahr und Sommer 2000 wurden die Firmen Yahoo und Amazon beispielsweise Opfer von Angriffen auf die Verfügbarkeit. Deren Dienstleistungen waren für mehrere Stunden nicht mehr erreichbar (http://www.heise.de/bin/tp/issue/dl-artikel.cgi?artikelnr=5785&mode=html). 2 Einem Betrüger gelang es mit Hilfe von ausgespähten Sozialversicherungsnummern, im Namen seiner Opfer Kreditschecks über Beträge von bis 44 000 Dollar zu erhalten und damit einkaufen zu gehen (http://www.heise.de/newsticker/data/jk-04.09.00-000/).

[7] R. Atkinson: Security Architecture for IP. RFC 1825, NRL, August 1995; R. Atkinson: IP Authentication Header. RFC 1826, NRL, August 1995; R. Atkinson: IP Encapsulating Security Payload. RFC 1827, NRL, August 1995.

Eine der heute schon verbreiteten Anwendungen von IPSec ist das Tunneling von firmeninternen Daten zwischen entfernten Unternehmensstandorten über sog. **Virtual Private Networks (VPN)**.

2. Die wichtigsten Anwendungen

2.1 Wichtige Internet-Dienste im Bereich E-Business

Electronic Mail (E-Mail)

Die elektronische Post ist ein netzübergreifender Mitteilungsdienst. Ein Internet-Benutzer erhält von seinem Provider oder seinem Systemadministrator eine E-Mail-Adresse, die aus einer persönlichen Kennung und dem Domain- oder Rechnernamen besteht, verbunden durch das at-Zeichen @. Beispiel: federrath@inf.tu-dresden.de.

Den Transport von elektronischer Post übernimmt im Internet das **Simple Mail Transfer Protocol (SMTP).** Es arbeitet nach dem Store-and-forward-Prinzip. Das bedeutet, der Rechner, von dem die E-Mail abgesetzt wird, schickt sie nicht direkt an den Zielrechner, sondern an einen günstig gelegenen Rechner auf dem Weg zum Zielrechner. Erhält ein günstig gelegener Rechner die E-Mail, so speichert er sie lokal (*store*), sucht einen weiteren günstig gelegenen Rechner und schickt die E-Mail dorthin weiter (*forward*). Dieses Store-and-forward wiederholt sich, bis der Zielrechner erreicht ist. Bei der Übertragung einer E-Mail kann jeder Store-and-forward-Rechner den Inhalt der Nachricht mitlesen. Es empfiehlt sich deshalb, den Nachrichteninhalt zu verschlüsseln. Hierzu können z.B. das Programm **Pretty Good Privacy (PGP)** oder der Standard **S/MIME (Secure Multipurpose Internet Mail Extensions)** eingesetzt werden.

Da nicht alle Teilnehmer auf jedem SMTP-Server im Internet eine Nutzerberechtigung haben können, sind diese Server meist über einen speziellen Zugang (meist TCP-Port 25) für das SMTP auch **ohne Login-Berechtigung** zugänglich. Ein Benutzer kann also mit einem SMTP-Server ohne Account kommunizieren. Diese Form des Zugangs zu einem SMTP-Server wird von den meisten PC-basierten E-Mail-Programmen verwendet, um E-Mails ins Internet zu schicken. Die im E-Mail-Programm angegebene Absenderadresse wird vom SMTP-Server meist ungeprüft übernommen. Deshalb kann der Empfänger von E-Mails nicht sicher davon ausgehen, daß der Absender der E-Mail authentisch ist.[8] Das Absenden von E-Mails mit gefälschter Absenderadresse wird **Mail-Spoofing** genannt. Um sich vor Mail-Spoofing zu schützen, muß der Absender seine Nachrichten authentisieren, z.B. indem er sie digital signiert. Hierfür können ebenfalls PGP oder S/MIME eingesetzt werden.

Das Abholen von E-Mails vom Mail-Server übernimmt das **Post Office Protocol (POP)**, oder neuerdings das **Interactive Mail Access Protocol (IMAP)**. Hierzu muß sich der Benutzer gegenüber dem Mail-Server authentisieren, um fremden Zugriff auf die empfangenen E-Mails zu verhindern. Dies wird z.B. durch Abfrage eines Passworts realisiert.

Hypertext Transfer Protocol (HTTP) und World Wide Web (WWW)

Das Hypertext Transfer Protocol (HTTP) ist das Kernprotokoll des World Wide Web (WWW). Die in einem Browser angezeigten Dokumente werden mit dem HTTP vom Webserver zum Browser übertragen. Die Dokumente im WWW enthalten Querverweise, sog. Links, zu anderen Hypertext-Dokumenten, die ihrerseits wieder Verweise enthalten können. Dadurch entsteht ein Netz (Web) von verlinkten Dokumenten. Als Dokumentbeschreibungssprache wird im WWW hauptsächlich die **Hyptertext Markup Language (HTML)** eingesetzt.

---

[8] Herbert Damker, Hannes Federrath, Michael J. Schneider: Maskerade-Angriffe im Internet. Eine Demonstration für Unsicherheit. Datenschutz und Datensicherung DuD 20/5 (1996) 286-294.

Protokoll	Aufbau der URL
HTTP	http://Rechneradresse:Port/Pfad/Datei Beispiel: http://www.inf.tu-dresden.de/~hf2
HTTPS (Verschlüsseltes HTTP mit dem Protokoll Secure Socket Layer, SSL)	https://Rechneradresse:Port/Pfad/Datei Beispiel: https://www.inf.tu-dresden.de/
FTP (File Transfer Protocol, dient meist dem Übermitteln großer Dateien)	ftp://Login:Passwort@Rechneradresse:Port/Pfad/Datei Beispiel: ftp://ftp.inf.tu-dresden.de/ (anonymes ftp)
E-Mail	mailto://Name@Domain
News	news:Newsgroup Beispiel: news:comp.os.macos
Telnet (Terminal Network)	telnet://Login:Passwort@Rechneradresse:Port Beispiel: feder:xv5d390n@server.inf.tu-dresden.de
Lokale Datei	file://Rechnername/Pfad/Datei

Tabelle 1. Ausgewählte Protokolle und ihre Verwendung in URLs

Ein Link wird durch einen **Uniform Resource Locator (URL)** realisiert. URLs vereinheitlichen die Bezeichnung von Dokumenten im Internet:

Aufbau einer URL: Protokoll://Rechneradresse:Port/Ressource

Protokolle können z.B. http, ftp, mailto sein (siehe auch Tabelle 1). Die Angabe des Ports mit führendem Doppelpunkt wird häufig weggelassen, wenn der Dienst auf dem standardisierten Port zur Verfügung steht. Die Ressource ist meist ein Dateiname mit führender Pfadangabe. Beispielsweise verweist die URL „http://www.inf.tu-dresden.de/~hf2/index.html“ auf eine Webseite auf dem Rechner www.inf.tu-dresden.de, von dem aus dem Verzeichnis „/~hf2/“ die Datei „index.html“ angefordert wird. Da der Webserver auf dem für HTTP standardisierten Port 80 arbeitet, kann die Portnummer weggelassen werden. Die Abbildung 4 zeigt vereinfacht den typischen Protokollablauf am Beispiel eines Abrufs einer HTML-Seite.

Mit HTTP können beliebige Datentypen (Texte, Bilder, Sound, binäre Dateien) übertragen werden.

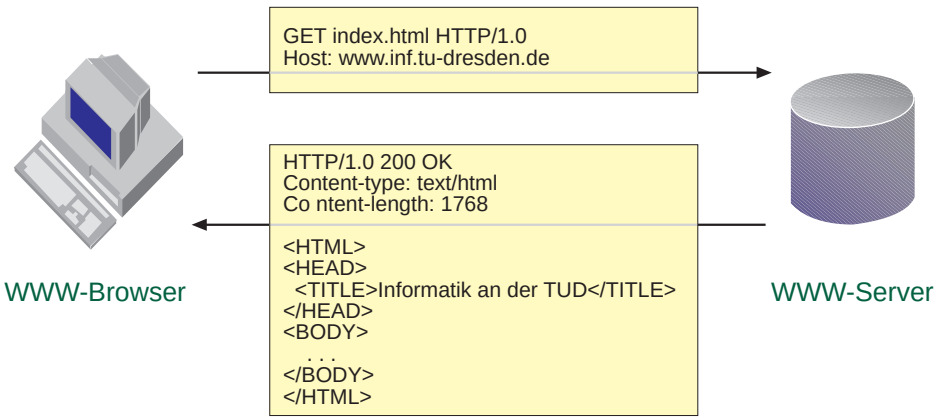

Abbildung 4. Beispielhafter Protokollablauf beim Abruf einer HTML-Seite

News

Der weltweite Informationsservice Usenet News arbeitet wie ein weltweites Schwarzes Brett und ist ein verteiltes Datenbanksystem, bei dem jeder News-Beitrag an alle News-Server weitergegeben und gespeichert wird. Hierzu wird das **Net News Transfer Protocol (NNTP)** verwendet. Es werden in themenspezifischen News-Gruppen vorrangig Diskussionen geführt, aber auch Daten (Bilder, Binärdateien etc.) öffentlich verteilt und verfügbar gemacht. Der Endbenutzer verbindet sich mit dem nächstgelegenen News-Server über seinen News-Client, der z.B. im Web-Browser integriert ist.

Die News-Gruppen sind hierarchisch nach Themengebieten geordnet. Beispielsweise beschäftigen sich die Newsgruppen, die mit der Bezeichnung sci.crypt beginnen, mit wissenschaftlichen Fragen zur Kryptographie. Deutschsprachige Diskussionsforen beginnen mit de. Beispielsweise beschäftigt sich die Newsgruppe de.talk.sex völlig offensichtlich mit in Deutsch geführten Diskussionen zum Thema Sex.

Telnet

Mit dem Dienst Telnet (Teletype Network) ist eine interaktive, aber rein **textbasierte Benutzung eines entfernten Rechners** möglich. Über eine sog. Konsolen- bzw. Terminalsoftware können Befehle auf dem entfernten Rechner über das Internet ausgeführt werden. Durch den Siegeszug von graphischen Benutzungsoberflächen (Graphical User Interfaces, GUI) verliert Telnet im Bereich E-Business zunehmend an Bedeutung.

Noch vor einigen Jahren arbeiteten nahezu alle Buchungs-, Abfrage- und Datenverwaltungssysteme textbasiert und kommunizierten teilweise über das Telnet-Protokoll mit dem Zentralrechner (Host). Heute wird Telnet vorzugsweise für die entfernte Administration von Rechnern angewendet und dürfte noch für Spezialanwendungen im Bereich B2B verwendet werden. Im Bereich B2C hat Telnet nahezu keine Bedeutung.

Wireless Application Protocol (WAP)

Das Wireless Application Protokoll (WAP) ermöglicht die Kommunikation zwischen einem WAP-Server und kleinen, mobilen, drahtlosen Endgeräten. Mobiltelefone, die mit einem WAP-Browser ausgestattet sind, können so Hypertext-Dokumente abrufen, deren Gestaltung sich an den sehr kleinen und bisher wenig Gestaltungsmöglichkeiten bietenden Displays der Geräte orientiert. Die Beschreibungssprache für WAP-Dokumente nennt sich **Wireless Markup Language (WML)**.

Prinzipiell könnten über WAP alle aus dem World Wide Web bekannten Dienste angeboten werden. In einigen Jahren werden sich die Darstellungsmöglichkeiten der mobilen Displays drastisch verbessert haben.

Durch die Verknüpfung von Ortsinformation (gegenwärtiger Aufenthaltsort des mobilen Teilnehmers) mit Stadtinformationssystemen ermöglicht WAP interessante Anwendungen, z.B. interaktive Wegweiser und Navigation: Der Kunde (bzw. sein Mobiltelefon) lokalisiert sich und ruft einen WAP-basierten Stadtplan auf, um sich zu seinem Ziel leiten zu lassen.[9] Ebenso ist der zielgenaue Ruf von Rettungsdiensten, Taxis, etc. möglich, wenn der Kunde (bzw. sein Mobiltelefon) seinen genauen Standort mitteilt.

Bei der **Lokalisierung** von Teilnehmern ist zu beachten, daß aus Datenschutzgründen keine ständige und keine unbemerkte Lokalisierung des Teilnehmers möglich sein darf. Es sollten nach Möglichkeit Techniken zum Schutz vor Lokalisierung[10] eingesetzt werden.

Mit der optionalen Erweiterung WTLS (Wireless Transport Layer Security) können WAP-Verbindungen auch **verschlüsselt** und authentisiert werden.

Virtuelle E-Shopping-Systeme

Eine komplexe Anwendung der Internet-Dienste stellen virtuelle E-Shopping-Systeme dar. Dabei handelt es sich um (meist) Web-basierte Einkaufssysteme. Auf einer Website werden dem Kunden in Katalogform verschiedene Waren angeboten. Artikelsuche und multimediale Präsentationsfähigkeiten erhöhen gegenüber herkömmlichen Home-Shopping-Möglichkeiten den Nutzwert von E-Shopping.

---

[9] Mit einem Navigationssystem kann man dies natürlich heute schon, doch erstens muß man ein solches besitzen und zweitens stets dabei haben. Dagegen könnte man über das Mobiltelefon einen solchen Service „on-demand“ nutzen und bezahlen.

[10] Hannes Federrath: Sicherheit mobiler Kommunikation. DuD Fachbeiträge, Vieweg, Wiesbaden 1999.

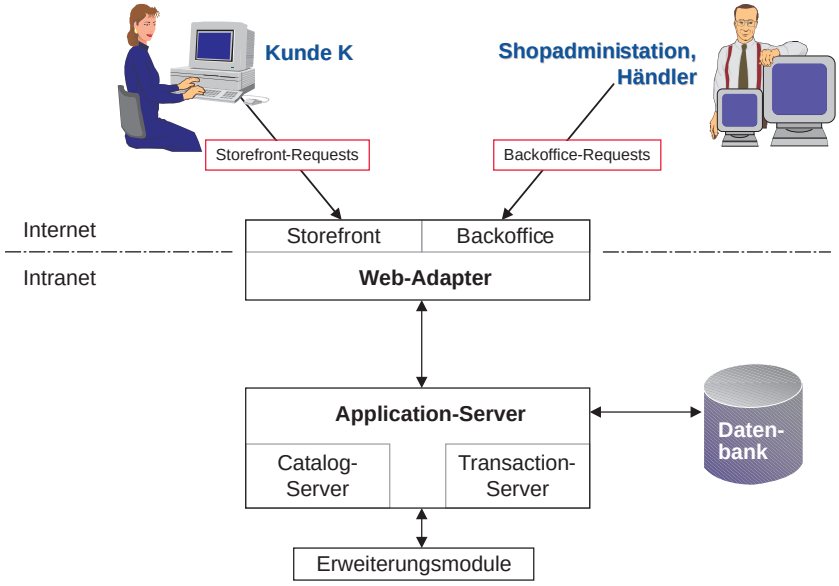

Abbildung 5. Komponenten eines E-Shops

Ein Online-Shop besteht aus folgenden Kernkomponenten (siehe auch Abbildung 5):

- **Web-Adapter mit Storefront und Backoffice:** Der für den Shop-Besucher sichtbare Teil des E-Commerce-Systems ist die „Storefront“. Händler und Administratoren greifen zur Einrichtung und Verwaltung des Shops auf das „Backoffice“ zu.
- **Application-Server:** Hier liegt die Hauptlast des Systems. Der Application-Server generiert aus den Katalogdaten dynamisch die im Browser des Kunden dargestellten Web-Dokumente und ist für die Abwicklung aller Transaktionsvorgänge bei Bestellungen zuständig. Außerdem verwaltet der Application-Server die Zugriffsrechte für die Datenbank.
- **Datenbank:** Die Datenbank enthält alle Katalogdaten und verwaltet die gespeicherten Kunden-, Händler- und Transaktionsdaten. Aufgrund der Sensibilität der personenbezogenen Daten darf die Datenbank niemals direkt mit dem Webadapter kommunizieren, sondern es müssen alle Anfragen über den Application Server laufen, der die Rechteverwaltung übernimmt.

Der **Zugriffsschutz** in E-Shopping-Systemen basiert bisher meist auf Passwörtern. Zum Schutz der Vertraulichkeit werden, sobald personenbezogene Daten zwischen dem Web-Adapter und dem Kunden bzw. dem Web-Adapter und dem Backoffice ausgetauscht werden, diese über das HTTPS-Protokoll übertragen. HTTPS ist ein verschlüsseltes HTTP auf der Basis des Protokolls Secure Socket Layer (SSL).

2.2 Beschreibungssprachen und Datentypen

Beschreibungssprachen

Hypertext-Dokumente sind heute die wichtigste Klasse von Inhalten im Internet. Der Aufbau eines Dokuments wird durch eine Sprache beschrieben, die eine Maschine (hier: der Browser, das WAP-Handy oder andere Anzeigeinstrumente) interpretieren kann und in eine multimediale Darstellung des Dokumentinhalts verwandelt. **HTML (Hypertext Markup Language)** ist die Beschreibungssprache für Webseiten im Internet.**WML (Wireless Markup Language)** ist die Beschreibungssprache für Inhalte, die auf kleinen mobilen, vorzugsweise WAP-fähigen Endgeräten verarbeitet/dargestellt werden. **XML (Extensible Markup Language)** ist eine universelle Dokumentbeschreibungssprache, die zunehmend auch im World Wide Web Anwendung findet. XML ist eine Teilmenge von **SGML (Standardized Generalized Markup Language)**.

Inhalte von Hypertext-Dokumenten

In Hypertext-Dokumenten können mindestens folgende drei **Klassen von Inhalten** enthalten sein:

- **Textdaten**, d.h. formatierter und unformatierter Text mit allen möglichen gestalterischen Elementen,
- **multimediale Objekte**, also Bilder, Grafiken, Animationen, Audio- und Video-Ströme und
- **aktive Inhalte**. Hierzu zählt ausführbarer Code (JavaScript, Java-Applets, ActiveX).

Die Abbildung 6 zeigt ein einfaches Hypertext-Dokument mit seiner Darstellung im Web-Browser. Links ist der Inhalt des HTML-Dokuments zu sehen, rechts dessen Anzeige im Web-Browser. Das Dokument enthält einen Hyperlink auf die URL http://www.datenschutz.de/ und das eingebettete Grafikobjekt note.gif sowie einige Textauszeichnungen.

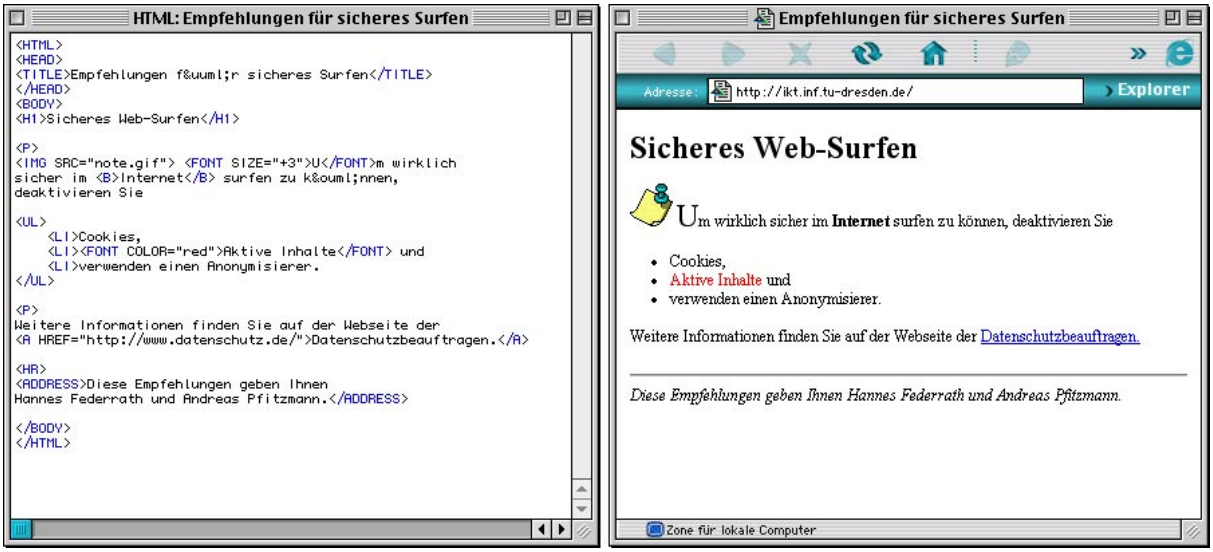

Abbildung 6. Beispiel eines Hypertext-Dokuments.

Durch die sog. **Plug-In-Architektur** der Browser ist es möglich, den Funktionsumfang der Hypertext-Dokumente zu erweitern. Die Installation von Plug-Ins und die Ausführung aktiver Inhalte birgt für den Endbenutzer einige **Sicherheitsprobleme und Risiken**. So ist es möglich, daß bösartige ActiveX-Controls und Plug-Ins unbemerkte Lese- und Schreibzugriffe auf lokale Dateien ausführen können. Auch Java-Applets dürfen nach Rückfrage an den Benutzer lokale Dateien schreiben und lesen. Dem Benutzer eines Betriebssystems ohne Zugriffs- und Speicherschutz ist deshalb dringend zu raten, nur Plug-Ins und ActiveX-Controls von vertrauenswürdigen Herstellern zu benutzen bzw. vom Browser ausführen zu lassen und genau darauf zu achten, welche Rechte ein Java-Applet bei der Ausführung erhält. Dies alles ist besonders wichtig, wenn der PC auch für sensible Anwendungen, z.B. Electronic Banking, benutzt wird. Eigentlich ist die vollständige Deaktivierung aktiver Inhalte empfehlenswert. Diese Empfehlung steht allerdings im Widerspruch zu dem Trend, mehr und mehr E-Business-Anwendungen (z.B. Electronic Banking) durch aktive Inhalte zur realisieren oder aufzupeppen, was praktisch einem Aktivierungszwang aktiver Inhalte gleich kommt. Der Benutzer muß folglich abwägen zwischen Sicherheit und Bequemlichkeit.

Grafiken und Bilder

Grafiken und Bilder im Internet werden meistens in den Formaten **GIF (Graphics Interchange Format)**, **PNG (Portable Network Graphics)** und **JPEG (Joint Photographic Expert Group)** verwendet.

GIF arbeitet mit einer verlustfreien Kompression der Bilddaten. Allerdings darf das Bild nur maximal 256 Farbwerte enthalten, was die Formate für Fotos ungeeignet, aber für graphische Darstellungen sehr geeignet macht. JPEG dagegen arbeitet mit verlustbehafteter Kompression und eignet sich für die komprimierte Übertragung und Speicherung von Fotos und anderem Bildmaterial mit Farbverläufen.

PNG setzt sich bei der Gestaltung von Web-Dokumenten aufgrund seiner Universalität immer mehr gegen GIF und JPEG durch.

Audio und Video

Bei Audio und Video ist zu unterscheiden, ob die Mediendaten bereits während der Übertragung konsumiert werden oder zunächst heruntergeladen und dann Offline verwendet werden.

Audio- und Videokommunikation über das Internet in Echtzeit, sog. **Streaming**, stellt hohe Anforderungen an die Dienstgüte und Übertragungsbandbreite des Kanals. Für eine hohe Übertragungsqualität werden Verfahren verwendet, die den Datenstrom hochkomprimiert und tolerant gegen Paketverluste machen. Ein erneutes Senden eines verlorengegangenen Datenpaketes wäre unnütz, da der fehlende „Abschnitt“ des Datenstroms zeitlich hinter dem aktuell gesendeten liegt.

Für Videoübertragung (inkl. des Audio-Streams) kommen hauptsächlich die standardisierten Formate der **Motion Picture Expert Group (MPEG)** sowie die proprietären Formate von **Quicktime** (Apple Computer Inc.) und **RealVideo** (Real Inc.) zum Einsatz. Allerdings sind die Software-Player von Apple und Real auch in der Lage, viele Standard-Formate wiederzugeben.

Als Datenformat für Audio im Internet ist derzeit das Format **MP3** (eigentlich richtig MPEG-1 Layer 3 genannt) gebräuchlich, das einen Audio-Stream auf etwa 1/10 der Datenmenge reduziert, nahezu ohne hörbare Qualitätsverluste.