Vorlesung
Internet-Datenbanken

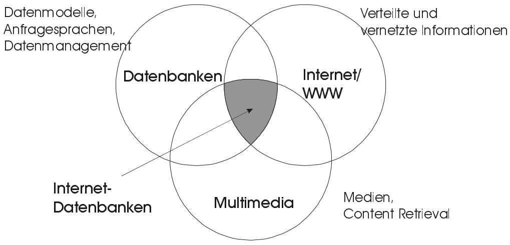
Wintersemester 1999/2000

Kai-Uwe Sattler
Universität Magdeburg
http://wwwiti.cs.uni-magdeburg.de/~sattler/

Überblick

- World Wide Web und Datenbanken
- Motivation
- Anwendungen
- Techniken
- Überblick zur Vorlesung
- Hinweise zur Übung

Internet-Datenbanken: Einflußgebiete



Datenbanken

- Vermeidung von Datenredundanz
- Effiziente Verwaltung großer Datenmengen
- Parallele Zugriffe auf gleiche Daten
- Datenschutz und Datensicherheit
- Datenunabhängigkeit
- Datenintegration

Internet/World Wide Web

- Internet
  - weltweiter Verbund von Rechnernetzen
  - Basis: Protokollfamilie TCP/IP
  - Dienste: ftp, Email, WWW
- World Wide Web
  - Internet-Dienst
  - Globales, vernetztes Informationssystem
  - Einfacher Zugriff auf weltweit verteilte Informationen

Multimedia

- Verwaltung und Präsentation unterschiedlicher Medien (Audio, Video, Bilder)
- Inhaltsbasierte Adressierung (Content Retrieval)
- Effiziente Speicherung
- Synchronisation
- Quality of Service

Motivation: Internet & Datenbanken

- WWW: weltweit größtes Informationssystem
  - 367 Mill. Internet-Hosts (7/98)
  - ca. 650 000 Websites (1/97)
  - 179 Mill. Online-Nutzer (7/99)
- Vorteile:
  - Einfacher Zugang zu verschiedenen Datenbanken
  - Präsentation von Datenbankinhalten
  - Nutzung von DBMS-Mechanismen

Probleme

- Protokoll (HTTP)
  - Zustandslosigkeit
  - GET/PUT-Primitive
  - URL-Adressierung, Parameterübergabe
- Benutzerschnittstelle (HTML, Forms)
  - Eingeschränkte Interaktivität
  - Konvertierung nach bzw. von HTML
- Weiterverarbeitung
  - Standard-API, Datenaustauschformat ?

Internet - Historie

- 30.8.1969: erster ARPANET Host angeschlossen (zweiter Host: 1.10.69)
- Juni 71: File Transfer Protocol (FTP)
- März 72: erste Email-Software
- Mai 74: erster Vorschlag TCP, Begriff Internet
- Januar 83: Einführung TCP/IP im ARAPNET
- November 83: Domain Name Service vorgestellt
- März 89: Idee des World Wide Web (Tim Berners-Lee, CERN)
- Dezember 90: Demo-Version des Web-Browsers (NeXT)

Internet - Historie

- Mai 91: Freigabe des WWW im CERN
- November 92: erster Entwurf für HTML
- April 93: WWW-Technologie wird Public Domain
- Juli 93: Internet-Entwurf für HTTP
- September 93: NCSA-Browser für X11, PC, Mac
- März 94: Gründung von Mosaic Corp. (später Netscape) durch Marc Andreesen
- Dezember 94: Navigator 1.0

Internet - Historie

- Mai 95: Freigabe von Java durch Sun Microsystems
- Dezember 95: Web-Server Apache freigegeben
- Dezember 95: JavaScript
- Januar 96: Freigabe MS Internet Explorer 2.0
- November 96: erster Entwurf für XML
- Dezember 97: HTML 4.0
- Februar 98: XML 1.0

Web-basierte Datenbankanwendungen

- E-Commerce, Shopping-Sites
  - Bestellkataloge
- Online-Banking
  - Kontenzugriffe, Überweisungen, Aktienhandel
- Online-Tracking
  - Verfolgen von Aufträgen
- Online-News, Foren
  - Schlagzeilen, Suche, Anmerkungen
- Web-basierte Geschäftsanwendungen

Anwendungsbeispiel: E-Commerce

- Buchhandel: Amazon.de
  - 4,7 Mill. Titel (Bücher, Musik etc.)
  - 8 Mill. Kunden
  - Suche nach verschiedenen Attributen (Titel, Autor, etc.)
  - Elektronischer Warenkorb
  - Empfehlungen, Kritiken von Kunden

Anwendungsbeispiel: E-Commerce

Anwendungsbeispiel: E-Commerce

Anwendungsbeispiel: Online-Banking

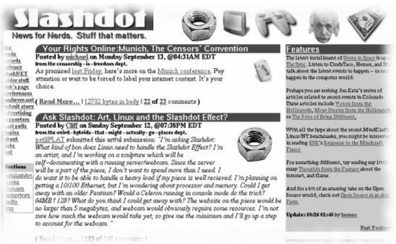
- Deutsche Bank 24
  - Verwaltung von Konten, Aktiendepots über Web
  - Überweisungen, Daueraufträge
  - Verschlüsselte Kommunikation

Anwendungsbeispiel: Online-Banking

Anwendungsbeispiel: Online-Forum

- slashdot.org
  - Aktuelle Informationen zu Software, Hardware, Büchern, Web, ...
  - Benutzer-Annotationen, Diskussionsforum
  - Suche nach Stichwörtern

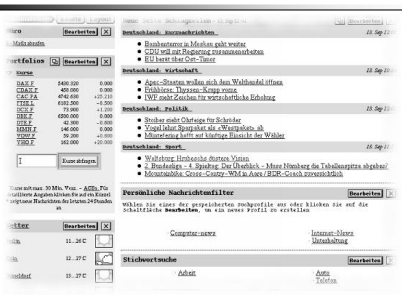
Anwendungsbeispiel: Online-Forum



Anwendungsbeispiel: Portale

- „Einstiegsseite" für Internet-Dienste
- Zugriff auf
  - Dienste (Suchmaschine, Kataloge, Email, Shopping)
  - Nachrichten
- Personalisierte Sichten
  - Layout
  - Inhalte
  - Dienste

Anwendungsbeispiel: Portale



Web-Datenbanken: Kriterien

- Zugriffsart (lesend, schreibend)
- Anzahl gleichzeitiger Zugriffe
- Sitzungslänge (Anfragen, Transaktionen)
- Überlappung der Zugriffe (Caching)
- Datentypen (Text, Multimedia)
- Änderungshäufigkeit
- Benutzeridentifikation, -authentisierung
- Sensitivität der Daten (Banking, Kreditkarteninfos)

Web-Datenbanken: Anforderungen

- Verwaltung großer, multimedialer Datenmengen
- Performance:
  - Hohe Anzahl gleichzeitiger Zugriffe
  - Effiziente Suche
- Integration Web – DBMS
  - Kopplungsmechanismen
  - Benutzerschnittstellen
  - Entwicklungsunterstützung
  - Einbindung von Legacy-Systemen

Web-Datenbanken: Anforderungen

- Sitzungsmanagement
- Sicherheit
- Zugriffsschutz
- Personalisierung von Web-Inhalten

Web-Datenbanken: Szenarien

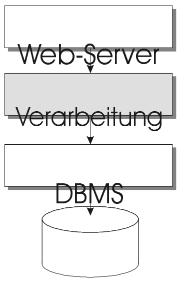
- Informationssystem als Sammlung statischer Seiten
  - Verwaltung von HTML-Seiten im Dateisystem
  - Generierung durch Report-Generatoren
- Verteilungsmechanismus für Anwendungen
  - Verteilung von Anwendungsfunktionen über HTTP (z.B. Applets)
  - Kommunikation mit DBMS über spezielle Protokolle und Schnittstellen (JDBC, CORBA)

Web-Datenbanken: Szenarien

- Web-Schnittstelle zum Datenbanksystem
  - Generierung dynamischer Seiten mit Datenbankinhalten
- Web-Seiten im Datenbanksystem
  - Speicherung von HTML-Seiten in der Datenbank als BLOBs oder mit eigenem Datentyp
- Web-Strukturen im Datenbanksystem
  - Abbildung komplexer Web-Netzwerke in Datenbankstrukturen

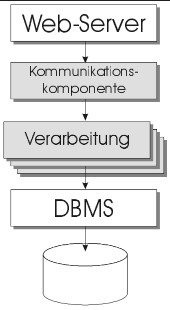
Architekturen: Einstufig

- Direkte Verbindung Web-Server – DBS
- Einfache Realisierung
- Langsamer Verbindungsaufbau
- Für kleinere Lösungen



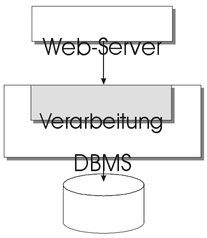
Architekturen: Zweistufig

- Spez. Verarbeitungs-prozesse
- Prozeßpooling
- Lastverteilung möglich
- Höherer Entwicklungsaufwand



Architekturen: Serverseitige Abarbeitung

- Erweiterung des DBS um Verarbeitungs-prozeduren
- Geringer Kommunikations-aufwand
- Belastung des DBS



Web-Datenbanken: Technologien

- Datenmodelle
  - Relationenmodell, objektorientierte Modelle, semistrukturierte Daten
- Anfragesprachen
  - SQL, Multimedia QL
  - XML-QL, XQL
- Präsentation, Datenaustausch
  - HTML, XML, Multimedia

Web-Datenbanken: Technologien

- Kopplungs- und Einbettungstechniken
  - CGI
  - Servlets
  - Einbettung von Skripten
  - Server Side Includes
- Middleware
  - CORBA
  - Applikationsserver

Applikationsserver

- Sammlung von Middleware-Komponenten
  - Object Request Broker, Messaging
  - Web-Server
  - DB-Zugriffskomponenten, Transaktionsmonitore
  - Sicherheitsmechanismen
- „virtuelles" Betriebssystem für Web-Anwendungen, spez. E-Commerce
- Bsp.: Oracle Application Server, IBM WebSphere, Sun NetDynamics, Zope, Midgard

Vorlesung: Zielstellung

- Vermittlung von Kenntnissen zur Entwicklung von Web-DB-Anwendungen
  - Grundlagen Datenbanken, Web
  - Datenbankentwurf und –implementierung
  - Techniken der Kopplung (Web – DBMS)
  - Architekturen
  - Praktische Beispiele

Vorlesung: Inhalt

1. Einleitung und Motivation
2. Grundbegriffe
   - Datenbanken
   - Multimedia
   - World Wide Web
3. Architekturen
   - Komponenten
   - Aufteilung

Vorlesung: Inhalt

4. Datenmodelle und Anfragesprachen
   - Relationenmodell, SQL
   - Multimedia-Anfragesprachen
   - Graphbasierte Modelle, XML
5. Präsentation und Datenaustausch
   - HTML
   - Präsentation und Transformation von XML
   - Multimedia-Präsentation mit SMIL

Vorlesung: Inhalt

6. Kopplungstechniken und Zugriffsschnittstellen
   - Dynamische Web-Seiten
   - CGI, Servlets
   - Einbettung mit PHP
7. Applikationsserver
   - Aufgaben, Techniken
   - Produktbeispiel: Zope

Vorlesung: Inhalt

8. Weitere Aspekte
   - Sicherheit, Sitzungsmanagement
   - Personalisierung
   - Web-Site-Management
9. Zusammenfassung, Ausblick

Vorlesung: Literatur

- Heuer/Saake: Datenbanken, MITP
- Saake/Heuer: Datenbanken – Implementierungstechniken, MITP
- Assfalg: Internet-Datenbanken, Addison-Wesley
- ...

Vorlesung: Organisatorisches

- Zeit/Ort:
  - Mittwoch 8.^{15}-9.^{45} Uhr, Raum 305 (IG)
- Skript:
  - Kopie der Foliensätze
- Prüfung:
  - Klausur
  - Prüfungsersatzleistung
- Aktuelle Informationen:
  - http://wwwiti.cs.uni-magdeburg.de/~sattler/

Hinweise zur Übung

- Ziele: Praktische Arbeit mit Web-Datenbanken
  - Entwurf und Implementierung einer Web-Datenbank
  - Kennenlernen von Methoden und Technologien
- Werkzeuge:
  - DBMS: Sybase
  - Einbettung: PHP (www.php.net, de.php.net)
  - Applikationsserver: Zope (www.zope.org)

Übung: Organisatorisches

- Ort:
  - Raum 224 (IG)
- Zeit:
  - Mittwoch $10.^{15}$-$11.^{45}$ Uhr (14-tägig)
  - 1. Woche: Gruppe 1
  - 2. Woche: Gruppe 2