

Veranstaltung
Pr.-Nr.: 10 1023 V

Wirtschaftsinformatik für
Wirtschaftswissenschaftler

Vernetzung, Internet und **WWW**

Dr. Chris Bizer
WS 2007/2008

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Kapitel 4: Vernetzung, Internet und WWW

- 1. Vernetzung**
 1. Aufbau von Rechnernetzwerken
- 2. Das Internet**
 1. Struktur des Internets
 2. Datenübertragung mit TCP/IP
- 3. Das World Wide Web (WWW)**
 1. Hypertext Markup Language (HTML)
 2. Hypertext Transfer Protocol (HTTP)
 3. Web-basierte Anwendungssysteme

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

1. Vernetzung

Rechnernetz: Verbindung von Rechnern unter Verwendung eines gemeinsamen Protokolls (Regeln zur Kommunikation).

■ Ziele von Rechnernetzen

- Datenübertragung (Filetransfer, E-Mail, Client/Server-Anwendungssysteme)
- Gemeinsame Ressourcennutzung (Daten, Speicher, Software, Peripheriegeräte)
- Last- und Leistungsverbund von Rechnern (z.B. Kopplung von Großrechnern)

■ Komponenten von Rechnernetzwerken

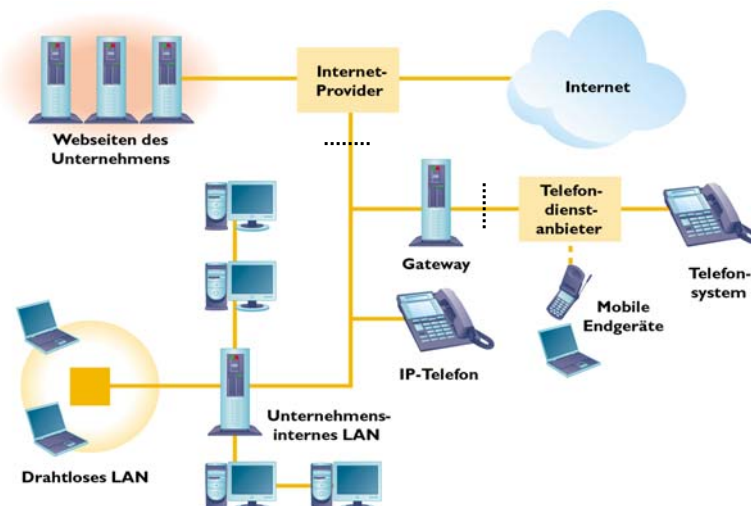
- Rechner mit Netzwerkkarte bzw. Modem
- Datenübertragungskanäle (Netzkabel, Lichtwellenleiter, Telefonleitung, Satelliten, Richtfunk)
- Übertragungsprotokolle (gemeinsame Sprache für den Verbindungsaufbau, den Datentransfer und das Handling von Übertragungsfehlern)

■ Arten von Rechnernetzen

- LAN (Local Area Network) verbinden Knoten in einem begrenzten Gebiet, z.B. Firma
- WAN (Wide Area Network) verbinden Knoten über größere Distanzen (> 10 km)
- MAN (Metropolitan Area Networks), beziehen sich auf eine Region (10-100 km)

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Beispielhafter Aufbau eines Unternehmensnetzwerks

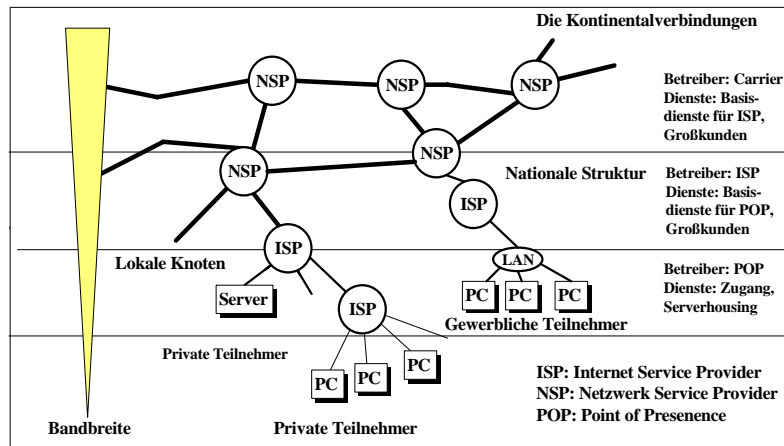


- Meist heterogene, gewachsene Netzwerkstrukturen
- Nutzung der TCP/IP- Protokollfamilie als gemeinsame Kommunikationssprache
- Verbindung über Gateways zum Internet und Telefonnetz

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

2. Das Internet

- Das Internet verbindet tausende voneinander unabhängige Netzwerke auf Basis der TCP/IP-Protokollfamilie.
- Strukturansicht des Internets:



- Internet Service Provider (ISP): Bieten Firmen und Endkunden Verbindungen zum Internet (Wahlverbindungen, Standleitungen) und Server-Dienste (Webspace, Applikationen) an.
- Network Service Provider (NSP): Betreiben Netzwerke (Backbones) zur Datenübertragung zwischen ISPs.

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

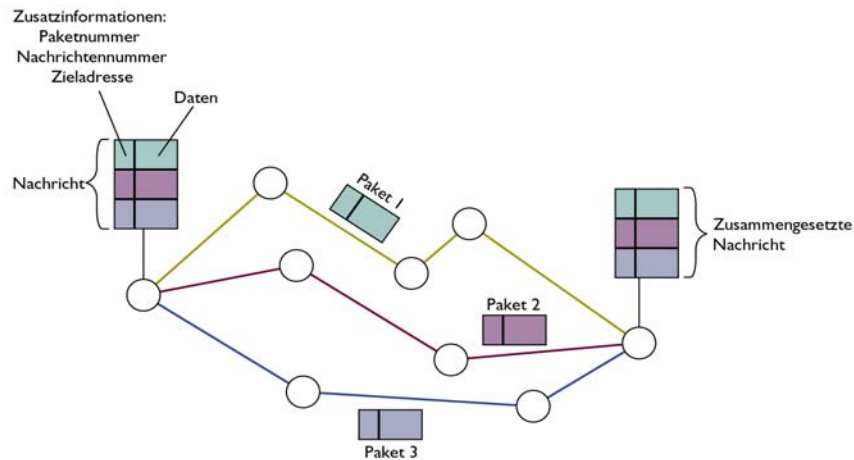
TCP/IP

Die TCP/IP-Protokollfamilie dient allen Internet-Anwendungen, wie z.B. WWW, eMail, Skype oder Chat, zur Datenübertragung.

- Internet Protocol (IP)
 - Zur Kommunikation über das Internet benötigen Rechner eine eindeutige IP-Adresse.
 - Diese IP-Adresse besteht aus vier natürlichen Zahlen 1-254, die zur Lesbarkeit durch einen Punkt getrennt werden, z.B. 160.45.141.165
 - Das Internet Protocol (IP) definiert den Aufbau und die Struktur dieser IP-Adressen und stellt Mechanismen zur Transport (Routing) von Datenpaketen von einer IP-Adresse zu einer Anderen zur Verfügung.
- Transmission Control Protocol (TCP)
 - Das Transmission Control Protocol (TCP) zerlegt die zu übertragende Daten in Pakete einheitlicher Größe.
 - Ein Datenpaket umfasst neben den zu übertragenden Nutzdaten auch Kopfdaten wie Absender-, Zieladresse und Paketnummer.

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Datenübertragung mit TCP/IP



- IP-Datenpakete werden unabhängig voneinander von Router zu Router weitergegeben, bis sie ihre Zieladresse erreicht haben.
- Der Zielcomputer setzt die Pakete entsprechend ihrer Nummerierung zusammen. Fehlende Pakete werden automatisch erneut angefordert.

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Datenübertragung mit TCP/IP

- Router analysieren die ankommenden Datenpakete und ermitteln an Hand der Zieladresse und ihrer internen Routing-Tabelle den weiteren Weg durch das Netz.
- Beispiel: Weg einer Anfrage an www.spiegel.de durch das Internet

1	160.45.141.165	wrz02614.wiwiss.fu-berlin.de
2	160.45.141.129	kuehl.wiwiss.fu-berlin.de
3	130.133.98.2	ice.spine.fu-berlin.de
4	188.1.33.33	ar-fuberlin1.g-win.dfn.de
5	188.1.18.21	cr-frankfurt1.g-win.dfn.de
6	80.81.192.14	BB3.Frankfurt.is-bone.net
7	194.163.251.75	IVR11-vl15.Hamburg.cr.ision.net
8	194.64.249.245	www.spiegel.de



- Sicherheitsrisiken bei der Datenübertragung mit TCP/IP
 - Die Datenpakete werden unverschlüsselt übertragen und können prinzipiell auf jedem Router gelesen werden.
 - Für Anwendungen, die eine sichere Datenübertragung erfordern (z.B. Online Shopping, Homebanking) ist daher der Einsatz wirkungsvoller kryptographischer Verfahren, wie SSL oder VPNs, unerlässlich (siehe Kapitel 6: Sicherheit).

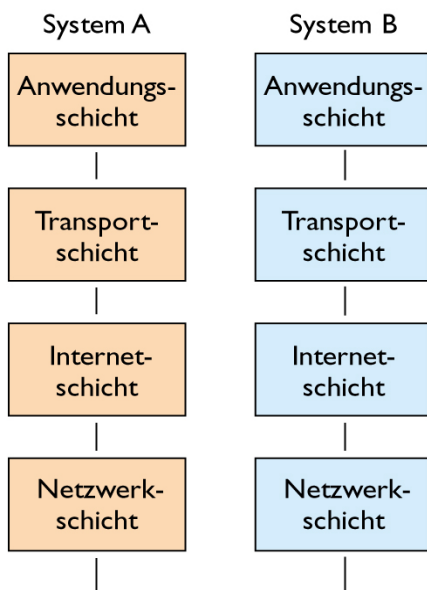
Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Das Domain-System

- Das Internet Protocol leitet Datenpakete auf Basis von IP-Adressen weiter.
- IP-Adressen sind für Menschen schlecht merkbar.
- Daher werden zur leichteren Adressierung hierarchisch angeordnete Domain-Namen verwandt, z.B. www.wiwiss.de
- Von Domain-Servern werden Domainnamen anhand einer Tabelle in IP-Adressen übersetzt.
- Für die Vergabe von Domainnamen ist in Deutschland die DE-NIC (<http://www.nic.de/>) zuständig.
- Auf internationaler Ebene regelt ICAN (<http://www.icann.org>) die Vergabe von Domainnamen durch Landesgesellschaften wie DE-NIC und die Einrichtung neuer Top-Level-Domains wie [.info](http://www.info) [.biz](http://www.biz) [.eu](http://www.eu)

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Die vier Schichten des TCP/IP-Referenzmodells



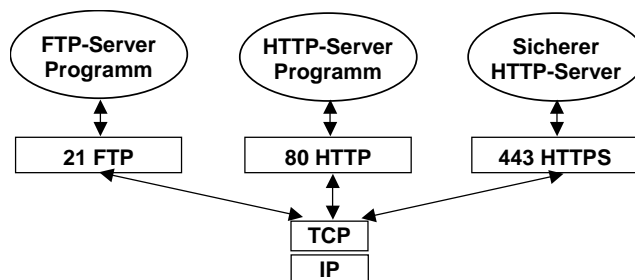
- **Anwendungsschicht**
 - regelt durch Nutzung entsprechender Protokolle die Kommunikation zwischen Anwendungsprogrammen wie z.B. Webbrowser und Webserver (HTTP Protokoll), Mailclient und Mailserver (POP bzw. SMTP Protokoll).
- **Transportschicht**
 - regelt die Aufteilung in Datenpakete und den Transport der Pakete zwischen dem sendenden und empfangenden System (TCP Protokoll).
- **Internetschicht**
 - regelt die Weiterleitung der Daten von einer Netzwerkkomponente zur nächsten bis hin zum Zielsystem (IP Protokoll).
- **Netzwerkschicht**
 - ist für die physikalische Datenübertragung zuständig (elektrische Impulse, Funk- oder Lichtwellen).

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Protokolle der Anwendungsebene und Port-Nummern

- Alle Protokolle der Anwendungsebene greifen zur Datenübertragung auf TCP/IP zurück.
- Damit mehrere Server-Programme (z.B. Web-Server und Mail-Server) auf einem Rechner gleichzeitig betrieben werden können, werden zusätzlich zur IP-Adresse des Rechners Portnummern verwendet: `http://160.133.141.36:80` `ftp://160.133.141.36:21`
- Anhand der Portnummer erkennt der Server welcher Dienst und damit welches Server-Programm angesprochen werden soll.

Port-Nr.	Protokoll	Bezeichnung
21,22	FTP	File Transfer Protocol
80	HTTP	Hypertext Transfer Protocol
110	POP	Post Office Protocol
443	HTTPS	HTTP über SSL



Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

3. Das World Wide Web (WWW)

Weltweiter Informationsraum, der aus Dokumenten besteht, die über Hyperlinks miteinander verbunden sind.

- Wichtiger Bestandteil des WWW sind einheitliche Adressen (URLs) zum Abruf von Dokumenten von unterschiedlichen Servern.
- **Uniform Resource Locators (URLs)** haben folgende allgemeine Form:
`protokoll://server/verzeichnis/datei`
- z.B. `http://www.wiwiss.fu-berlin.de/suhl/index.html`
- Dokumente werden im **Hypertext-Format (HTML)** als Dateien auf dem Server gespeichert.
- Über das **Hypertext Transfer Protokoll (HTTP)** werden die Dokumente vom Browser des Nutzers angefordert und zu diesem übertragen.

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

HTML (Hypertext Markup Language)

- HTML ist eine Auszeichnungssprache für Web-Dokumente

- Auszeichnungssprachen

- Auszeichnen (Markup) bezeichnet den Prozess der Informationsanreicherung von Textdokumenten.
- Durch die Auszeichnung von Teilen des Textes durch sogenannte „Tags“ (Marken) wird sowohl das Erscheinungsbild als auch die inhaltliche Struktur des Dokumentes festgelegt.

```
<H1>Beispiel</H1>
<P><B>Hallo</B>, ich bin <I>HTML</I></P>
```

Start-Tag: Fett

End-Tag: Fett

- Einführung in HTML

- Stefan Münz: SelfHTML 8.0, <http://selfhtml.org/>

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Aufbau eines HTML-Dokuments

```
<html>
  <head>
    <title>Beispieldokument</title>
    <meta name="author" content="Chris Bizer">
    <meta name="keywords" content="HTML, Beispiel">
  </head>
  <body bgcolor="#FFFFFF">
    <H1>Beispiel</H1>
    <P>
      <B>Hallo</B>, ich bin ein ganzes
      <I>HTML-Dokument</I>.
    </P>
  </body>
</html>
```

Head Bereich:
Dokument-
informationen
(Autor, Key-
words etc.)

Body Bereich:
Eigentliches
Dokument

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Hyperlinks

■ Relativer Link zwischen Dokumenten auf einem Server

<p>Das ist ein `Link`.</p>

Start-Tag:
Hyperlink

Ziel des
Hyperlinks

Text

End-Tag:
Hyperlink

■ Beispiel

- Verzeichnis1/Dokument1.htm
- Verzeichnis1/Dokument2.htm

■ Absoluter Link zwischen unterschiedlichen Servern

<p>Und das ist ein `absoluter Link`.</p>

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Webbrowser und Webserver

■ Aufgaben des Webbrowsers

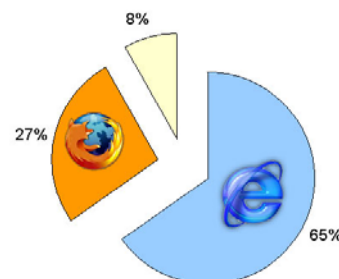
- Abrufen von HTML-Seiten von einem Webserver mittels des HTTP Protokolls
- Darstellen der HTML-Seiten
- ggf. ausführen von Programmcode in der HTML-Seite (z.B. Javascript, Java Applet)

■ Aufgaben eines Webserver

- Warten auf Anfragen: Der Server horcht an Port 80 nach HTTP-Requests.
- Anfragen beantworten, in dem er ein HTML-Dokument, das direkt aus dem Dateisystem stammt (statische Webseite) oder von einem Serverprogramm erzeugt wird (dynamische Webseite) als HTTP-Response an den Browser zurückschickt.
- Zugriffskontrolle auf Dateien und Verzeichnisse
- Protokollierung von Seitenaufrufen (logging)

Marktanteile Webbrowser

(Deutschland, 6/2006, BusinessWeek)



Marktanteile Webserver

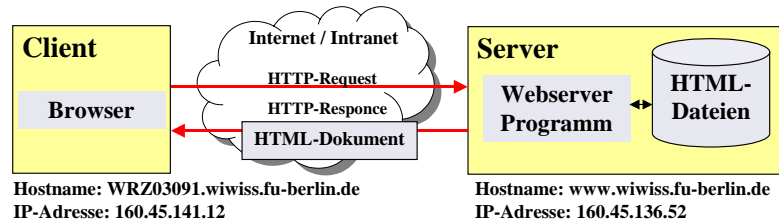
(Global, 11/2007, Netcraft)

Server	Marktanteil
Apache	50.48%
Microsoft IIS	34.94 %

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Das HTTP-Protokoll

■ Gemeinsame Sprache zwischen Webbrowser und Webserver.



■ Inhalt eines HTTP-Requests

- Dokumentenanforderung : „Bitte schicke mir das Dokument /suhl/index.htm“
- Informationen über den Client: „Ich bin ein Internet Explorer und spreche am liebsten deutsch.“

■ Inhalt einer HTTP-Response

- Angefordertes Dokument: „Alles hat geklappt, anbei schicke ich das angeforderte Dokument. Bei dem Dokument handelt es sich um eine HTML-Seite“
- Informationen über den Server: „Ich bin ein Apache Server, unterstütze PHP und laufe unter Windows XP.“

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

HTTP-Beispiel-Request

```
GET /books/index.htm HTTP/1.1
Accept: image/png, image/gif, *.*
Accept-Language: de, en
Accept-Encoding: gzip
User-Agent: Mozilla/4.0 (compatible; MSIE 5.5; Windows 98)
Host: www.amazon.com
Referer: http://www.amazon.com/index.htm
```

■ Dokumentenanforderung

```
GET /books/index.htm HTTP/1.1
```

Methoden Dokument Protokoll-Version

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Die HTTP-Request-Headers

Header	Bedeutung
Accept: image/png, image/gif, *.*	Bevorzugte Dokumentenformate. Angabe mittels MIME-Types (Multipurpose-Internet-Mail-Extension-Standard)
Accept-Language: de	Bevorzugte Sprache
Accept-Encoding: gzip	Unterstützte Übertragungskodierungen
User-Agent: Mozilla/4.0	Browser-Kennung (hier Internet Explorer)
Referer: http://www.amazon.com	URL des Dokuments, das auf das angeforderte Dokument verweist (Von woher kommt der Surfer?)

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Aufbau einer HTTP-Response

```
HTTP/1.1 200 OK
Date: Mon, 18 Jan 2005 20:54:26 GMT
Server: Apache/1.3.6 (UNIX)
Last-Modified: Mon, 06 Dec 2003 14:06:11 GMT
Content-length: 6345
Content-Type: text/html

<html>
  <head><title>CB CD-Shop</title></head>
  <body><h1>Willkommen beim CB CD-Shop</h1> ....
```

■ Erste Zeile der Response

HTTP/1.1 200 OK



Protokoll-Version Statuscode Erläuterung

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

HTTP-Response-Headers und Statuscodes

Header	Bedeutung
Date: Mon, 18 Jun 2001 20:54:26 GMT	Zeitpunkt der Erstellung der Response.
Server: Apache/1.3.6 (UNIX)	Servertyp und Betriebssystem
Content-length: 6345	Größe des Dokuments in Bytes
Last-Modified: Mon, 06 Dec 1999 14:06:11 GMT	Zeitpunkt der letzten Modifikation des angeforderten Dokuments
Content-Type: text/html	Erklärt Browser, um was für ein Dokument es sich handelt (application/pdf, image/gif)

Statuscode	Erklärung
200 OK	OK, die Response enthält das angeforderte Dokument
400 Bad Request	Syntaxfehler im Request
401 Unauthorized	Die Ressource ist passwortgeschützt. Der Client wird aufgefordert, sich mit Benutzername und Kennwort zu autorisieren.
403 Forbidden	Zugriff verweigert, ohne Möglichkeit zur Autorisation.
404 Not Found	Das angeforderte Dokument existiert nicht.

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Aufruf einer HTML-Seite (1)

1. Eingabe einer URL durch den Benutzer in den Browser

http://www.wiwiss.fu-berlin.de/suhl/index.htm
Protokoll Domain-Name Dokumentenpfad

2. Browser erkennt das HTTP-Protokoll

- Aufgabe: Anfragen an den Webserver stellen/Übertragung der Antwort

3. Browser stellt Anfrage (Request) an einen Domainname-Server

- Aufgabe: Übersetzung von Domain-Namen in IP-Adressen

4. Domain Name Server übersetzt den Domainnamen

- www.wiwiss.fu-berlin.de ⇒ IP 160.45.136.52
- Vorgehensweise: Tabelle mit Domainnamen und zugehörigen IP-Adressen

5. Aufbau einer TCP/IP-Verbindung mit dem Server

- Notwendige Informationen: IP-Adresse, Port-Nummer
- Standard-Port des HTTP Protokolls: 80

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Aufruf einer HTML-Seite (2)

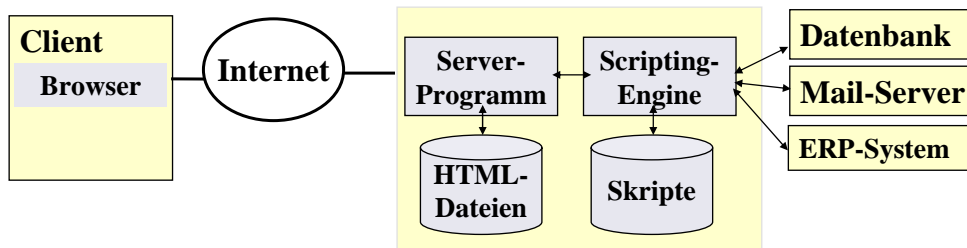
6. Browser fordert Dokument an
 - HTTP-Request: „Bitte schicke mir das Dokument /suhl/index.htm“
7. Server prüft HTTP-Request
 - Gibt es die angeforderte Datei?
 - Darf der Benutzer auf die Datei zugreifen?
8. Server protokolliert die Anfrage in der Log-Datei
9. Server sendet HTTP-Response (Antwort)
 - Möglichkeiten:
 - „OK. Schicke Datei“
 - „Habe Datei nicht gefunden!“
 - „Zugriff auf Datei nicht erlaubt“
 - „Server Fehler“
10. Browser stellt die empfangene Datei dar
 - Anzeigen der HTML-Seite
 - ggf. weitere Browser-Requests für Seitenelemente wie Grafiken
 - ggf. Ausführen von Programmcode in der Datei
 - ggf. Zugriff auf externe Programme zur Darstellung (Flash, PDF)

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Web-basierte Anwendungssysteme

Ein web-basiertes Anwendungssystem (WBAS) ist ein Anwendungssystem, das einen Webbrowser als Userinterface verwendet.

- Beispiele: E-Shop, Homebanking, Blackboard LMS
- Der Webserver liefert kein vorhandenes HTML Dokument, sondern erzeugt die Antwort unter Rückgriff auf eine Datenbank oder sonstige Backoffice-Anwendungen.
- Architektur web-basierter Anwendungssysteme



- Die Applikationslogik lässt sich auf den Server und den Client verteilen:
 - Programmausführung auf dem Client, z.B. Javascript, Flash
 - Programmausführung auf dem Server z.B. PHP, ASP, Java Servlets

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Ablauf: Dynamische Erzeugung einer HTML-Seite

1. Eingabe einer URL durch den Benutzer in den Browser
 - Z.B. <http://www.wiwiss.fu-berlin.de/suhl/anmeldung.php>
2. Der Webserver erkennt das Skript an der Dateiendung und gibt Aufruf an die Scripting-Engine weiter.
3. Scripting-Engine führt den Programmcode des Skripts aus. Dabei:
 - HTML-Templates dynamisch mit Daten aus Datenbanken füllen
 - Verarbeitung von Formulareingaben
 - Zugriff auf Backoffice-Anwendungen wie Warenwirtschaftssysteme
 - Automatisches Verschicken von eMails Z.B. als Bestätigung einer Bestellung
5. Als Ergebnis erzeugt das Skript eine HTML-Seite.
6. Webserver schickt die dynamisch erzeugte HTML-Seite per HTTP an den Client

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)

Exkurs: Wie hacke ich einen Server?

■ Einbruch in den Server

■ Möglichkeit 1: Zugangskennungen erraten

- Mittels Hack-Tools, die automatisch Millionen Passwörter durchprobieren

■ Möglichkeit 2: Über Sicherheitslücken (Exploits) in der Server-Software in den Rechner einbrechen

- Beispiel Angriff auf den Internet Information Server:

```
GET /scripts/..%00../winnt/system32/cmd.exe /c+dir
```

■ Denial-of-Service Attacken

- Schritt 1: Einbruch in eine Reihe schlecht gesicherter Server und Installation von Angriffs-Tools auf diesen Servern (Botnetz).
- Schritt 2: Die Tools senden zu einem festgelegten Zeitpunkt Millionen von Anfragen mit gefälschten Absenderangaben an den Server, dem der eigentliche Angriff gilt.
- Folge: Angegriffener Server wird überlastet und reagiert auch auf echte Benutzeranfragen nicht mehr.

Freie Universität Berlin –Bizer: Wirtschaftsinformatik – WS07/08 (Version vom 15.11.07)