



Replikation in einem homogenen strukturierten Chord Peer-to-Peer Netz

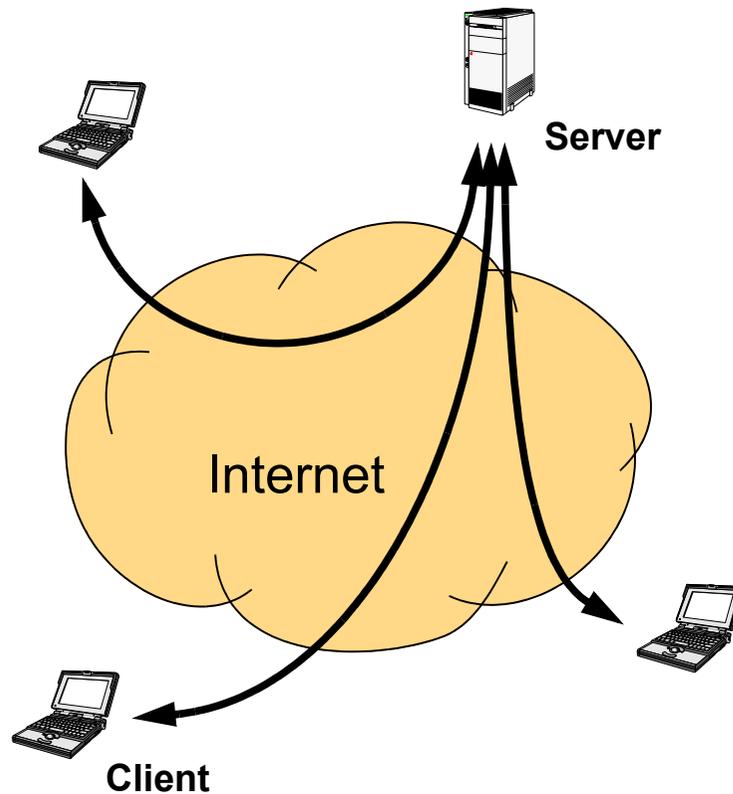
VFF IND/IKR-Workshop

Andreas Reifert, Michael Miksch
reifert@ikr.uni-stuttgart.de

3. März 2006

- Chord
- Replikationsmechanismen
- Untersuchung und Bewertung

Klassisches Client-Server Konzept



Anwendungen

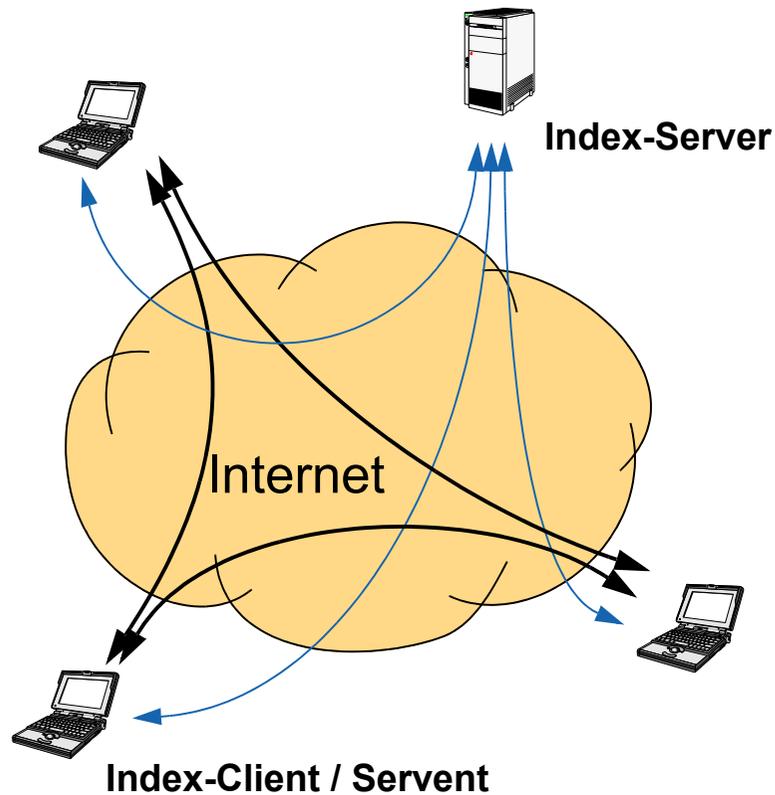
- Web Server
- Mail Server
- File Server

Server

- (Index-) Anfrage
- Daten

Peer-to-Peer Netze

Zentralisiertes Peer-to-Peer Konzept



z.B. Napster

zentralisiert

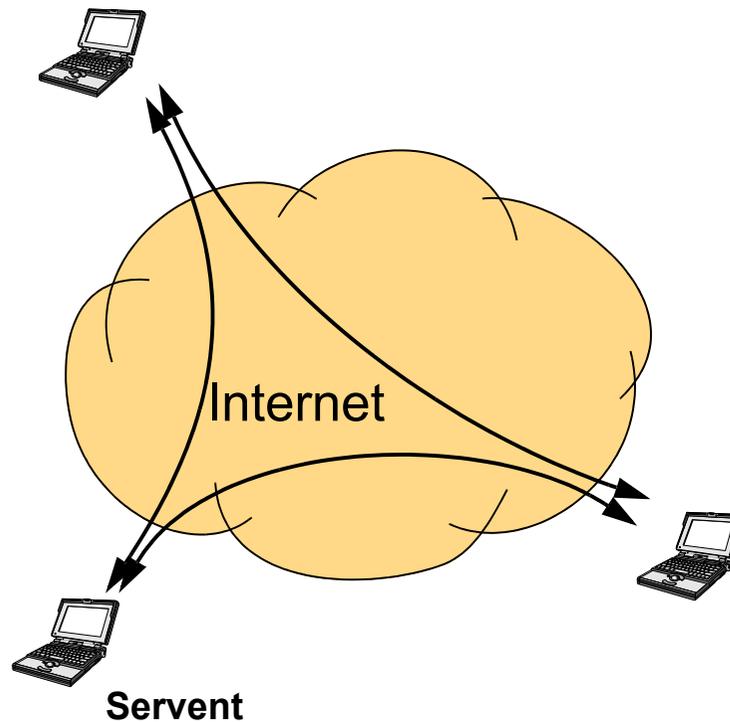
Server

(Index-) Anfrage

Peers

Daten

Dezentralisiertes Peer-to-Peer Konzept



z.B. Chord

dezentralisiert

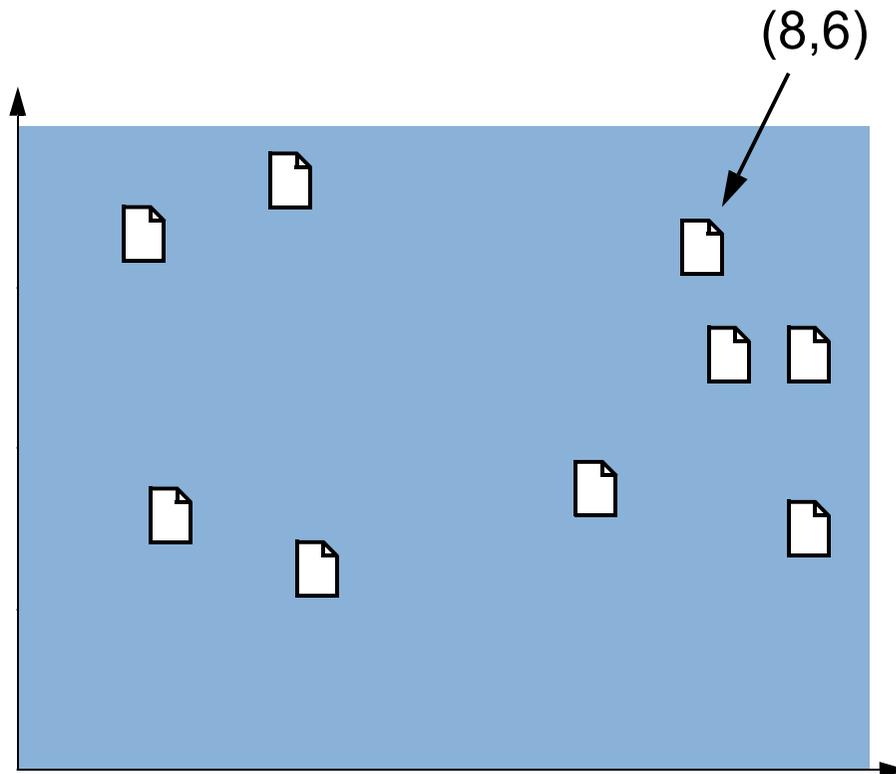
Peers

- (Index-) Anfrage
- Daten

Anwendungen

- File Sharing
- Datenbanken
- Datei-/Speichersysteme

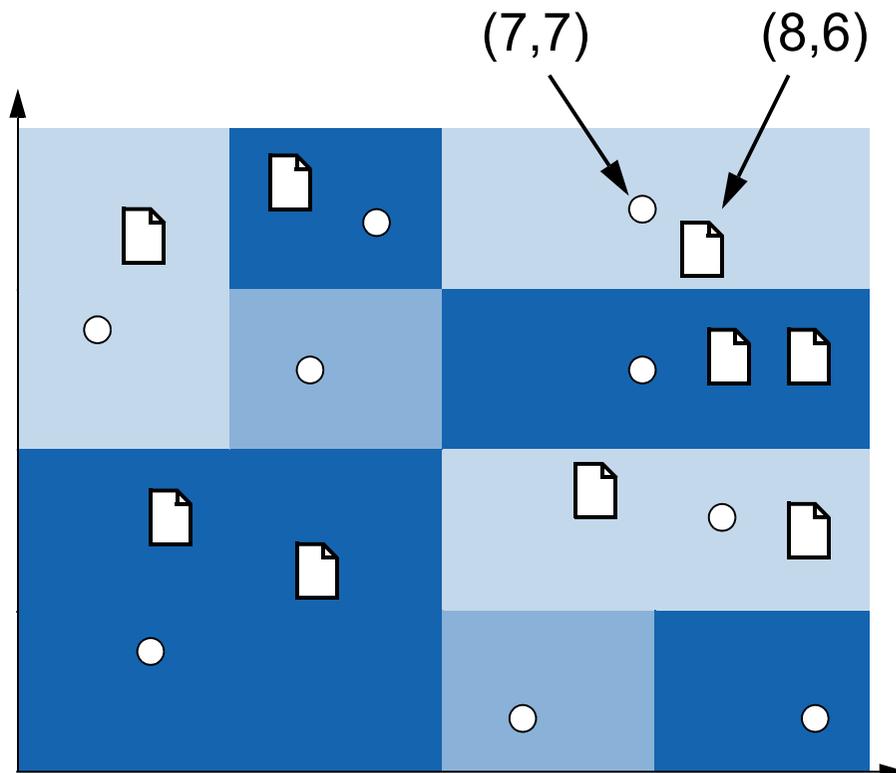
Organisation der Anfrageinformationen / Dokumente



- **Auffinden von Dokumenten**
 - Broadcast
 - Dokumentenbasiertes Routing
- **Dokumente haben eindeutigen Schlüssel (Key)**
Dokumentenraum

Strukturierte Peer-to-Peer Netze

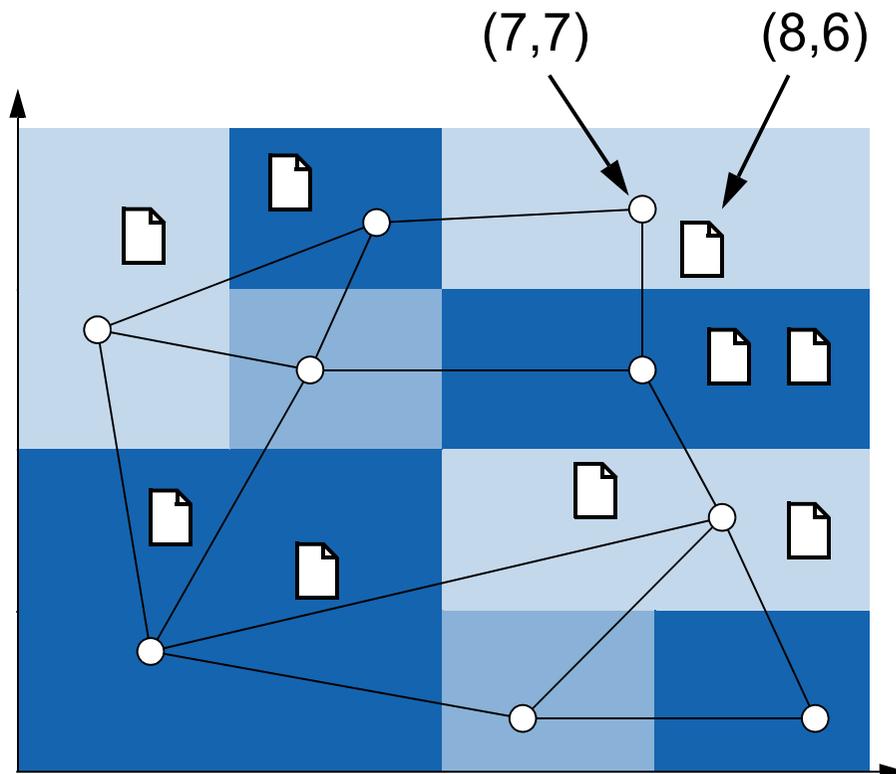
Organisation der Anfrageinformationen / Dokumente



- **Auffinden von Dokumenten**
 - Broadcast
 - Dokumentenbasiertes Routing
- **Dokumente haben eindeutigen Schlüssel (Key)**
Dokumentenraum
- **Partitionierung des Dokumentenraums**
 - Zuordnung zu Peers
 - Peer hat eindeutige Schlüssel aus seinem Bereich (ID)

Strukturierte Peer-to-Peer Netze

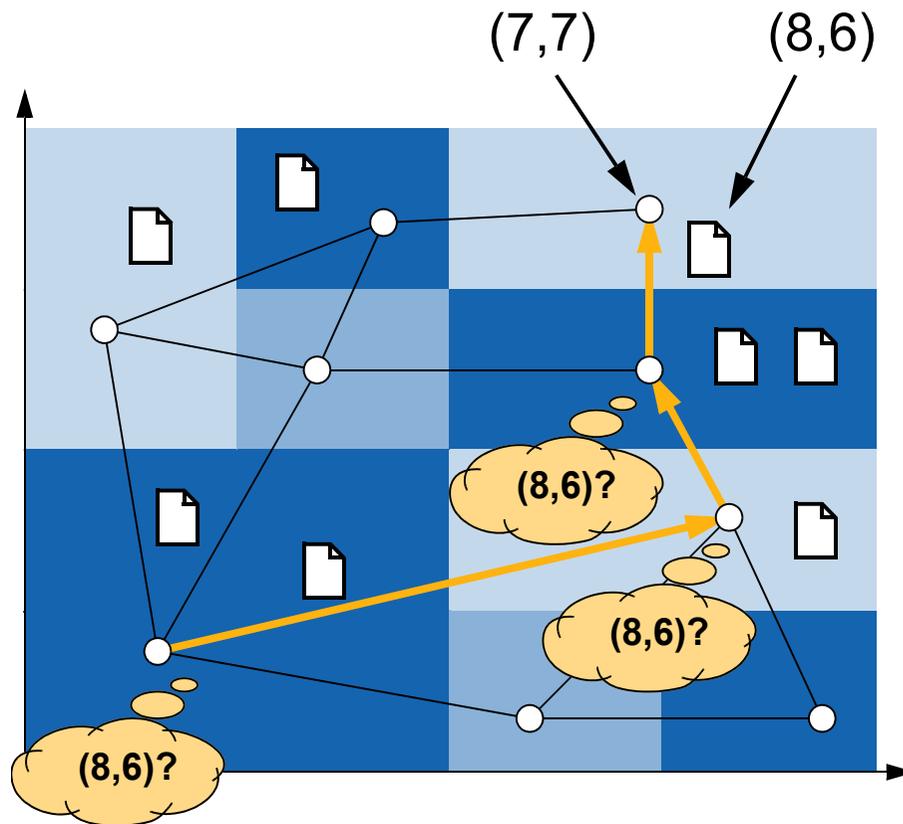
Organisation der Anfrageinformationen / Dokumente



- **Auffinden von Dokumenten**
 - Broadcast
 - Dokumentenbasiertes Routing
- **Dokumente haben eindeutigen Schlüssel (Key)**
Dokumentenraum
- **Partitionierung des Dokumentenraums**
 - Zuordnung zu Peers
 - Peer hat eindeutige Schlüssel aus seinem Bereich (ID)

Strukturierte Peer-to-Peer Netze

Organisation der Anfrageinformationen / Dokumente



- **Auffinden von Dokumenten**
 - Broadcast
 - Dokumentenbasiertes Routing
- **Dokumente haben eindeutigen Schlüssel (Key)**
Dokumentenraum
- **Partitionierung des Dokumentenraums**
 - Zuordnung zu Peers
 - Peer hat eindeutige Schlüssel aus seinem Bereich (ID)

Strukturierte Peer-to-Peer Netze

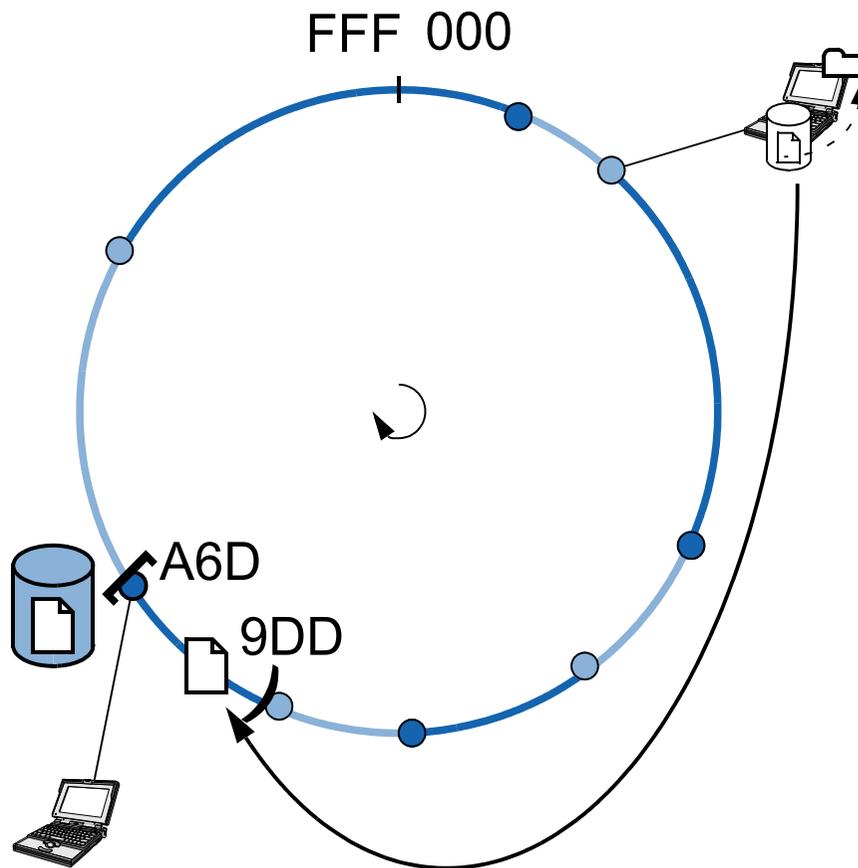
Merkmale

- **Keine zentralen Server**
- **Dokumentenbasiertes Routing**
- **Selbstorganisation der Peers**
 - Routing Topologie
 - Partitionierung des Dokumentenraums
- **Robust gegenüber Dynamik der Peers**
 - An- und Abmeldeverhalten der Peers
 - Adaption der Routing Topologie
 - Dokumente im Vordergrund

Verfügbarkeit: Garantierte Existenz der Dokumente aus Sicht aller Peers

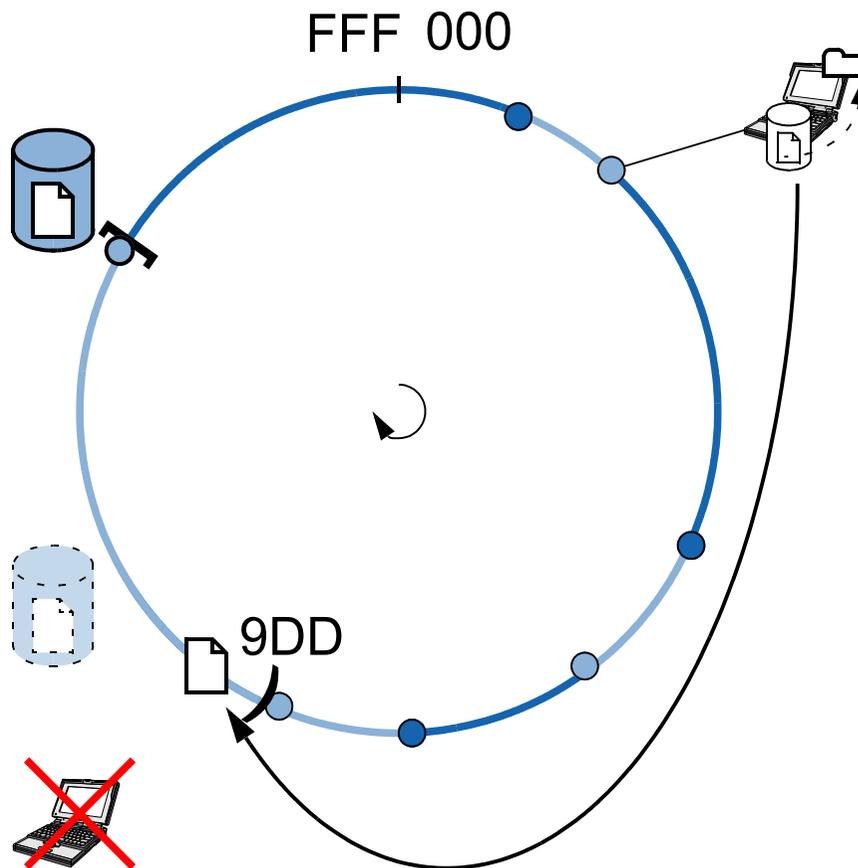
Erreichbarkeit: Kurze Pfadlänge zu Dokumenten (→ Lastverteilung)

Chord Ring



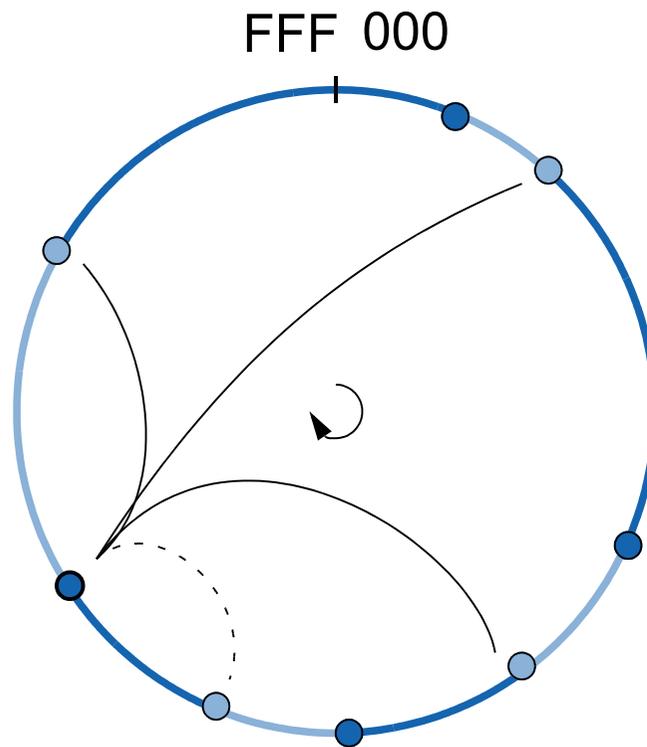
- **Auffinden von Dokumenten**
 - Broadcast
 - Dokumentenbasiertes Routing
- **Dokumente haben eindeutigen Schlüssel (Key)**
Dokumentenraum
- **Partitionierung des Dokumentenraums**
 - Zuordnung zu Peers
 - Peer hat eindeutige Schlüssel aus seinem Bereich (ID)

Chord Ring / Peerausfall



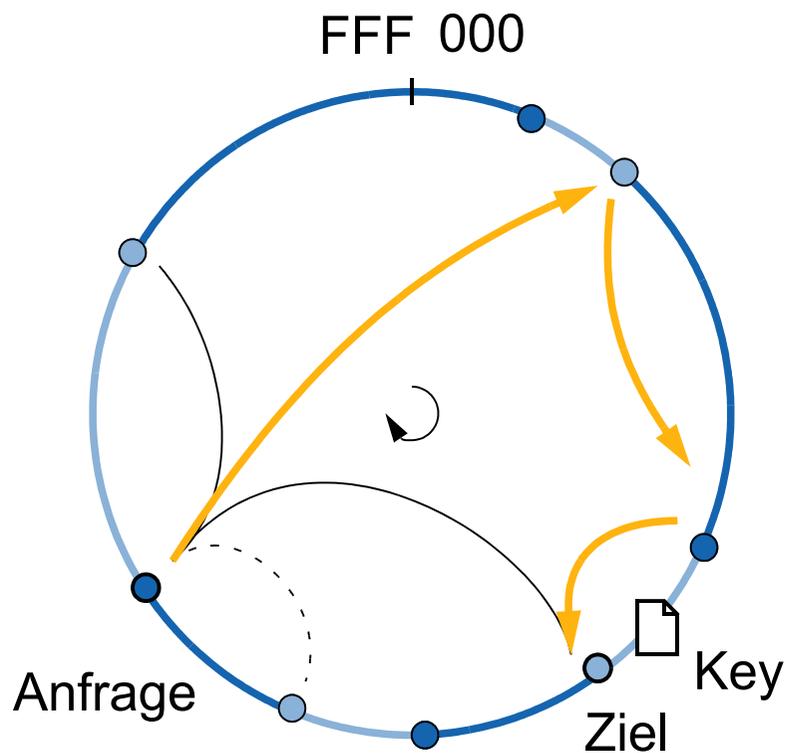
- **Auffinden von Dokumenten**
 - Broadcast
 - Dokumentenbasiertes Routing
- **Dokumente haben eindeutigen Schlüssel (Key)**
Dokumentenraum
- **Partitionierung des Dokumentenraums**
 - Zuordnung zu Peers
 - Peer hat eindeutige Schlüssel aus seinem Bereich (ID)

Routing



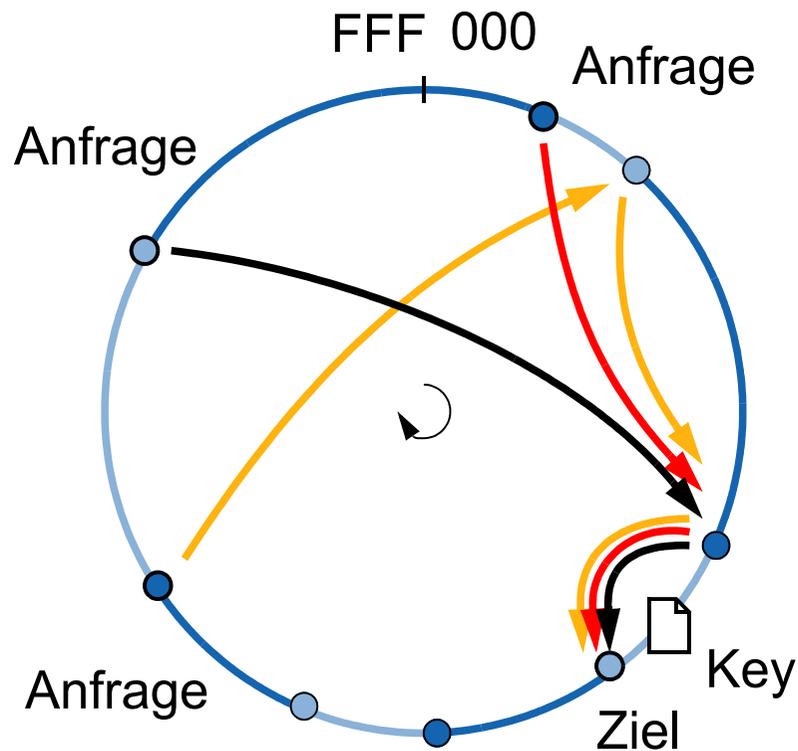
- **Topologie**
 - Nachfolger
 - Weitere Peers
 - (Vorgänger)

Routing



- **Topologie**
 - Nachfolger
 - Weitere Peers
 - (Vorgänger)
- **Routing**
 - Nächster Peer vor Key
 - Immer über Vorgänger

Routing



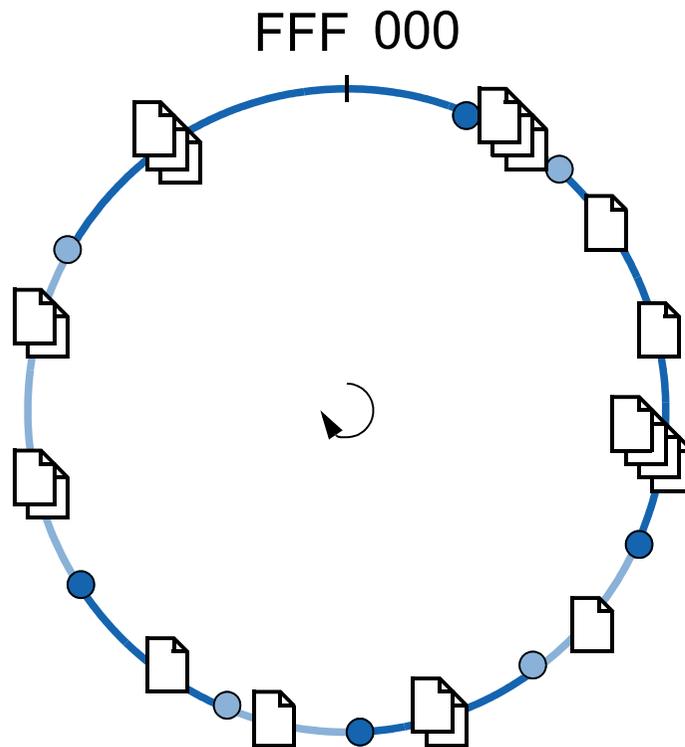
- **Topologie**

- Nachfolger
- Weitere Peers
- (Vorgänger)

- **Routing**

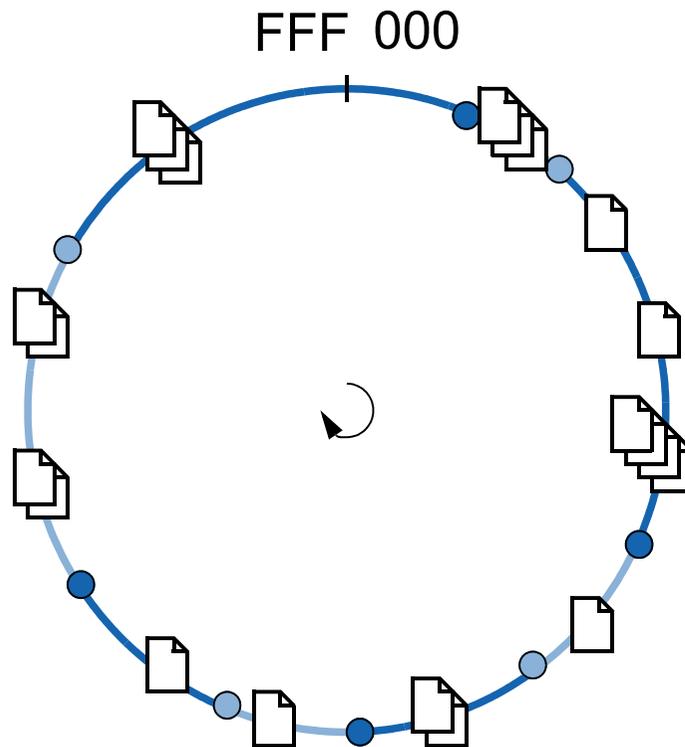
- Nächster Peer vor Key
- Immer über Vorgänger

Verteilung von Keys, IDs und Anfragen



- Zuweisung der Keys und IDs über HASH Wert
- Negativ exponentiell verteilte Abstände
- Individuelle Anfragehäufigkeit von Keys

Ungleiche Verkehrslast an Peers

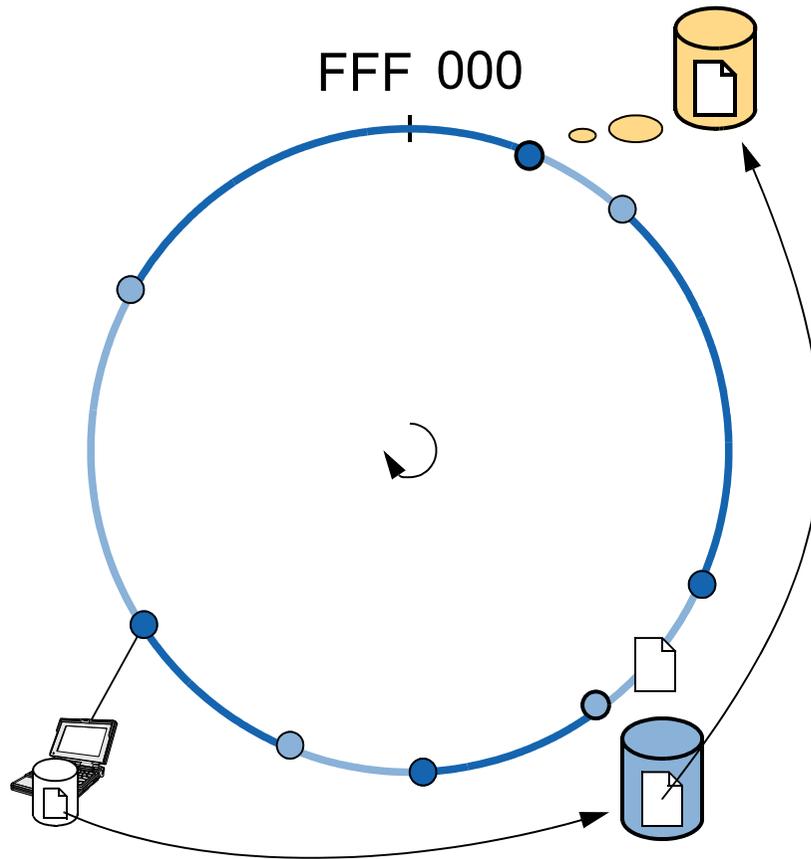


- **Größe des zu verwaltenden Teilrings (Partition)**
- **Anfragehäufigkeit an Keys**
- **Konzentration der Routingpfade vor Ziel**

Lösungsansätze zur Lastverteilung

- **Veränderter Routingmechanismus**
- **Adaptive Zuordnung von Keys und IDs**
- **Fragmentierung der Dateien**
- **Replikations- und Caching Mechanismen**

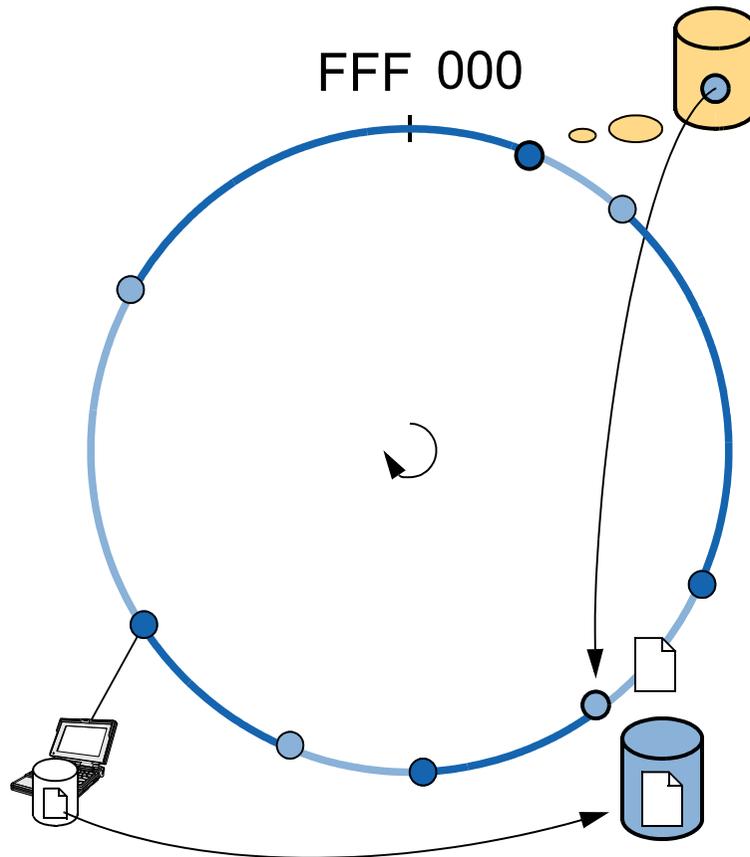
Verteilung von Informationen



Daten

- Kopie der Daten
- Erhöhung der Erreichbarkeit
- Optimierung des Routings
- Erhöhung der Verfügbarkeit

Verteilung von Informationen



Daten

- Kopie der Daten
- Erhöhung der Erreichbarkeit
- Optimierung des Routings
- Erhöhung der Verfügbarkeit

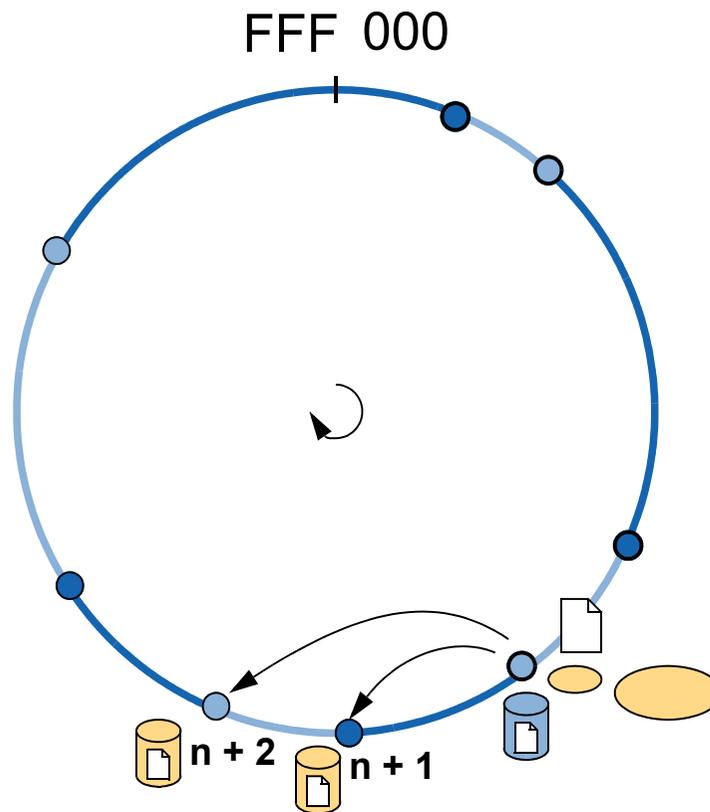
Referenzen

- Referenz auf Peer
- Optimierung des Routings
- Geringes Datenvolumen

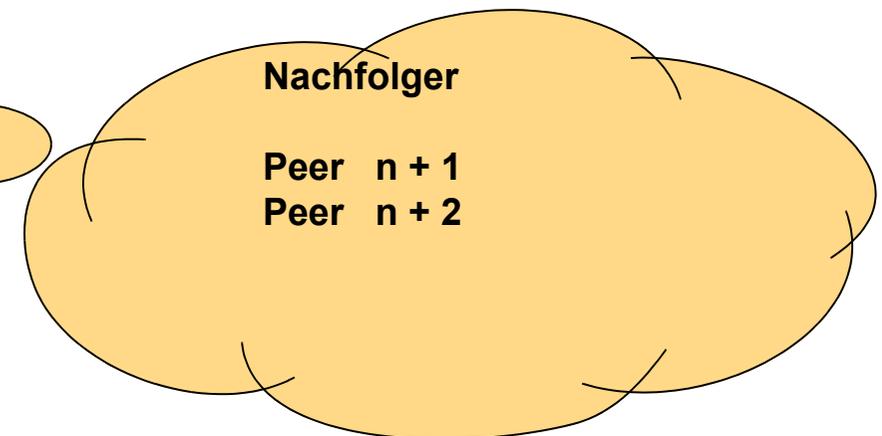
Klassifizierung

- **Zweck (Verfügbarkeit oder Erreichbarkeit)**
- **Daten oder Referenzen**
- **Anzahl der Replikationen (fest oder adaptiv)**
- **Ort des Replikation (fest oder adaptiv)**

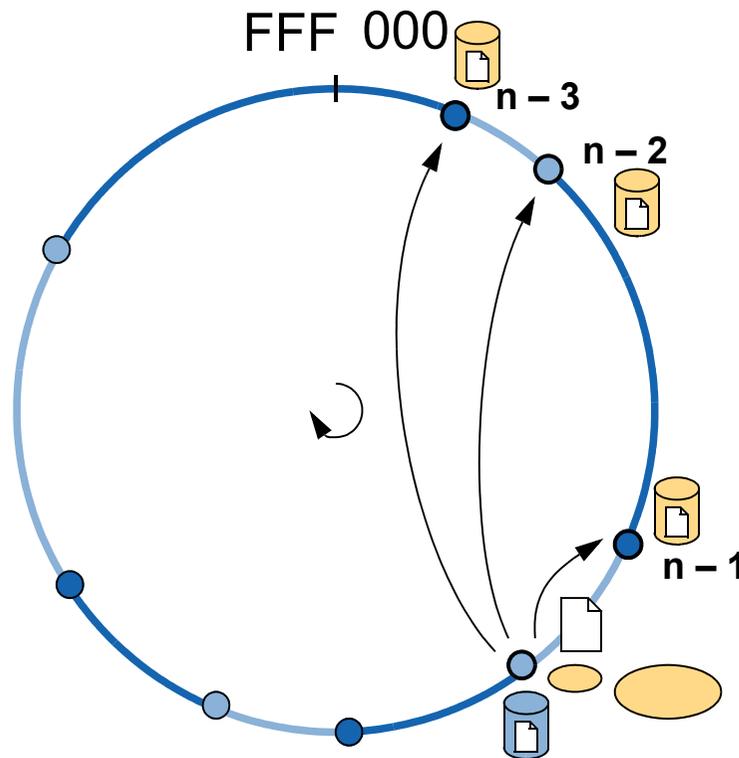
Nachfolger Replikation



Ein Peer **repliziert** eine konstante Anzahl (z.B. 2) an Informationen auf seinen **direkten** Nachfolgern



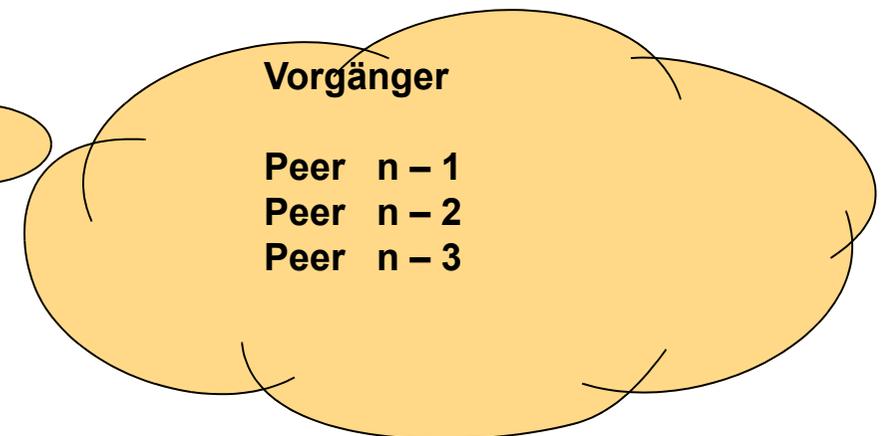
Vorgänger Replikation



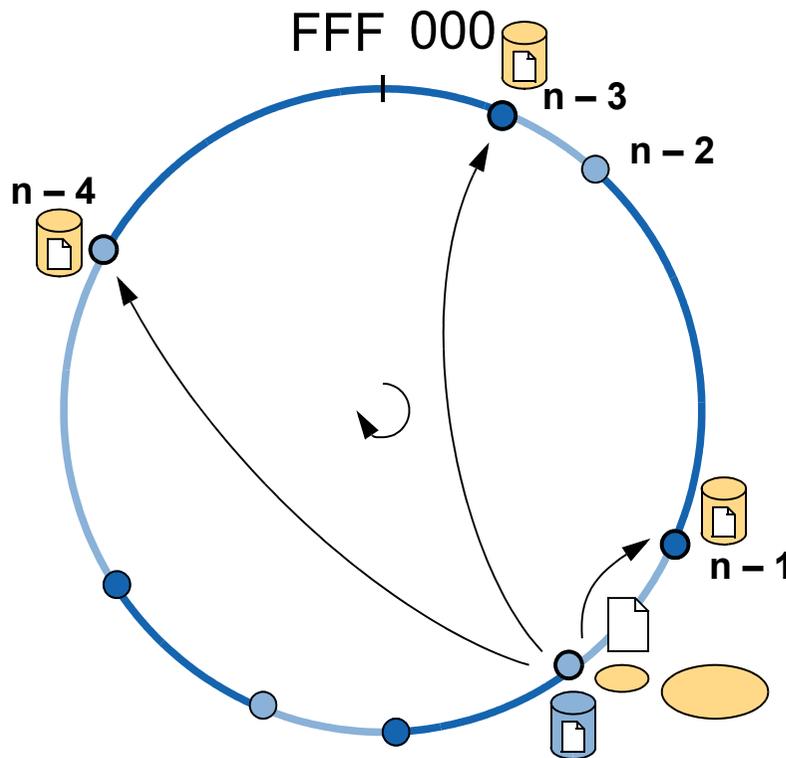
Ungleiche Verkehrslast an Peers:

„Konzentration der Routingpfade vor Ziel“

Ein Peer repliziert eine konstante Anzahl (z.B. 3) an Informationen auf seinen direkten Vorgängern



Adaptive Replikation



Ungleiche Verkehrslast an Peers:

„Konzentration der Routingpfade vor Ziel“

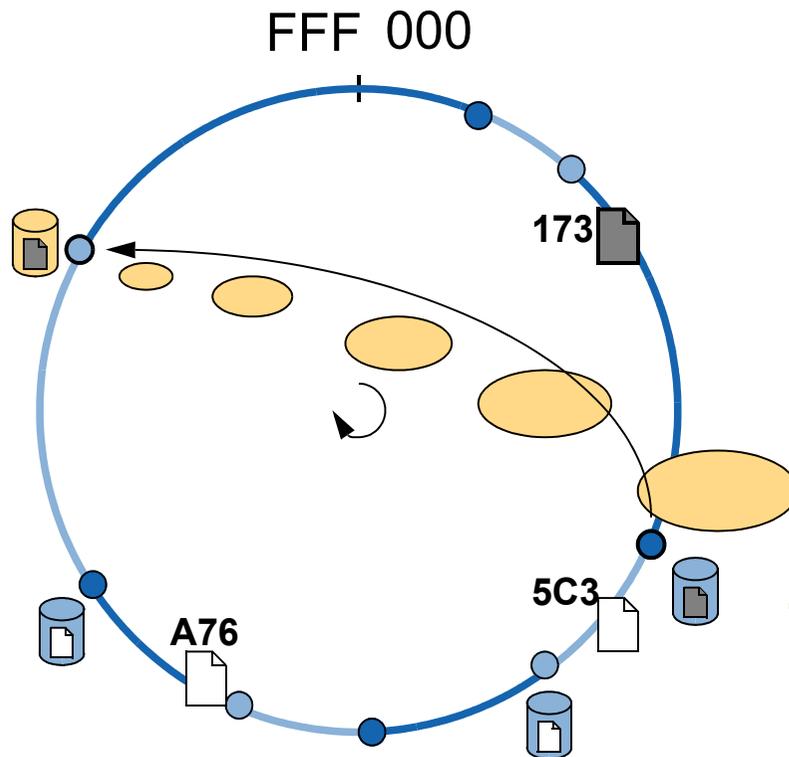
Ein Peer repliziert eine konstante Anzahl (z.B. 3) an Informationen auf am besten geeigneten Peers

Auswertung der Pfade

Peer	n - 1	26 mal
Peer	n - 2	4 mal
Peer	n - 3	17 mal
Peer	n - 4	13 mal

... ..

Caching



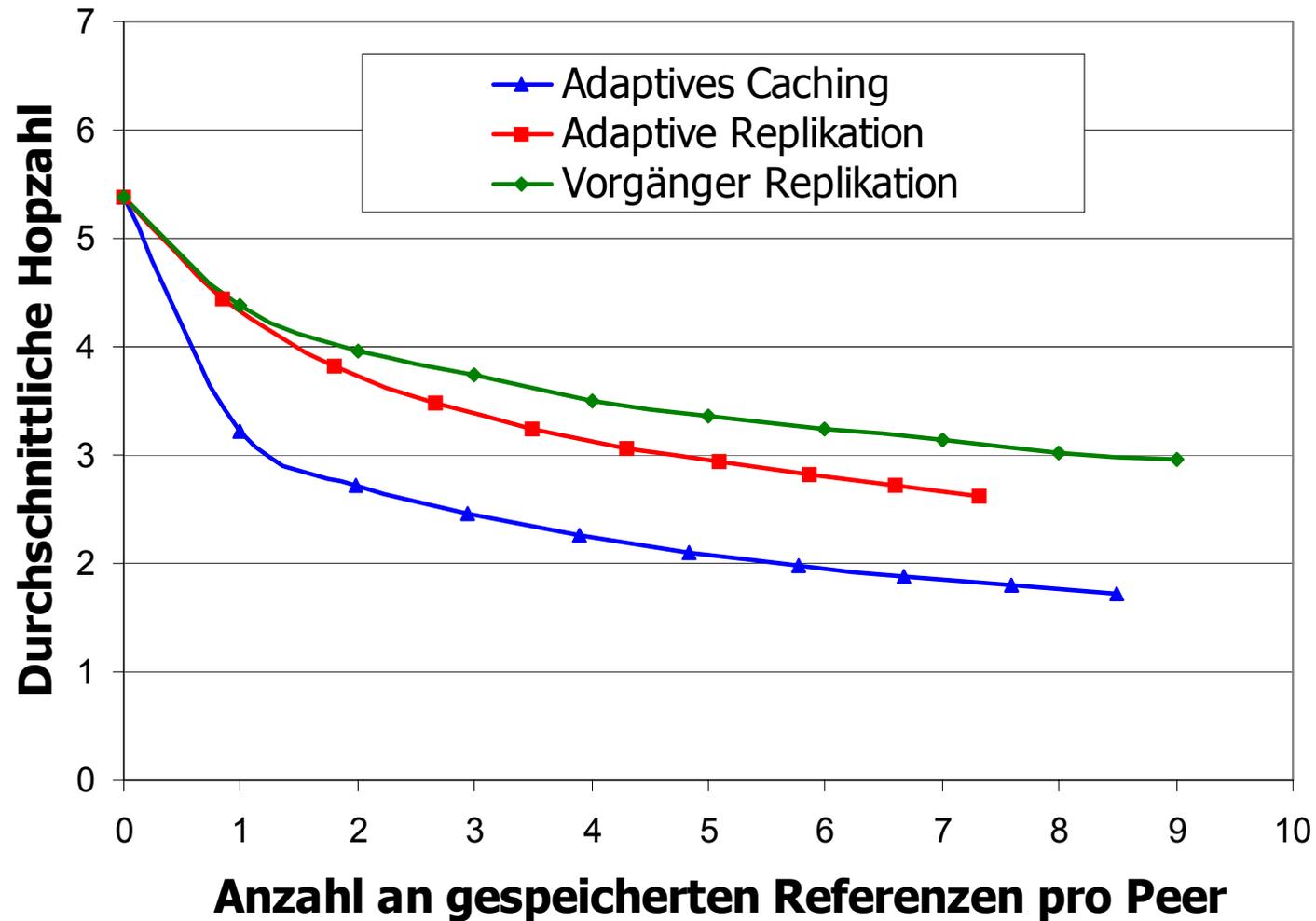
Ungleiche Verkehrslast an Peers:
„Anfragehäufigkeit an Keys“

Ein Peer **cached** selbst eine
konstante Anzahl (z.B. 1) an
Informationen

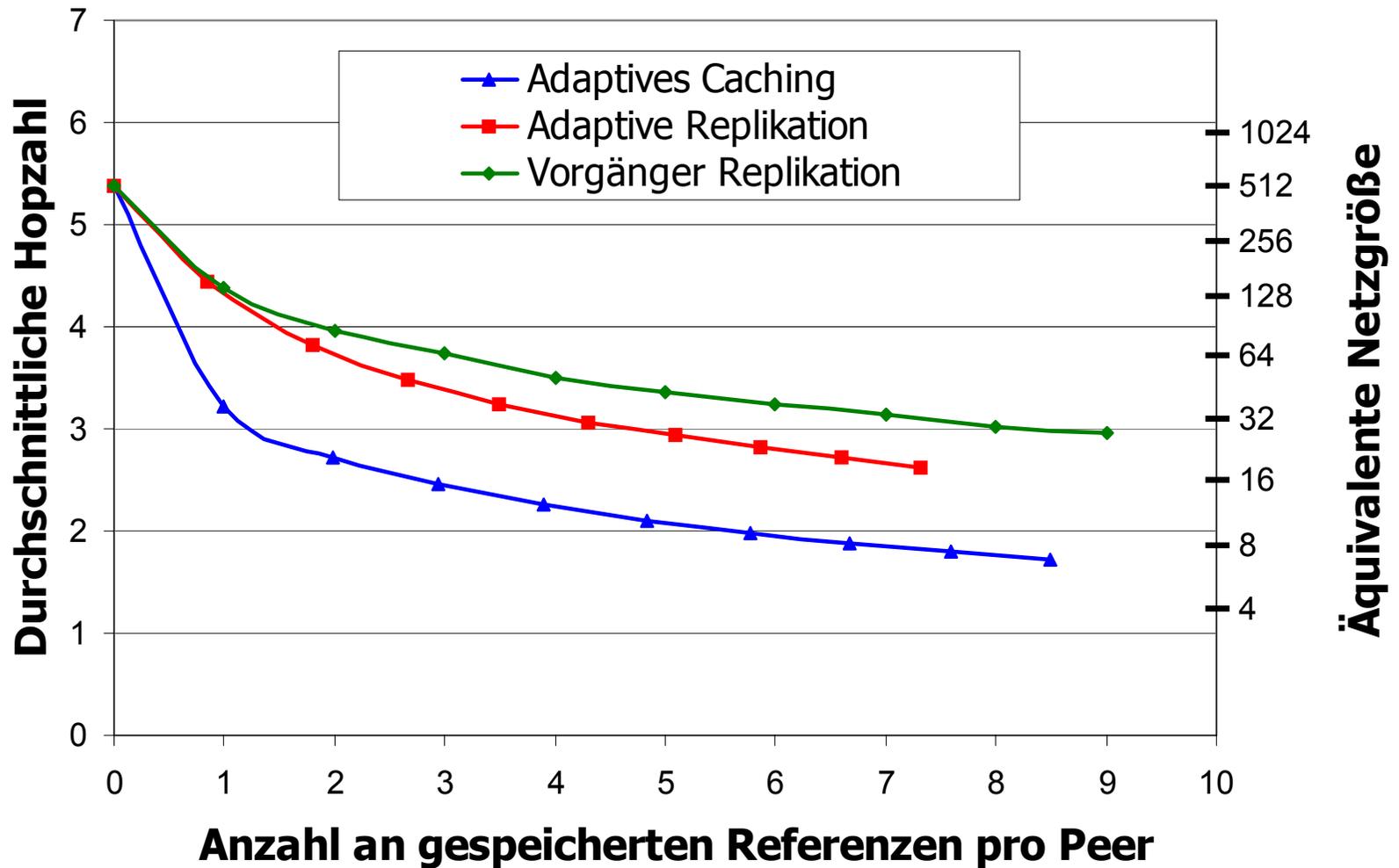
Auswertung der Anfragen

Key	173	26 mal
Key	5C3	21 mal
Key	A76	5 mal
...	...	

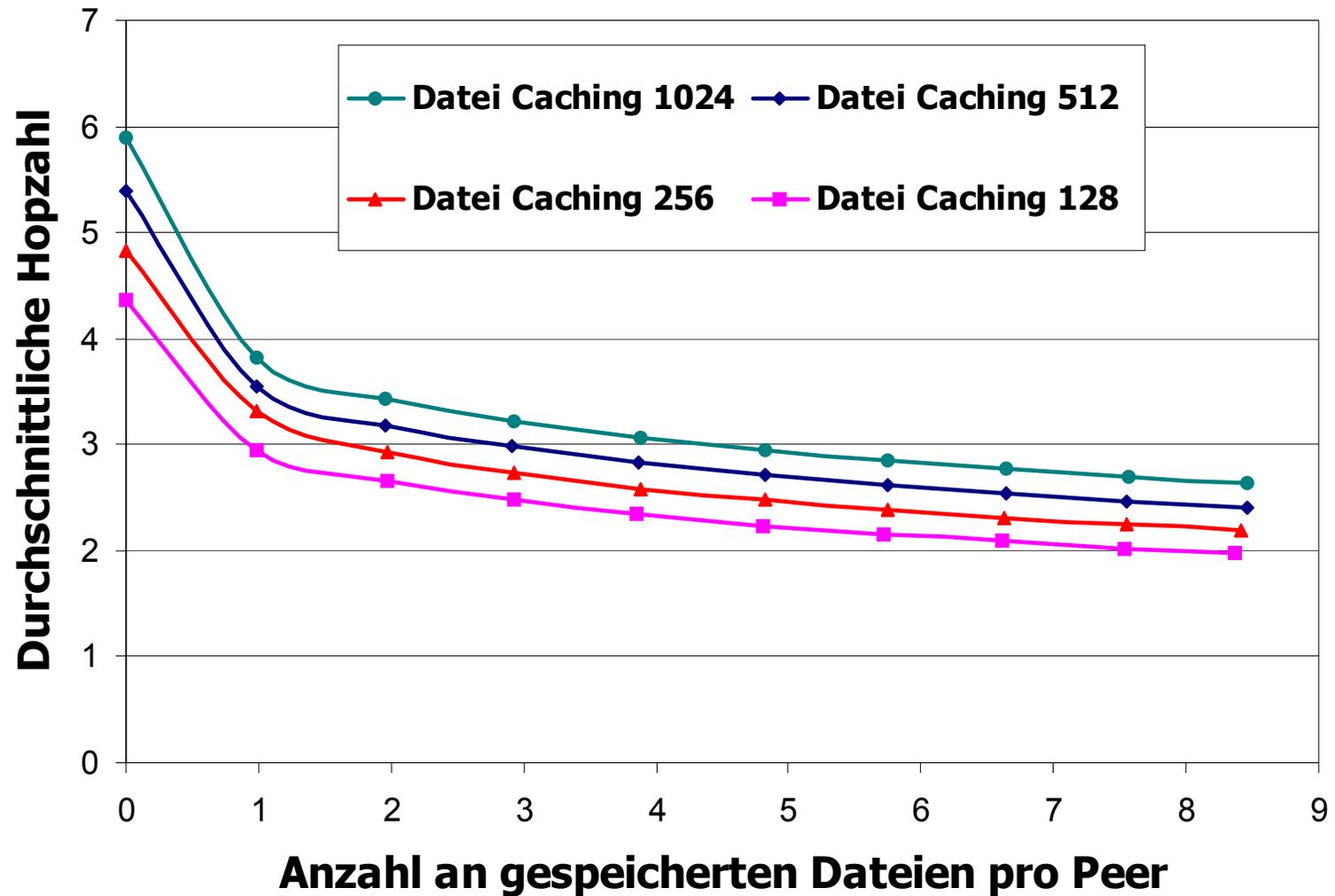
Vergleich der Mechanismen (512 Peers)



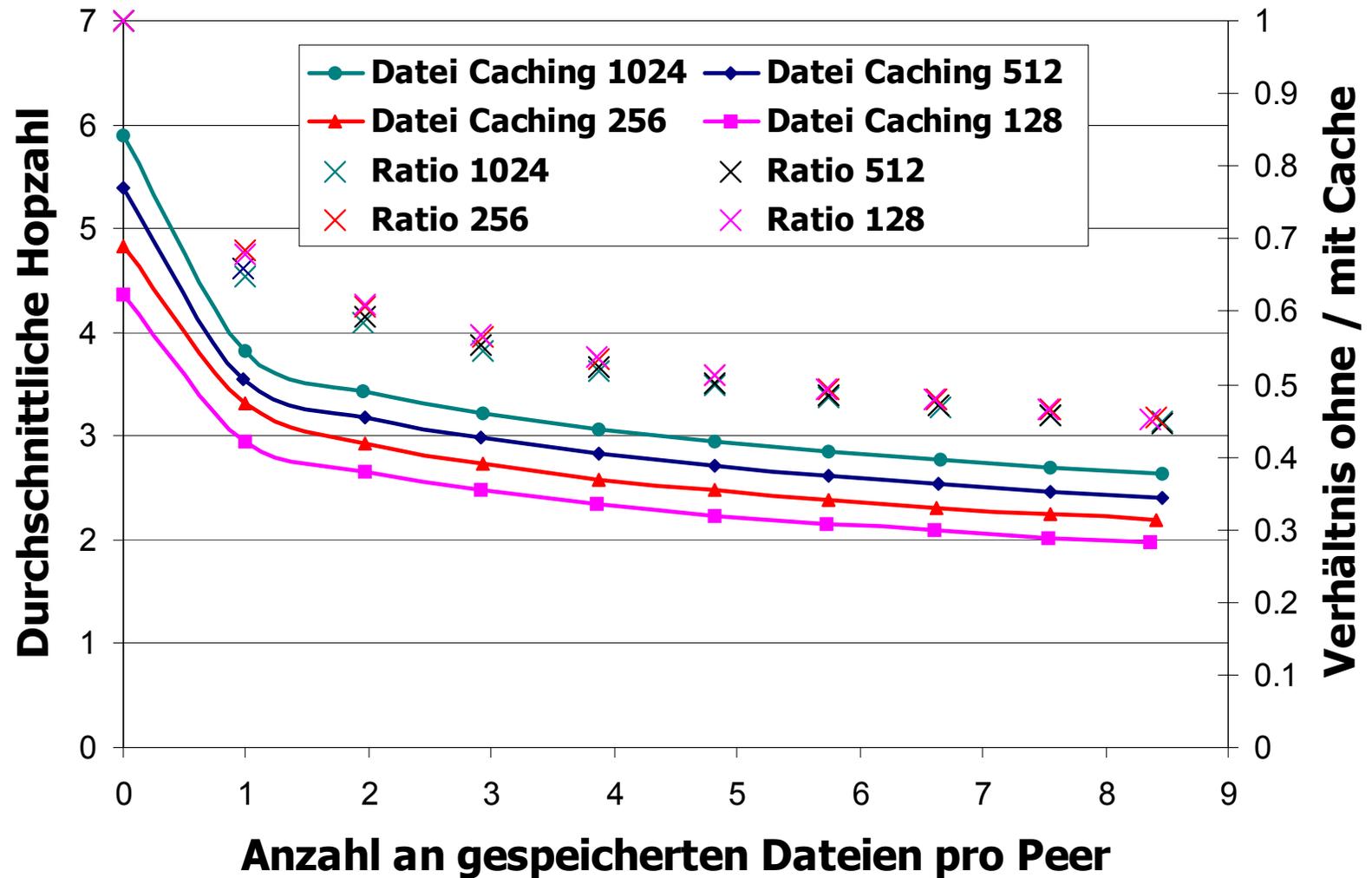
Vergleich der Mechanismen (512 Peers)



Bewertung Caching (128 bis 1024 Peers)



Bewertung Caching (128 bis 1024 Peers)



Klassifizierung und Bewertung

	Primärer Zweck	Daten / Referenz	Ort adaptiv	Anzahl adaptiv	Lastverteilung	Verfügbarkeit
Nachfolger	Verfügbarkeit	Daten	Nein	Nein	-	+++
Vorgänger	Erreichbarkeit	Beides	Nein	Nein	+	++ / -*)
Adaptive Replikation	Erreichbarkeit	Beides	Ja	Nein	++	+ / -*)
Caching	Erreichbarkeit	Beides	Ja	Ja	+++	-

*) Daten / Referenz

Zusammenfassung

- **Effizientes Auffinden von Dokumenten in strukturierten Peer-to-Peer Netzen am Beispiel Chord**
- **Klassifikation von Replikation nach Zweck**
 - Verfügbarkeit
 - Erreichbarkeit
- **Verbesserung der Erreichbarkeit mittels Caching**
Verbesserung der Verfügbarkeit mittels Replikation
→ Kombination mehrerer Mechanismen notwendig

Weiterführende Untersuchungen

- **Dynamik**
- **Konsistenz und Overhead**
- **Benutzerverhalten (Bereitstellung und Anfrage welcher Daten?)**



Replikation in einem homogenen strukturierten Chord Peer-to-Peer Netz

VFF IND/IKR-Workshop

Andreas Reifert, Michael Miksch
reifert@ikr.uni-stuttgart.de

3. März 2006

- Chord
- Replikationsmechanismen
- Untersuchung und Bewertung