

Vorlesung Netzbasierte Informationssysteme (WS 2004/05) Nutzung und Betrieb

Robert Tolksdorf
Freie Universität Berlin
Institut für Informatik
Netzbasierte Informationssysteme
mailto:tolk@inf.fu-berlin.de
<http://www.robert-tolksdorf.de>
<http://nbi.inf.fu-berlin.de>

[1] © Robert Tolksdorf, Berlin

Überblick

[2] © Robert Tolksdorf, Berlin

Überblick

- Nutzung und Nutzer von Web-Sites
- Betriebsaspekte sehr grosser Dienste

[3] © Robert Tolksdorf, Berlin

Nutzung und Nutzer von Web-Sites

[4] © Robert Tolksdorf, Berlin

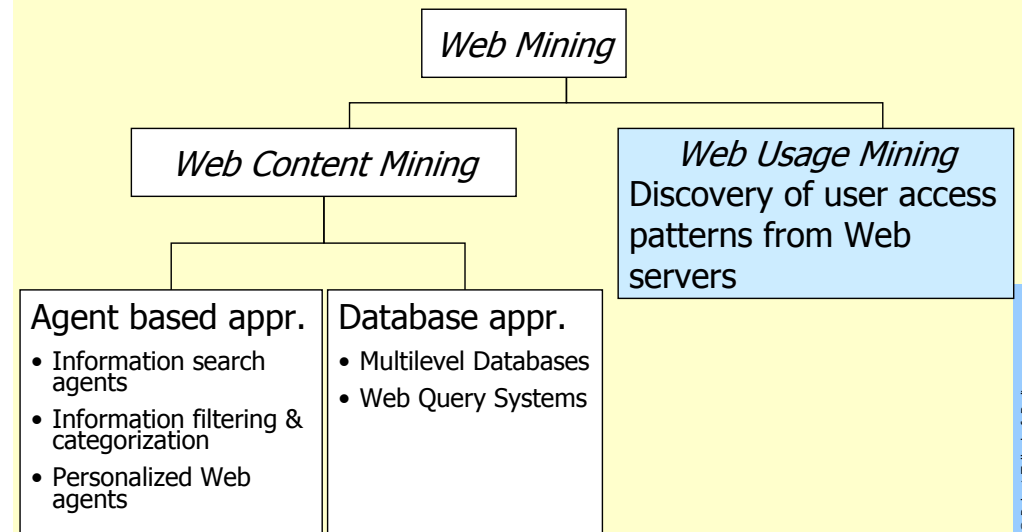
Nutzungs-/Nutzerinformationen

- Nutzer von Web-Sites sind für den Server anonym
 - Keine Identifikation des tatsächlichen Rechners: Proxies, Caches, private Netze, dynamische IP-Nummern
 - Keine Identifikation des Nutzerprozesses: Mehrbenutzerrechner, Proxies, Caches
 - Keine Identifikation des Nutzers: Account-Informationen lokal
- Informationen über Nutzer sind aber nützlich
 - Personalisierung
 - Optimierung des Angebots
 - Grundlage des Geschäftsmodells

[5] © Robert Tolksdorf, Berlin

Web Usage Mining

- *Web Mining*: The discovery and analysis of useful information from the Web



[6] © Robert Tolksdorf, Berlin

Logfiles auf Web-Servern

- Logfiles werden zeilenweise geschrieben
 - Mögliches Format: Common Logfile Format
 - remotehost: IP-Nummer oder Name des Client-Rechners
 - rfc1413: Nutzer-ID auf Quellrechner (ident Dienst)
 - authuser: Nutzer-ID für Web-Session
 - [date]: Datum des Eintrags
 - "request": HTTP-Request Zeile
 - status: HTTP Antwortcode
 - bytes: Größe der Antwort
- ```
--- [19/Dec/2002:10:07:30 +0100] ↵
"GET /~tolk/cool2.gif HTTP/1.1" 200 4942
--- [19/Dec/2002:10:08:06 +0100] ↵
"GET /~tolk/%22http://www.dcs.ed.ac.uk/home/cdw/
phdproject/SECD/Applet/lispkit.html%22 HTTP/1.1" 404 -
```

[7] © Robert Tolksdorf, Berlin

## Logfiles auf Web-Servern

- Extended Common Logfile Format
    - CLF Felder
    - "referer": Seite von der Link verfolgt wurde
    - "user agent": Client-Software
- ```
--- [19/Dec/2002:10:07:30 +0100] ↵
"GET /~tolk/cool2.gif HTTP/1.1" 200 4942 ↵
"http://grunge.cs.tu-berlin.de/~tolk/vmlanguages.html" ↵
"Mozilla/4.0%20(compatible;%20MSIE%206.0;%20windows%20N
T%205.1)"
```

[8] © Robert Tolksdorf, Berlin

Logfiles auf Web-Servern

- Probleme:
 - remotehost:
Nummer des Rechners, der einen Socket zum Server aufbaut ist noch nicht Rechner an dem der Nutzer ist
 - rfc1413:
Läuft ident-Dienst überhaupt? Was soll man mit Ergebnis anfangen?
 - [date]:
Nicht eindeutig bei vielen Zugriffen in kurzen Abständen
 - "request":
GET mit IfModifiedSince-Header, Caches
 - "referer":
Nicht bei Direkteingabe, Bookmarks
 - "user_agent":
Keine zuverlässige Angabe, was ist mit Crawlern?

[9] © Robert Tolksdorf, Berlin

Messgrößen

- Auf Basis von Logfiles lassen sich verschiedene Aussagen über die Nutzung einer Site treffen
- Insbesondere sind diese Aussagen Basis für die Preisfindung der Werbewirtschaft
- Diese Aussagen sind von unterschiedlicher Güte

[10] © Robert Tolksdorf, Berlin

Messgrößen

- Hits
 - Anzahl der Abrufe von Informationen
 - Summe der Anzahl der Requests mit 200 und 304 Antwort
 - Nicht sehr aussagekräftig, weil nicht jede Datei eigenständige Informationseinheit
- Pageviews/Page Impressions
 - Anzahl der abgerufenen HTML-Seiten
 - Anzahl der Hits mit HTML Dateien als Antwort
 - Beschränkt auf einen Medientyp

[11] © Robert Tolksdorf, Berlin

Messgrößen

- 4 Hits, 1 Pageview:
 - - - [19/Dec/2002:12:05:51 +0100]
"GET /~tolk/vmlanguages.html HTTP/1.1" 200 81671
"http://search.msn.com/results.asp?FORM=SCPN&RS=CHECKE D&un=doc&v=1&q=java%20window%20commands"
"Mozilla/4.0%20(compatible;%20MSIE%206.0;%20Windows%20NT%205.1)"
 - - - [19/Dec/2002:12:05:51 +0100]
"GET /~tolk/unclear.gif HTTP/1.1" 200 988
"http://flp.cs.tu-berlin.de/~tolk/vmlanguages.html"
"Mozilla/4.0%20(compatible;%20MSIE%206.0;%20Windows%20NT%205.1)"
 - - - [19/Dec/2002:12:05:51 +0100]
"GET /~tolk/new.gif HTTP/1.1" 200 907
"http://flp.cs.tu-berlin.de/~tolk/vmlanguages.html"
"Mozilla/4.0%20(compatible;%20MSIE%206.0;%20Windows%20NT%205.1)"
 - - - [19/Dec/2002:12:05:51 +0100]
"GET /~tolk/cool2.gif HTTP/1.1" 200 4942
"http://flp.cs.tu-berlin.de/~tolk/vmlanguages.html"
"Mozilla/4.0%20(compatible;%20MSIE%206.0;%20Windows%20NT%205.1)"
- Pageviews und Framesets
 - Erster Abruf des Framesets ist 1 Pageview
 - Jedes Neuladen eines Inhaltsframes ist 1 Pageview
 - Zum Messen immer nur einen Frame neuladen (DMMV)

[12] © Robert Tolksdorf, Berlin

Messgrößen

- Visits / Sessions
 - Zusammenhängende Abrufe in einem Zeitraum
 - Navigationspfade aus Logfile
 - Nicht zuverlässig identifizierbar
 - Problem: Wann ist Visit beendet?
- Heuristiken
 - Zeitorientiert:
 - Gesamtdauer einer Visit ist nach oben begrenzt
 - Verweildauer auf einer Seite ist nach oben begrenzt
 - Navigationsorientiert
 - Topologische Begrenzung: Sitzungsende, wenn Seite nicht von vorherigen Seiten aus erreicht werden konnte
 - Begrenzung durch Referrer: Sitzungsende, wenn Seite nicht durch Navigation von vorheriger Seite erreicht wurde

[13] © Robert Tolksdorf, Berlin

Messgrößen

- Unique Visitors
 - Abrufe von gleicher IP Adressen als 1 Besucher gezählt
 - Objektiv nicht aussagefähig (Proxies, Dynamische IP Adressen, etc.)
- AdImpressions / Clickthroughs
 - Klick auf Werbebanner
 - Messbar beim Werbekunden
 - Quelle durch Referer ermittelbar
 - Abrechnung
 - Preis nach Attraktivität des Werbeträgers: Pageviews und Visits als Maß
 - Preis nach Effizienz des Werbemittels: Clickthroughs als Maß

[14] © Robert Tolksdorf, Berlin

Messgrößen

- Viewtime
 - Dauer des Verweilens auf einem Angebot
 - Kaum aus Logfile messbar
 - Klientenseitige Unterstützung notwendig (z.B. Skripting)
 - Sitzt der Nutzer vor dem Rechner?
- Durch zusätzliche direkte Befragung ermittelbar:
 - Qualified visits: Bestätigte Besuche
 - Regionale Herkunft
 - Alter, Geschlecht etc.
 - Interessen
 - Akzeptanz

[15] © Robert Tolksdorf, Berlin

Wer misst?

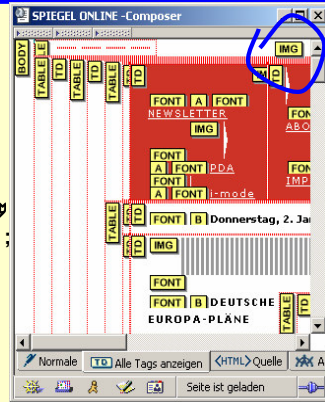
- Serverbetreiber nach eigenem Verfahren und eigener Auswertung
- Serverbetreiber oder Externer nach standardisiertem Verfahren und Auswertung
 - Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V." (IVW) (<http://www.ivwonline.de/>)
 - Messung z.B. durch transparente Grafiken („IVW-Pixel“) auf Seiten
 - ``
 - ``
 - Lösen Messung aus
 - IVW Zahlen sind Grundlage für Preisgestaltung

[16] © Robert Tolksdorf, Berlin

```
<body bgcolor="#ffffff" text="#000000"
link="#b20a15" vlink="#b20a15" alink="#ff0000"
marginheight="0" marginwidth="4" leftmargin="4"
topmargin="0" rightmargin="4" bottommargin="0">
<!-- IWV VERSION="1.2" -->
<script language="JavaScript">
<!--
var IWV="http://spiegel.iwvbox.de/cgi-bin/iwv/CP/
spiegel;/home/c-18/be-PB64-ag9tZXBhZ2UvY2VudGvy";
document.write('<IMG SRC="'+IWV+'?r='+
escape(document.referrer)+'" WIDTH="1,,
HEIGHT="1" BORDER="0" ALIGN="RIGHT">');
// -->
</script>

<noscript>
<IMG SRC="http://spiegel.iwvbox.de/cgi-bin/iwv/CP/spiegel;/home/c-18/be-
PB64-ag9tZXBhZ2UvY2VudGvy" WIDTH="1" HEIGHT="1" BORDER="0" ALIGN="RIGHT">
</noscript>
<!-- /IWV -->
<!-- IWV VERSION="prev" -->

<!-- /IWV -->
```



- Datenaufbereitung
 - Extraktion relevanter Zugriffe, also z.B. nicht Hits auf eingebettete Daten etc. (jpg, map, robots.txt)
 - Hinzufügen verlorengangener Zugriffe
 - Zusammenführen mit Cookie-Informationen
 - Zusammenführen mit Registrierungsinformationen
 - Heuristiken zur Cache Nutzung
- Sitzungsermittlung
 - Ziel: Sequenzen von zusammengehörigen (gleicher Nutzer, gleiche Nutzung) Zugriffen als Sitzung (Session, Visit) identifizieren
 - Problem vergleichbar mit dem Problem der Identifizierung eines Nutzers

- Path analysis
 - Ermittlung von Pfaden in Graphen (oder Graphen aus Pfaden), die Web-Site repräsentieren
 - Link-Struktur einer Site
 - Ähnlichkeitsstruktur von Seiten einer Site
 - Linkverfolgungsstruktur einer Site
 - Weitere Zusammenhänge ermitteln:
 - 70% der Nutzer, die /inst/ag-nbi/lehre/03/S_SW/ zugegriffen haben kamen über den Pfad /inst, /inst/ag-nbi (20% über /lehre/, ...)
 - 5% der Nutzer haben ihren Besuch bei /inst/ag-nbi begonnen
 - 70% der Nutzer haben ihre Sitzung nach einem Pfad der Länge 5 beendet
 - Nutzung dieser Zusammenhänge für die Struktur der Site

- Association rule
 - Ermittlung von Korrelationen zwischen Zugriffen einer Sitzung
 - 30% der Nutzer die /inst/ag-nbi besucht haben, haben auch /inst/ag-tech besucht
 - 2% der Nutzer von /inst/ag-nbi/lehre/0203/V_NBI/ haben sich danach in die Mailingliste eingetragen auf http://lists.spline.inf.fu-berlin.de/mailman/listinfo/nbi_v_nbi
- Sequential pattern
 - Ermittlung von Zusammenhängen zwischen Sitzungen
 - 20% der Nutzer, die sich über http://lists.spline.inf.fu-berlin.de/mailman/listinfo/nbi_v_nbi in eintragen, haben das innerhalb von 10 Tagen über http://lists.spline.inf.fu-berlin.de/mailman/listinfo/nbi_s_xml auch für die andere Mailingliste eingetragen

Analysetechniken

- Classification rules
 - Ermittlung von Profilen von Nutzergruppen
 - 80% derjenigen, die sich unter http://lists.spline.inf.fu-berlin.de/mailman/listinfo/nbi_v_nbi eingetragen haben, studieren Diplom-Informatik
 - Bachelor-Studierende besuchen eher Seiten unter /inst/ag-nbi als unter /inst/ag-bio
- Clustering
 - Gruppierung ähnlicher Nutzer und Daten
 - Interesse an hochpreisiger Consumer-Electronic
 - Nutzung für Marketing und Site-Personalisierung (z.B. amazon)

[21] © Robert Tolksdorf, Berlin

Zusammenfassung

- Ermittlung von Informationen über Nutzer und Nutzung notwendig
- Logfiles als Datenbasis bei Servern, verschiedene Format
- Verschiedene Messgrößen verbreitet
- Ermittlung teilweise sehr schwer
- Web Usage Mining zur Ermittlung komplexerer Zusammenhänge

[22] © Robert Tolksdorf, Berlin

Literatur

- *Common Logfile Format*. <http://www.w3.org/Daemon/User/Config/Logging.html#common-logfile-format>
- Mike StJohns. *Identification Protocol*. Request for Comments 1413. February 1993 <http://www.ietf.org/rfc/rfc1413.txt?number=1413>
- DMMV. *Messgrößen*. http://www.dmmv.de/de/7_pub/homepagedmmv/themen/emarketing/media/zielemedia.cfm
- Bettina Berendt, Bamshad Mobasher, Myra Spiliopoulou, and Jim Wiltshire. *Measuring the accuracy of sessionizers for web usage analysis*. In Workshop on Web Mining at the First SIAM International Conference on Data Mining, pages 7-14, April 2001. <http://maya.cs.depaul.edu/~mobasher/papers/wm-siam01.pdf>
- R. Cooley, B. Mobasher, J Srivastava. *Web Mining: Information and Pattern Discovery on the World Wide Web*. Proceedings of the 9th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'97), November 1997. <http://maya.cs.depaul.edu/~mobasher/papers/webminer-tai97.ps>

[23] © Robert Tolksdorf, Berlin

Betriebsaspekte sehr grosser Dienste

[24] © Robert Tolksdorf, Berlin

Grosse Dienste ("Giant Scale Services")

- Reale große Dienste sind um Dimensionen grösser als beispielsweise www.inf.fu-berlin.de
 - www.cnn.com
 - www.msn.com
 - www.google.com
 - ...
- Werden zu Infrastruktur Diensten
 - imode Gateways
 - email Dienste
 - ...
- [Brewer2001] berichtet über Erfahrungen
 - Eric Brewer (UC Berkeley) ist Mitgründer von Inktomi, Hersteller von (OEM) Suchmaschinentechnologie



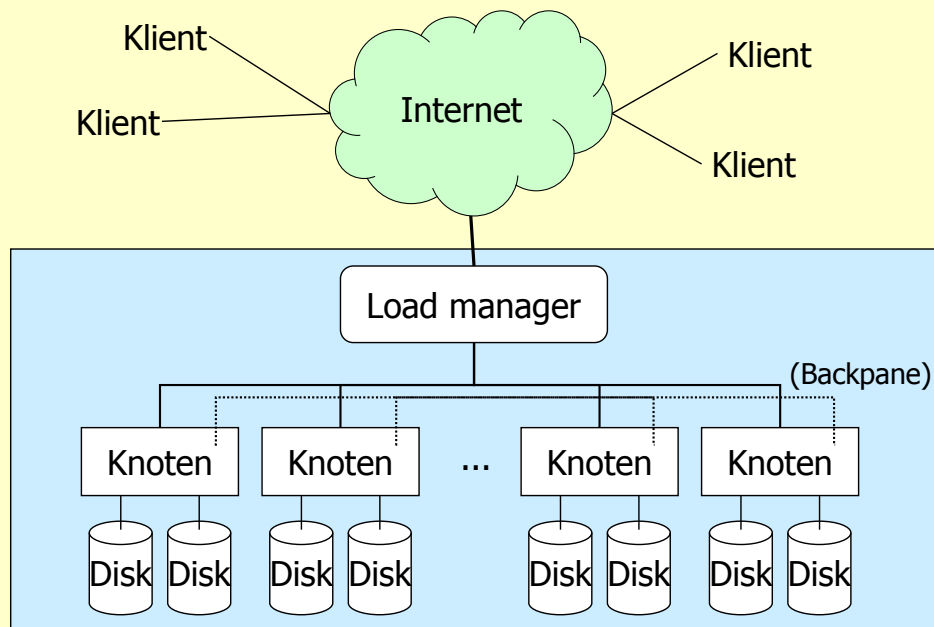
[25] © Robert Tolksdorf, Berlin

Warum grosse Dienste?

- Zugriff von unterschiedlichen Orten (ubiquitäre Dienste)
- Zugriff von unterschiedlichen Geräten aus
- Zusätzliche Dienstmöglichkeiten durch Zentralisierung (z.B. Gruppenkalender)
- Geringere Kosten
 - Infrastruktur wird für mehrere Nutzer genutzt
 - Auslastung der Infrastruktur höher als die einzelner Geräte
 - Einfachere Administrationsstrukturen
- Einfachere Dienstewartung

[26] © Robert Tolksdorf, Berlin

Grundlegendes Modell großer Server



[27] © Robert Tolksdorf, Berlin

Annahmen

- Annahmen
 - Klienten und Netz außerhalb der Kontrolle des Betreibers
 - Dienste werden vom Klienten aus initiiert (HTTP: GET)
 - Erheblich mehr Lesezugriffe als Schreibzugriffe
 - Cluster verwendet
- Datenhaltung
 - *Replikation*
 - Alle Knoten haben komplette Kopie des gleichen Datenbestands
 - *Partitionierung*
 - Knoten haben jeweils einen Teil des kompletten Datenbestandes
 - Alle Knoten zusammen haben kompletten Bestand
 - *Mischformen*

[28] © Robert Tolksdorf, Berlin

Cluster

- *Cluster* (auch: Serverfarm):
Zusammengeschaltete
Standard-Hardware



Service	Nodes	Queries	Nodes
AOL Web cache	>1,000	10B/day	4-CPU DEC 4100s
Inktomi search engine	>1,000	>80M/day	2-CPU Sun Workstations
Geocities	>300	>25M/day	PC Based
Anonymous Web-based e-mail	>5,000	>1B/day	FreeBSD PCs

[29] © Robert Tolksdorf, Berlin

Cluster - Vorteile

- Absolute Skalierbarkeit
 - Notwendig wegen Nutzerwachstum
- Kosten
 - Cluster sind günstig
 - Netz- und Betriebskosten deutlich höher als Hardware-Kosten
- Isolierte Fehler
 - Fehler treten auf einem Rechner des Clusters auf, nicht auf einem zentralen großen Server
- Skalierung ist inkrementell
 - Knoten leicht austauschbar
 - Knoten haben ca. 3-jährige Lebenszeit
 - Ausgetauscht, wenn sie ihren Raumbedarf im Serverschrank nicht mehr rechtfertigen (Moore's Law: vierfache Leistung je Raumbedarf alle 3 Jahren)

[30] © Robert Tolksdorf, Berlin

Load management

- Ziel: Dienste gleichmäßig auf verfügbare Server verteilen
- Erste Ansätze: *Round-robin DNS*
 - DNS: Rechnername -> IP-Nummer
 - Idee: DNS Eintrag zyklisch ändern
 - Name wird zyklisch zu jeweils anderen Servern aufgelöst (IP-Nummer₁, IP-Nummer₂, ..., IP-Nummer_n)
- Nachteile
 - Löst auch nach Servern auf, die gerade nicht erreichbar sind
 - Klienten kann auf gleicher Nummer Versuche wiederholen
 - Kurzes Time-to-live von DNS-Einträgen -> Viele DNS-Lookups
 - Nicht zuverlässiges Klientenverhalten

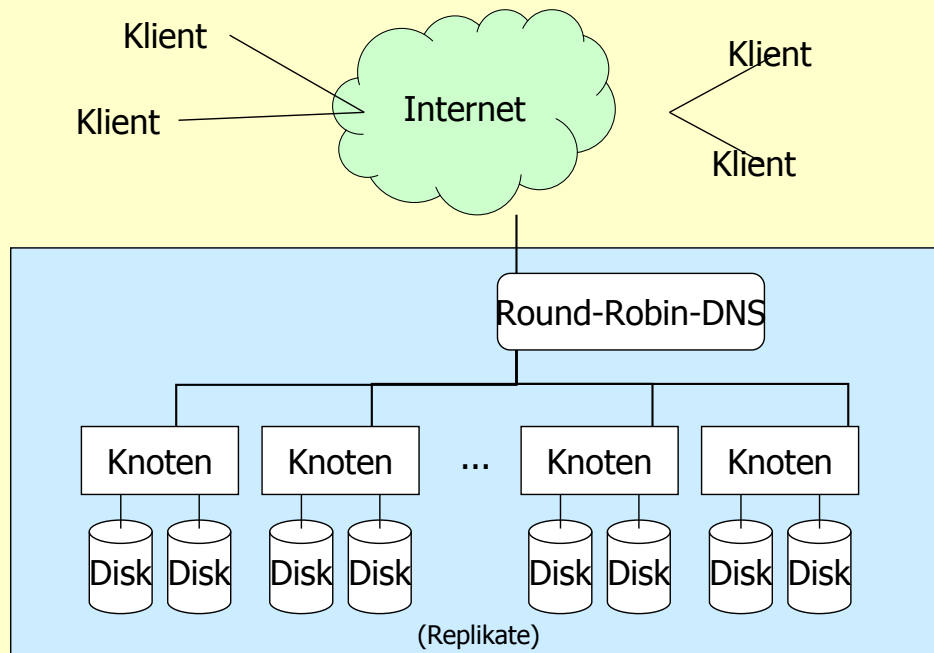
[31] © Robert Tolksdorf, Berlin

Load management

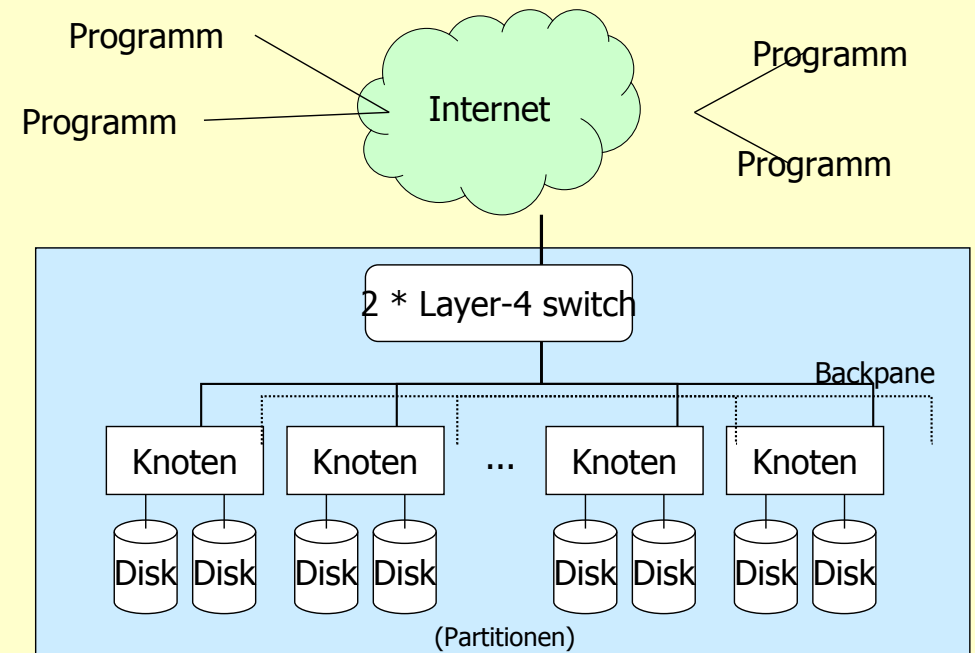
- Layer-4-switches
 - Hardware die TCP "versteht"
 - Switch leitet Pakete aufgrund von TCP-Dienste-Feld und Portnummern an unterschiedliche Server weiter
- Layer-7-switches
 - Hardware die HTTP "versteht"
 - Können URLs mit Netzbandbreite parsieren und leiten Pakete entsprechend weiter
- Meistens als Paar vorhanden
- >20Gbits/s Durchsatz
- Automatisches Monitoring von Knoten

[32] © Robert Tolksdorf, Berlin

Einfache Web-Farm mit Round-Robin-DNS



Inktomi Suchdienst



Verfügbarkeit

- Zentrale Anforderung an grosse Dienste:

Verfügbarkeit (Availability)

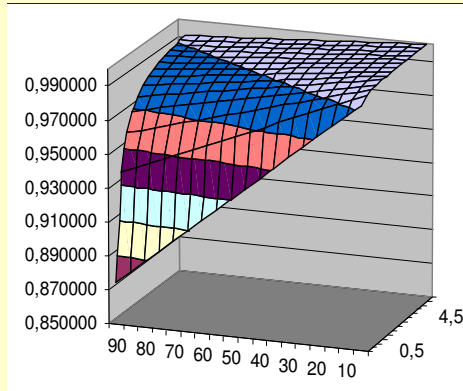
- Gemessen in "Neunern":
 - Vier Neuner: 0,9999 Verfügbarkeit (<60 Sek. Ausfall/Woche)
 - Fünf Neuner: 0,99999
- Ähnlich geleitete Systeme:
 - Telefonsystem
 - Zugverkehr
 - Wasserversorgung

Weitere Maße

- MTBF: Mean-time-between-failures, mittlerer zeitlicher Abstand zwischen Fehlern
- MTTR: Mean-time-to-repair, mittlere Dauer der Fehlerbehebung
- $$\text{uptime} = \frac{(\text{MTBF} - \text{MTTR})}{\text{MTBF}}$$
- MTBF = 6 Tage (518400s), MTTR = 20 Min (1200s)
$$\text{uptime} = \frac{(518400\text{s} - 1200\text{s})}{518400\text{s}} = 0,997685$$

Weitere Maße

- Uptime erhöhen durch:
 - Erhöhung der MTBF
 - MTTR senken
- MTBF = 7 Tage
MTTR = 20 Min
uptime=0,998016
- MTBF = 6 Tage
MTTR = 10 Min
uptime= 0,998843



- MTBF schwerer zu ermitteln, MTTR klarerer Prozess
- MTTR stabiler auch bei neuen Komponenten
- Fokus auf MTTR, Best-Effort bei MTBF

[37] © Robert Tolksdorf, Berlin

Yield und Harvest

- Weiteres Maß: Yield – vielen Anfrageergebnisse?
 - $$\text{yield} = \frac{\text{bearbeitete Anfragen}}{\text{gestellte Anfragen}}$$
 - entspricht Nutzererfahrung
 - gewichtet Uptime-Sekunden
- Weiteres Maß: Harvest – welcher Teil der Datenbank ist nutzbar?
 - $$\text{harvest} = \frac{\text{zugreifbare Daten}}{\text{gesamte Daten}}$$
 - Erweiterbar zum Anteil der nutzbaren Dienste

[38] © Robert Tolksdorf, Berlin

Yield und Harvest

- Ziel: 100% Yield und 100% Harvest
- Anpassbar:
 - Fehler bei Replikaten
 - Sinkendes Yield
 - Gleichbleibendes Harvest
 - Fehler bei Partitionen
 - sinkendes Harvest
 - Gleichbleibendes Yield

[39] © Robert Tolksdorf, Berlin

DQ Prinzip

- Daten pro Anfrage (D) × Anfragen pro Sekunde (Q) → C
- Reflektiert physikalische Beschränkung des Gesamtsystems
 - I/O Bandbreite
 - Plattengeschwindigkeit
 - ...
- Ist bandbreitenorientiert
- Höheres DQ bewirken
 - Mehr Knoten (linear, also Veränderungen leicht testbar)
 - Software-Optimierungen
- Niedrigeres DQ bewirken
 - Fehler
- Relatives DQ entscheidend
 - Auswirkung von Fehlern ermittelbar
 - Zielvorgaben für Optimierungen/Ausbau ermittelbar

[40] © Robert Tolksdorf, Berlin

Replikation vs. Partitionierung

- Fehleranalyse: 1 Ausfall bei 2 Knoten
 - Replikation
 - Harvest bleibt bei 100%
 - Yield fällt auf 50%
 - D bleibt gleich, Q halbiert sich
 - Partitionierung
 - Harvest fällt auf 50%
 - Yield bleibt bei 100%
 - D halbiert sich, Q bleibt gleich
- Replikation und Partitionierung sind nach DQ Prinzip gleich

[41] © Robert Tolksdorf, Berlin

Replikation vs. Partitionierung

- Replikation benötigt zusätzlich DQ Kapazität zum Zugriff auf Replikat
- Partitionierung aber auch kein Vorteil: Keine Ersparnis durch Einsparung von Datenplatz, weil DQ eigentlicher Kostenfaktor ist
- Nach anfänglicher Partitionierung Übergang zu mehr Replikation
 - Harvest steuerbar
 - Datenverlustwahrscheinlichkeit steuerbar
 - Skalierung leichter handhabbar

[42] © Robert Tolksdorf, Berlin

Kontrollierter Leistungsrückgang

- Graceful degradation:
Bei hoher Last Verfügbarkeit erhalten
- Daten:
 - Hohe Last zu Normallast zwischen 1.6:1 und 6:1
 - Vorhaltung von Leistung zu teuer
 - z.B. SMS an Silvester
 - Extremelast durch einzelne Ereignisse
 - Vorgehaltene Leistung immer noch zu klein
 - z.B. Kartenreservierung Star Wars
 - Zusammenhängende Fehler
 - z.B. Stromausfall

[43] © Robert Tolksdorf, Berlin

cnn.com 9.11.01 und 2.1.03: 28757 Bytes vs. 280133 Bytes



The screenshot shows the CNN.com homepage on September 11, 2001. The main headline is "AMERICA UNDER ATTACK" with a sub-headline "DEVELOPING STORY". Below the headline is a photograph of people running in a city street. The text below the photo reads: "At 9:36 a.m. EDT, the first of two airliners crashed into the World Trade Center, opening a horrifying and apparently coordinated terrorist attack on the United States, which saw the collapse of the two 110-story towers into surrounding Manhattan streets and a later attack on the Pentagon." Below this is a "COMPLETE COVERAGE:" section with a list of links: "Chronology of terror", "President Bush: U.S. will 'hunt down and punish those responsible'", "World Trade Centers collapse after planes hit, 10,000 emergency workers head to scene", "Plane hits Pentagon, part of the Pentagon collapses", "American, United both confirm losing two planes each", "Federal buildings, United Nations evacuated", "FAA grounds all U.S. flights, sends trans-Atlantic flights to Canada", "Sen. John McCain calls attacks 'act of war'", "Israel evacuates embassies", "Non-essential NATO employees asked to leave Brussels HQ", "Taliban issues statement to tell U.S. 'Afghanistan feels your pain'", "Kennedy Space Center, LAX, Disney Florida parks closed", and "World shook over U.S. attacks".



The screenshot shows the CNN.com homepage on February 1, 2003. The main headline is "Edwards in 2004?" with a sub-headline "FULL STORY". Below the headline is a photograph of John Edwards speaking. The text below the photo reads: "Democratic Sen. John Edwards said today that he's taking the first step toward a 2004 bid for the White House by forming an exploratory committee. The North Carolina lawmaker joins Vermont Gov. Howard Dean and Massachusetts Sen. John Kerry as likely Democratic candidates." Below this is a "FULL STORY" section with a link to "Video: Profile of Sen. Edwards" and another link to "Interactive: Democratic contenders".

[44] © Robert Tolksdorf, Berlin

DQ Prinzip z. Modellierung kontrollierten Leistungsrückgangs

- D (Daten pro Anfrage) verkleinern um Q (Anfragen pro Sekunde) zu erhöhen
 - Datenbank verkleinern
 - Dadurch Yield erhalten
 - siehe CNN am 9.11.01
- Q limitieren um D zu erhalten
 - Durch Zugangsbeschränkungen Q senken
 - Anfragekosten schätzen, aufwändige verbieten
 - Bestimmte Anfragen priorisieren
 - Anfragekosten durch mindere Datenqualität senken
 - Dadurch Harvest erhöhen

[45] © Robert Tolksdorf, Berlin

Upgrades / Wartung

- Grosse Dienste müssen fortlaufend gewartet werden
 - Software-Erweiterungen/-Korrekturen
 - Hardware-Erweiterungen/-Ersatz
- Wartung ist kontrollierter Ausfall
- Verlust an DQ Kapazität:
$$\Delta DQ = n * u * \overline{DQ/Knoten} = DQ * u$$

[46] © Robert Tolksdorf, Berlin

Upgrades / Wartung

- Schneller Neustart
 - Maschinen werden in neue Konfiguration gebootet
 - Yield geht verloren
 - Optimierte durch geeigneten Zeitpunkt (off-peak)
- Rolling Upgrade
 - Jeweils ein Knoten upgraden
 - Replikation: Yield sinkt minimal
 - Partitionierung: Harvest sinkt
 - Zwei Versionen müssen verträglich koexistieren
- „Big Flip“
 - Halber Cluster wird neu gestartet, danach andere Hälfte
 - Durch Layer-4 Switch transparent nach aussen
 - 50% DQ Verlust

[47] © Robert Tolksdorf, Berlin

Zusammenfassung

- Sehr große Dienste sind durch sehr große Cluster realisiert
- Lastverteilung durch spezielle Switches
- Verfügbarkeit zentral
- Bewegte Daten pro Sekunde als Obergrenze von Systemen (DQ Prinzip)
- DQ Prinzip: Replikation besser als Partitionierung
- DQ Prinzip zur Modellierung von Ausnahmemanagement
- Upgrades unterschiedlich handhabbar

[48] © Robert Tolksdorf, Berlin

Literatur

- Eric A. Brewer. *Lessons from Giant-Scale Services*. IEEE Internet Computing. July/August 2001. 46-55.
- Nils Peters. *"The Internet under Attack"*
<http://520066654038-0001.bei.t-online.de>
- Matrix.Net. *Internet Withstands Attack on America*. Presseerklärung. http://520066654038-0001.bei.t-online.de/webmirror/Internet%20Withstands%20Attack%20on%20America%20--%20Matrix_%20Net.htm

Zusammenfassung

Zusammenfassung

- Nutzung und Nutzer von Web-Sites
- Betriebsaspekte sehr grosser Dienste