

Replikation und Caching für mobile Anwendungen

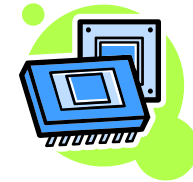
Seminar: Mobile and Context-aware Database Technologies and Applications
Lehrgebiet Informationssysteme
Freitag, den 13.07.2006
Christian Doppstadt

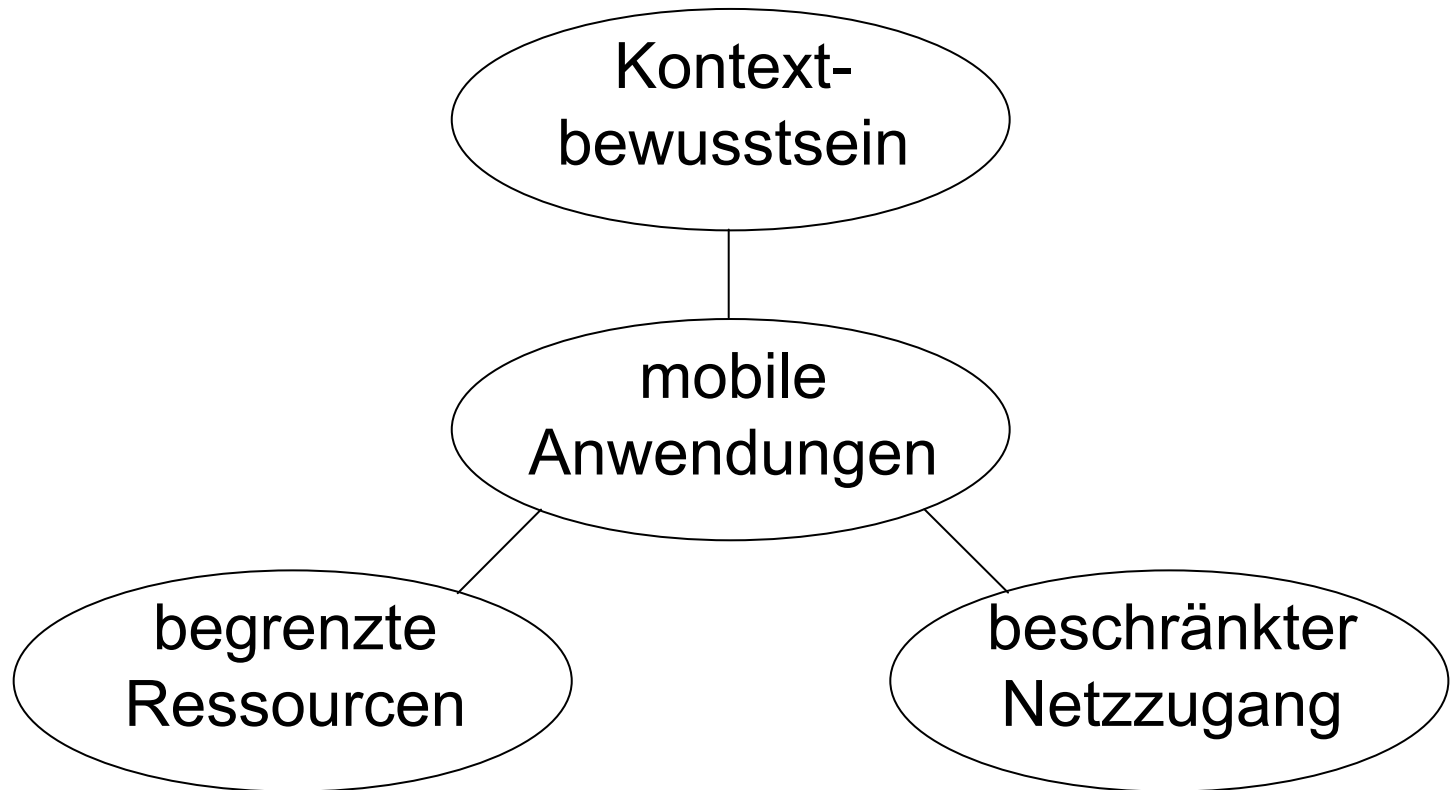
- Motivation
- Caching und Replikation
- Verfahren
- Fazit

- **Motivation**
 - **Caching und Replikation im Alltag**
 - **Einordnung**
 - **Anforderungen**
- Caching und Replikation
- Verfahren
- Fazit

Caching und Replikation im Alltag

- Caching
 - Proxy-Server
 - CPU-Cache
 - Festplatten-Cache
- Replikation
 - Datensicherung (Backup)
 - MP3-Player
 - Verteilte Systeme





- Vorhalten von Daten
 - Minimierung des Übertragungsvolumens
 - Ressourcen des mobilen Endgerätes schonen
- ➔ Nutzung einer mobilen Anwendung trotz eingeschränkter Ressourcen und beschränktem Netzzugang

- Motivation
- **Caching und Replikation**
 - **Caching**
 - **Replikation**
 - **Abgrenzung**
 - **Einschränkungen**
 - **Exkurs: Synchronisation**
- Verfahren
- Fazit

- Einsatz einer mehrschichtigen Speicherstruktur
- Zugriff erfolgt über die höchste Stufe

Ziele

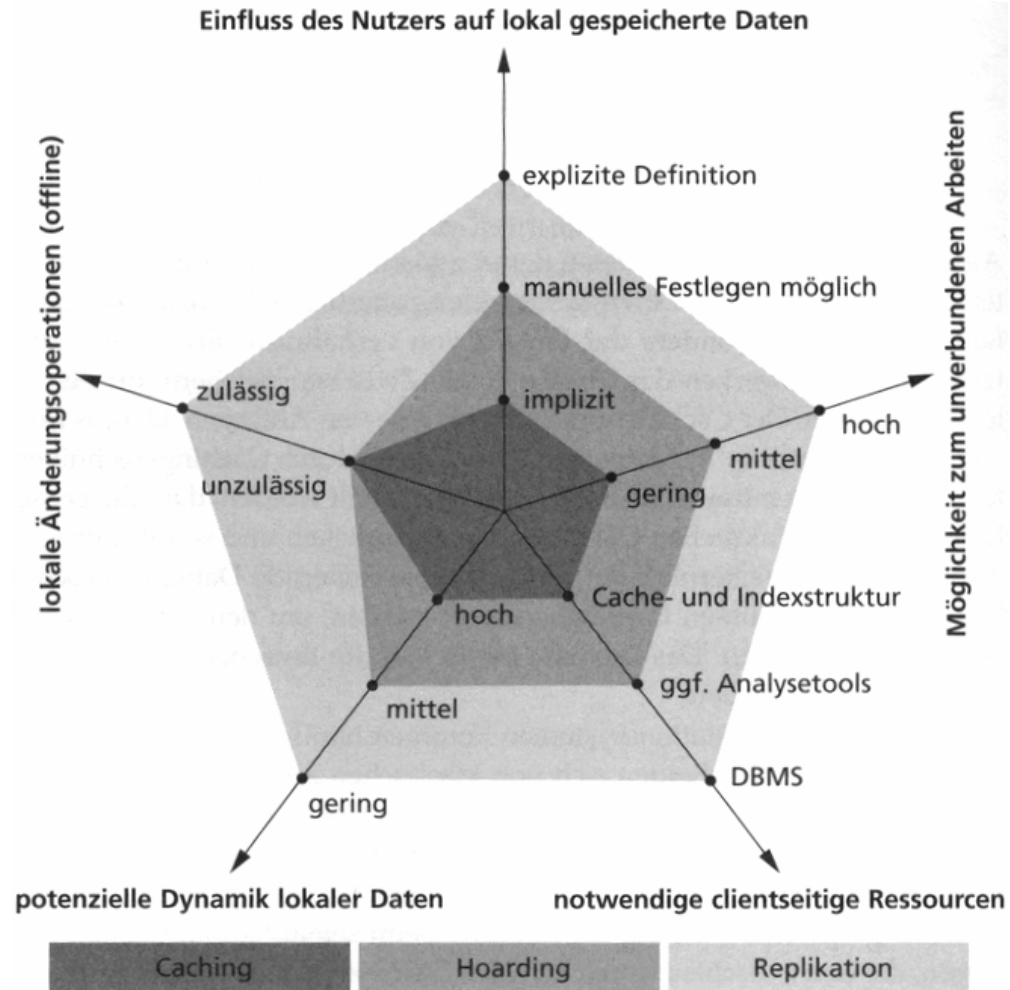
- Beschleunigung des Zugriffs auf die unterste Speicherebene
- Reduzierung der zwischen den Ebenen zu transportierenden Daten

- Erzeugung einer physikalischen Kopie der Ausgangsdaten
- Die ursprünglichen Daten werden nach dem Kopieren weiter genutzt

Ziele

- Erhöhung der Verfügbarkeit
 - ➔ Höhere Leseraten
 - ➔ Sinkende Netzauslastung

Abgrenzung



Einschränkungen

- Keine Garantie gesuchte Information im Cache zu finden
- Replikation aller benötigten Daten unmöglich
- Schreibende Zugriffe erfordern Synchronisation

- Synchrone Replikation
- Asynchrone Replikation
 - Konfliktvermeidende Replikation
 - Konfliktauflösende Replikation



- Motivation
- Caching und Replikation
- **Verfahren**
 - **Check-out/Check-in**
 - **Escrow-Verfahren**
 - **Änderung der Granularität**
 - **Semantisches Caching**
 - **Key-Pool-Verfahren**
 - **Optimistische Synchronisation**
- Fazit

- Versicherungsvertreter
- Tätigkeit im Außendienst
- Kundendaten liegen in einer Datenbank auf dem Firmenserver



Aufgabe 1:

- Besuch eines Kunden
- Überprüfung der Stammdaten
- Kein Netzzugang beim Kunden



Check-out/Check-in (I)

- Langzeit-Lese-/Schreibsperren
- Zusätzliche Check-in-Sperren

	R	W	LR	LW	C
R	+	-	+	+	-
W	-	-	-	-	-
LR	+	-	+	-	-
LW	+	-	-	-	-
C	-	-	-	-	-

PRO

- Gewährleistet Serialisierbarkeit
- Primärdaten können gelesen werden

CONTRA

- Sperrprotokoll
- Primärdaten können nicht verändert werden

Aufgabe 2:

- Verkauf einer neuen Versicherung
- Nur begrenzte Anzahl verfügbar



Escrow-Verfahren (I)

- Zugriff auf Hot-Spot Objekte
- „Reservierung“ von Ressourcen

T1	T2	INF	Q	SUP
		10	10	10
-5		5	5	10
	-3	2	2	10
	COMMIT	2	2	7
ABORT		7	7	7

PRO

- Zugriff ohne Sperren
- Garantiert Erfüllbarkeit

CONTRA

- Nur zum In- und Dekrementieren geeignet
- Gefahr zu hoher „Reservierung“

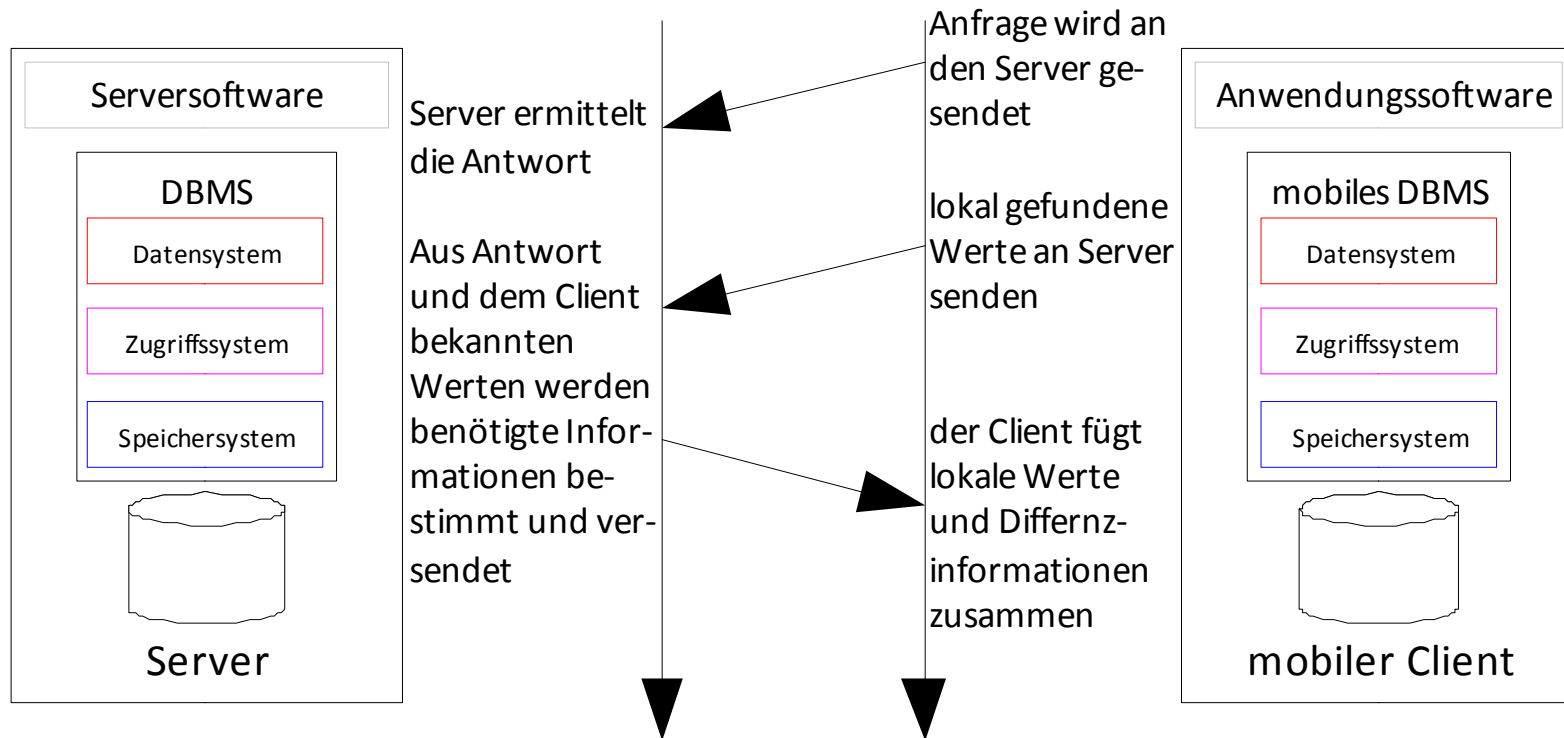
Aufgabe 3:

- Ungeplanter Besuch bei einer Kundin in der selben Gegend
- Zusätzliche benötigte Vektorgrafiken für die Anfahrt herunterladen
- Netzzugang über das Handy möglich



Änderung der Granularität (I)

- Von Seiten zu Attributen oder Objekten



Änderung der Granularität (II)

PRO

- Höhere Hit-Rate als bei Seiten-Caching
- Geringere Netzwerklast

CONTRA

- Netzverbindung zwingend notwendig
- Cache-Verwaltung wird komplexer

Aufgabe 4:

- Durch die Planänderung ist ein neues Restaurant für die Mittagspause zu suchen
- Das Restaurant soll nicht zu teuer sein



Semantisches Caching (I)

- Semantische Indexierung
- Beschreibung des Cache-Inhaltes mit Prädikaten
- Aus Anfrage und Cache-Beschreibung werden Kompensations- und Filteranfrage bestimmt

Datenbasis Restaurant

<u>Name</u>	Adresse	PLZ	Ø Preis
Bürgershof	Lintorfer Markt 24	67650	14
Cedric's	Brachter Str. 85	67656	17
Haus Langhardt	Kaiserswerther Str. 121	67656	13
Schlüssel am Markt	Düsseldorfer Str. 1	67657	12

Cache-Inhalt

<u>Name</u>	Adresse	PLZ
Bürgershof	Lintorfer Markt 24	67650
Cedric's	Brachter Str. 85	67656
Haus Langhardt	Kaiserswerther Str. 121	67656

Semantisches Caching (III)

Q = ({Name, Adresse, Ø Preis},
{(PLZ ≥ 67656), (Ø Preis < 15)}, Restaurant)

Datenbasis Restaurant

<u>Name</u>	Adresse	PLZ	Ø Preis
Bürgershof	Lintorfer Markt 24	67650	14
Cedric's	Brachter Str. 85	67656	17
Haus Langhardt	Kaiserswerther Str. 121	67656	13
Schlüssel am Markt	Düsseldorfer Str. 1	67657	12

Semantisches Caching (IV)

Q = ({Name, Adresse, Ø Preis},
{(PLZ ≥ 67656), (Ø Preis < 15)}, Restaurant)

C = ({Name, Adresse, PLZ},
{PLZ ≤ 67656}, Restaurant)

Cache-Inhalt

<u>Name</u>	Adresse	PLZ
Bürgershof	Lintorfer Markt 24	67650
Cedric's	Brachter Str. 85	67656
Haus Langhardt	Kaiserswerther Str. 121	67656

Semantisches Caching (V)

- Q = ({Name, Adresse, Ø Preis},
{(PLZ ≥ 67656), (Ø Preis < 15)}, Restaurant)
- C = ({Name, Adresse, PLZ},
{PLZ ≤ 67656}, Restaurant)
- Q^F = ({Name, Adresse},
{PLZ ≥ 67656}, Restaurant)

Q^F

<u>Name</u>	Adresse
Cedric's	Brachter Str. 85
Haus Langhardt	Kaiserswerther Str. 121

Semantisches Caching (VI)

- Q = ({Name, Adresse, Ø Preis},
{(PLZ ≥ 67656), (Ø Preis < 15)}, Restaurant)
- C = ({Name, Adresse, PLZ},
{PLZ ≤ 67656}, Restaurant)
- Q^F = ({Name, Adresse},
{PLZ ≥ 67656}, Restaurant)
- Q^{K1} = ({Name, Ø Preis},
{(PLZ ≤ 67656), (PLZ ≥ 67656), (Ø Preis < 15)},
Restaurant)

Q^{K1}

<u>Name</u>	Ø Preis
Haus Langhardt	13

Semantisches Caching (VII)

- $Q =$ ($\{\text{Name, Adresse, } \emptyset \text{ Preis}\}$,
 $\{(\text{PLZ} \geq 67656), (\emptyset \text{ Preis} < 15)\}$, Restaurant)
- $C =$ ($\{\text{Name, Adresse, PLZ}\}$,
 $\{\text{PLZ} \leq 67656\}$, Restaurant)
- $Q^F =$ ($\{\text{Name, Adresse}\}$,
 $\{\text{PLZ} \geq 67656\}$, Restaurant)
- $Q^{K1} =$ ($\{\text{Name, } \emptyset \text{ Preis}\}$,
 $\{(\text{PLZ} \leq 67656), (\text{PLZ} \geq 67656)\}$, Restaurant)
- $Q^{K2} =$ ($\{\text{Name, Adresse, } \emptyset \text{ Preis}\}$, $\{\neg(\text{PLZ} \leq 67656),$
 $(\text{PLZ} \geq 67656), (\emptyset \text{ Preis} < 15)\}$, Restaurant)
- Q^{K2}

<u>Name</u>	Adresse	\emptyset Preis
Schlüssel am Markt	Düsseldorfer Str. 1	12

Semantisches Caching (VIII)

Q^{K2}

<u>Name</u>	Adresse	Ø Preis
Schlüssel am Markt	Düsseldorfer Str. 1	12

Q^{K1}

<u>Name</u>	Ø Preis
Haus Langhardt	13

Q^F

<u>Name</u>	Adresse
Cedric's	Brachter Str. 85
Haus Langhardt	Kaiserswerther Str. 121

Q

<u>Name</u>	Adresse	Ø Preis
Cedric's	Brachter Str. 85	-
Haus Langhardt	Kaiserswerther Str. 121	13
Schlüssel am Markt	Düsseldorfer Str. 1	12

PRO

- Anfragen oder Teilanfragen können ohne Netzzugang beantwortet werden
- Genaues Wissen über den Cache-Inhalt ermöglicht gezielte Verdrängung

CONTRA

- Sehr hoher Rechenaufwand für den Prädikatvergleich
- Problem der Konsistenzsicherung

Aufgabe 5:

- Abschluss einer neuen Versicherung mit der Kundin
- Eintragen der benötigten Daten
- Kein Netzzugang verfügbar



Key-Pool-Verfahren (I)

- Reservierung von Primärschlüsseln beim zentralen DBMS
- Einfügen neuer Datensätze zu diesen Primärschlüsseln
- Neue Datensätze ins zentrale DBMS übernehmen

PRO

- Zugriff ohne Sperren
- Garantiert Erfüllbarkeit

CONTRA

- Nur zum Einfügen neuer Datensätze geeignet
- Verwendung von synthetischen Primärschlüsseln

Aufgabe 6:

- Bei einer bestehenden Kfz-Versicherung sollen Anpassungen vorgenommen werden



Optimistische Synchronisation (I)

- Änderungen werden auf ein Replikat ausgeführt
- Überprüfung auf Konflikt ggf. Zurücksetzen der Transaktion
- Änderungen auf die Ausgangsdaten übernehmen



PRO

- Keine Behinderung von Transaktionen auf dem Server
- Beliebige Operationen möglich

CONTRA

- Globales COMMIT erst bei der Synchronisation möglich
- Keine Erfolgsgarantie

- Motivation
- Caching und Replikation
- Verfahren
- **Fazit**

- Caching kann den Datenverkehr bei mobilen Anwendungen erheblich verringern und das Netzwerk entlasten
 - Replikation ermöglicht das Arbeiten mit mobilen Anwendungen ohne Netzwerkverbindung
- ➔ Arbeiten mit eingeschränkter Konnektivität wird ermöglicht



ABER: Mit steigender Leistungsfähigkeit der Technik nimmt auch die Komplexität für den Anwender zu

→ Notwendigkeit zur Vereinfachung der Technik für den Endbenutzer als Schwerpunkt für zukünftige Entwicklung

- Motivation
 - Caching und Replikation im Alltag
 - Einordnung
 - Anforderungen
- Caching und Replikation
 - Caching
 - Replikation
 - Abgrenzung
 - Einschränkungen
 - Exkurs: Synchronisation
- Verfahren
 - Check-out/Check-in
 - Escrow-Verfahren
 - Änderung der Granularität
 - Semantisches Caching
 - Key-Pool-Verfahren
 - Optimistische Synchronisation
- Fazit