

Apache Cassandra

Kursangebot 2013



Geroba Data Technologies e.U.

DI Gernot R. Bauer +43 650 26 44 013

DI Philipp Potisk +43 660 46 02 632

bigdata@geroba.at

Kaiserfeldgasse 17/25

8010 Graz



Was ist Apache Cassandra?

Die Datenmenge, die von Applikationen verarbeitet werden muss, hat sich in den letzten Jahren vervielfacht. Die Erfahrung hat gezeigt, dass klassische Datenbanksysteme wie Oracle oder Microsoft SQL Server ab einer gewissen Größe nicht mehr kostengünstig skalierbar sind. Zugleich müssen Anwendungen auch bei großen Datenmengen performant arbeiten.

Apache Cassandra ist ein Vertreter der sogenannten NoSQL-Datenbanken. Diese Art von neuartigen Datenbanksystemen zeichnen sich dadurch aus, dass sie auf große Datenmengen, hohe Verfügbarkeit, gute Performance und einfache, dynamische sowie kostengünstige Skalierbarkeit ausgelegt sind. Dies erlaubt es Applikationen, auch große Datenmengen zu verarbeiten, ohne auf teure und hochspezialisierte Server-Hardware zurückgreifen zu müssen.

Apache Cassandra ist ein erfolgreicher und oft eingesetzter Vertreter von NoSQL-Datenbanken. Cassandra wurde von Facebook entwickelt und kombiniert Ideen von Googles „Bigtable“-Datenbank sowie Amazons Dynamo-Datenbank. Cassandra hat sich in der Praxis als flexible Datenbank erwiesen, die mit wenig Aufwand als Cluster betrieben werden kann. Dies erlaubt den Betrieb von hochverfügbaren Anwendungen, die auch beim Ausfall einzelner Server oder sogar ganzer Datacenter noch immer ihre Dienste tut. Dabei werden (auch unter Last) Lese- und insbesondere Schreibgeschwindigkeiten erreicht, die sich mit klassischen Datenbanksystemen schwer bis gar nicht realisieren lassen.

Apache Cassandra ist ein praxiserprobtes Datenbanksystem, das u.a. von Unternehmen wie Twitter, eBay, Cisco oder Walmart erfolgreich eingesetzt wird.

Das Kursangebot

Überblick

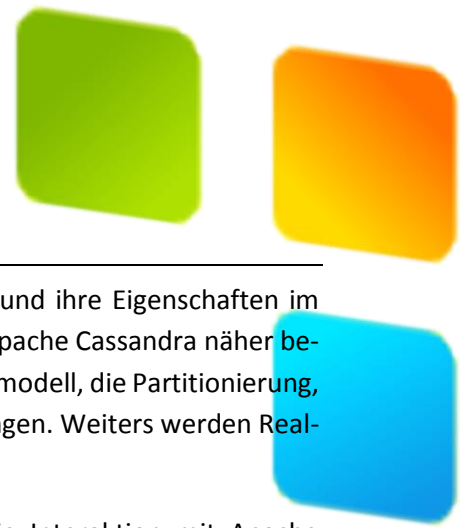
Unser zweiteiliges Kursangebot bietet Architekten, Entwicklern, Datenbankadministratoren sowie IT-Verantwortlichen eine Mischung aus Theorie, Hands-On-Beispielen sowie Use Cases aus der Praxis. Dadurch erhalten die Schulungsteilnehmer einen Einblick in die Architektur und Theorie hinter Cassandra sowie die Vor- und Nachteile gegenüber traditionellen Datenbanklösungen. Durch die Hands-On-Beispiele wird dieses theoretische Wissen in Übungen gefestigt. Use Cases sowie Praxiserfahrungen runden die Schulung ab und geben den Teilnehmern die Mittel, um erfolgreich Projekte mit Cassandra durchführen zu können.

Zielgruppen

Die Kurse richten sich an Software-Architekten, Software-Entwickler, Datenbankadministratoren und IT-Verantwortliche.

Trainer

DI Gernot Bauer und DI Philipp Potisk arbeiten seit Abschluss ihres Softwareentwicklung/Wirtschaftsstudiums an der TU Graz als Software-Architekten und Java-EE-Entwickler. Seit 2010 haben sie intensive Praxiserfahrung mit Apache Cassandra. Sie sind beide aktive Mitglieder der Cassandra Community und stets am Fortschritt des Open Source Projekts interessiert.



Grundlagenkurs

Im ersten Teil werden grundlegende Konzepte von NoSQL-Datenbanken und ihre Eigenschaften im Vergleich zu traditionellen SQL-Stores vermittelt. Darauf aufbauend wird Apache Cassandra näher beleuchtet. Hierbei werden auf die grundlegende Ring-Architektur, das Datenmodell, die Partitionierung, Replikationsmechanismen sowie Konsistenzstufen von Cassandra eingegangen. Weiters werden Real-Life Use Cases vorgestellt.

Im Workshop-Format werden die Installation, die Konfiguration und die Interaktion mit Apache Cassandra präsentiert. Weiters werden mehrere Client-Frameworks anhand von Praxisbeispielen erarbeitet. Das Hauptaugenmerk wird auf die Java-Client-Frameworks gelegt und anhand einfacher Implementierungsaufgaben werden deren Stärken und Schwächen beleuchtet.

Ziel dieses Grundlagenabschnittes ist es, Entwicklern und operativem Personal das Mindset, das hinter NoSQL-Stores und speziell hinter Cassandra steht, zu vermitteln. Dies geschieht insbesondere auch im Vergleich zu etablierten SQL-Datenbanken. Das Hands-On erlaubt den Entwicklern, die theoretischen Grundlagen anhand von praktischen Beispielen ausgiebig zu testen.

Agenda

- Einführung in BigData & NoSQL Datenbanken
- Eigenschaften von Cassandra und Real-Life Use Cases
- Datenmodell von Cassandra
- Hands-on: Installation, Konfiguration, Command Line-Interface
- Architektur von Cassandra
- Überblick über Client-Frameworks
- Java Beispiele für Thrift, Hector, Astyanax, Kundera

Dauer

1 Tag (ca. 8 Stunden inklusive Pausen)

Unterlagen

Die Präsentationsunterlagen sowie der Source Code der Übungsbeispiele werden den Teilnehmern zur Verfügung gestellt.

“They have the hands-on experience to answer real-life questions. The course features well prepared code snippets & examples and an exciting experiment.”

(Feedback eines Schulungsteilnehmers)

Sie haben Interesse?

Weitere Details und einen maßgeschneiderten Schulungsaufbau für Ihre MitarbeiterInnen besprechen wir gerne mit Ihnen.

Geroba Data Technologies e.U.

DI Gernot Bauer

+43 650 2644013

bigdata@geroba.at

Kaiserfeldgasse 17/25

8010 Graz



Fortgeschrittenenkurs

Im Fortgeschrittenenkurs werden verschiedene Aspekte von Cassandra im Detail betrachtet. Diese beinhalten die Verarbeitungsschritte von Lese- und Schreiboperationen und sämtliche damit in Verbindung stehenden Mechanismen, wie z.B.: Tombstones, Compaction, Read Repair, Hinted Handoff, etc.

Im Bereich der Datenmodellierung werden Best Practices gezeigt. Anhand ausgewählter Praxisbeispiele werden mögliche Modellierungsansätze sowie deren Vor- und Nachteile diskutiert.

Die neue Schnittstelle zu Cassandra (CQL 3.0) wird anhand von Übungsbeispielen vorgestellt. Weiters wird gemeinsam eine einfache Java-Anwendung entwickelt, die die Verwendung des CQL-Driver von Datastax demonstriert und in einem Live-Cluster von 4 Nodes ausgeführt wird.

Es werden mehrere Tools für die Administration von Cassandra-Clustern in der laufenden Anwendung vorgestellt. Ebenso werden Tuning-Tips, Patterns sowie Anti-Patterns besprochen.

Ziel dieses Kurses ist es, Entwickler und operativen Mitarbeitern einen tieferen Einblick in Cassandra zu geben, sowie einzelne Aspekte detailliert auszuführen. Darüber hinaus sollen die im Grundlagenkurs erworbenen Kenntnisse bezüglich „Thinking in NoSQL/Cassandra“ anhand praxisnaher Use Cases gemeinsam diskutiert werden.

Agenda

- Use Case Ebay
- Cassandra Intern (Verarbeitung von Lese-/Schreib-Operationen, etc.)
- Tools für die Administration (nodetool, JMX, OpsCenter) inkl. Demo am Live-Cluster
- Hands-on Einführung in Cassandra Query Language 3.0
- Entwicklung einer Java-Applikation und Inbetriebnahme im Live-Cluster
- Tuning, Patterns, Anti-Patterns
- Optional: Kundenspezifische Use Cases
- Optional: Cassandra & Hadoop, Cassandra & Solr

Dauer

1 Tag (ca. 8 Stunden inklusive Pausen)

Unterlagen

Die Präsentationsunterlagen sowie der Source Code der Übungsbeispiele werden den Teilnehmern zur Verfügung gestellt.

“The guys’ style is fresh combined with deep technical knowledge.”
(Feedback eines Schulungsteilnehmers)

Sie haben Interesse?

Weitere Details und einen maßgeschneiderten Schulungsaufbau für Ihre MitarbeiterInnen besprechen wir gerne mit Ihnen.

Geroba Data Technologies e.U.

DI Gernot Bauer

+43 650 2644013

bigdata@geroba.at

Kaiserfeldgasse 17/25

8010 Graz