Designing Data■Intensive Applications – Studienguide

Konzepte, Kapitelhighlights, Checklisten, EdTech■Bezug

Autor der Studienzusammenstellung: ChatGPT

Hinweis

Diese Datei ist eine eigenständige, paraphrasierte Zusammenstellung als Lernhilfe. Sie enthält keine langen, urheberrechtlich geschützten Volltextauszüge.

Einführung

Überblick über robuste, skalierbare Datensysteme und Kernziele: Zuverlässigkeit, Skalierbarkeit, Wartbarkeit.

Datenmodelle

Relational, Dokumente, Graphen; Schema∎on∎Write vs. Schema∎on∎Read.

Speicherung & Indizes

B■Trees vs. LSM■Trees, Log■strukturierte Ansätze, Schreib-/Lese■Pfade.

Replikation

Single■Leader, Multi■Leader, Leaderless; Konfliktauflösung und Quorums.

Partitionierung

Hash vs. Range, konsistentes Hashing, Rebalancing, Hot≣Keys vermeiden.

Transaktionen

ACID, Isolation, Anomalien; Sagas und Idempotenz in verteilten Abläufen.

Zeit & Ausfälle

Unzuverlässige Netze, Clocks, Timeouts, Backoff, Circuit Breaker.

Konsens & Konsistenz

Linearizierbarkeit, eventual consistency, Paxos/Raft in der Praxis.

Batch■Verarbeitung

MapReduce■Stil, Reprocessing, unveränderliche Logs.

Stream Verarbeitung

Event■Logs, Windowing, State, Exactly■once, Materialized Views.

Architektur■**Muster**

Outbox/Inbox, CDC, Event Sourcing, CQRS.

AntiPatterns

Geteilte DB als Kopplung, schemalose Änderungen ohne Versionierung.

Checklisten

 $Work load profil, \ Datenmodell, \ Replikation/Partitionierung, \ Observability, \ Kosten.$

EdTech■**Bezug**

Kurs-/Nutzer∎Daten, Reporting (Batch) und Live∎Metriken (Stream).

ReviewFragen

Wann LSM vs. B■Tree? Unterschied SI vs. RR? Wann Multi■Leader?

Glossar

ACID, CAP, Quorum, Snapshot Isolation, CDC, Backpressure, SLO/SLA.

Eigene Notizen zu Designing Data■Intensive Applications – Studienguide.

Quellen & weiterführende Links

- O'Reilly - Buchseite: https://www.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/