

Agenda für ein Seminar “Apache Cassandra”

Dr. Rainer Sawitzki, 7.3.2023

Dauer: 3 Tage, 12 Blöcke mit jeweils 90 Minuten. 9:00 - 16:15

Termin: Werden noch vereinbart

Ort: Online

Zielgruppe: Entwickler, Architekten, Systembetreiber

Vorkenntnisse: Grundkenntnisse in relationalen Datenbanken und SQL sind von Vorteil

Methode:

- Vortrag, Präsentation, Diskussion, eigenständige Übungen

Sprache: Deutsch

Technische Voraussetzungen:

- Jeder Teilnehmer hat eine eigene Umgebung mit
 - Teams
 - Zugriff auf die Remote Rechner der Cegos Integrata
- Die Seminarumgebung sowie ein Cassandra-System wird auf Rechnern der Cegos Integrata bereitgestellt

Inhalte:

- NoSQL (4 Blöcke)
 - Kategorisierung
 - CAP-Theorem
 - Architektur von NoSQL-Datenbanken
 - Unterschiede/Vor- und Nachteile von NoSQL zu Standarddatenbanken (Relationalen Datenbanken) bzgl. Architektur/Performance/Ausfallsicherheit
 - Platzierung von Apache Cassandra im NoSQL-Umfeld
- First Contact (1 Block)
 - Zugriff auf eine vorbereitete Cassandra-Instanz
 - Client-Werkzeuge zum Zugriff
 - Absetzen elementarer Befehle zur Daten-Manipulation
- Datenhaltung in Cassandra im Detail (3 Blöcke)
 - Grundlagen CQL
 - Transaktionen
 - Tombstones
 - Wie entstehen Sie
 - Wie kann man sie vermeiden
 - Wann werden sie kritisch für das System/für die Performance
 - Kann man sie optimieren?
 - Kann man das Löschen optimieren(CQL/TTL/anderes?)?
- Betrieb (2 Blöcke)
 - Installation
 - Backup und Recovery
 - Netzwerkverschlüsselung und Sicherheit (Authentifizierung, Autorisierung, TLS Verschlüsselung Server-Server und Client-Server)
 - Cassandra Performance-Benchmarking und -Tuning
 - Werkzeuge
 - Migration von Cassandra Versionen
 - Hochverfügbares Cluster-Setup und Replikationsstrategien
- Übersicht: Betrieb in der Cloud (1 Block)
 - Cluster-Deployment in der Cloud (Single-Datacenter Setup)

- Cluster-Monitoring mit OpsCenter
 - Cluster-Deployment in der Cloud (Multi-Datacenter Setup)
- Frage- und Diskussionsrunde (1 Block)