

Benchmarking von NoSQL-Datenbanksystemen

Holger Wegert

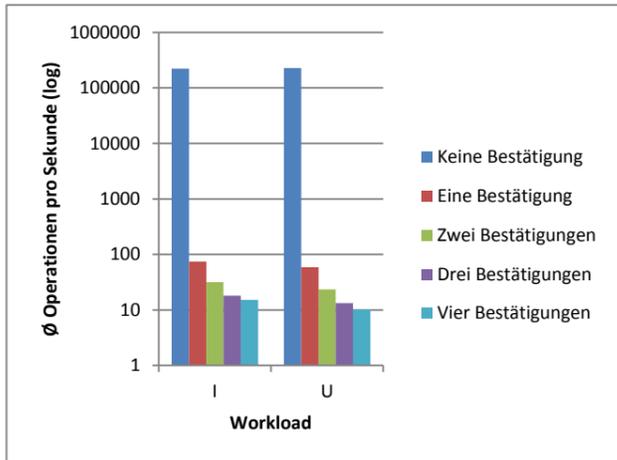


Abbildung 1: Couchbase – Abgewartete Persistenzbestätigungen von unterschiedlichen Mengen von Cluster-Nodes

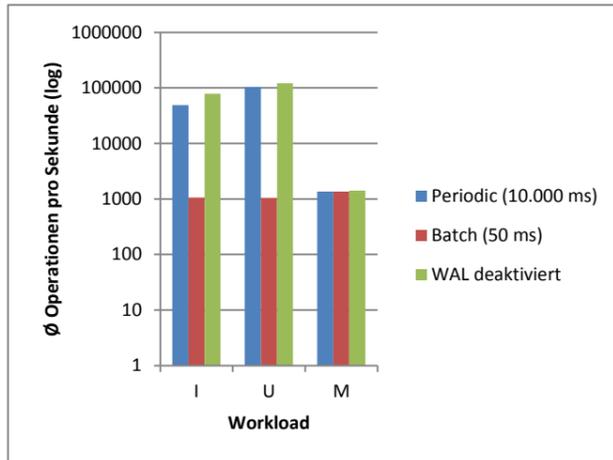


Abbildung 2: Cassandra – Write-Ahead-Logging

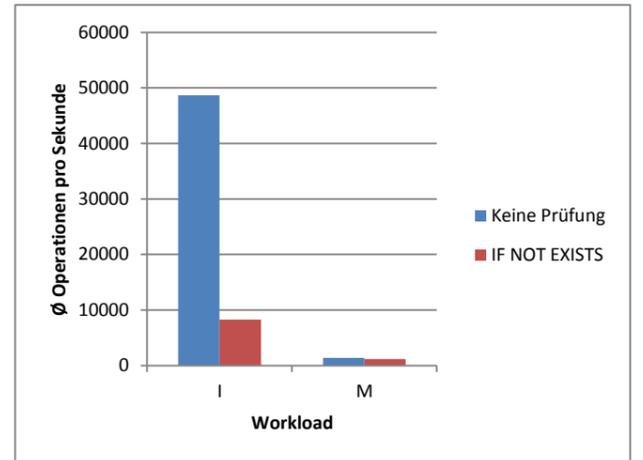


Abbildung 3: Cassandra – Existenzprüfung bestehender Keys bei Insert-Operationen

Motivation

Als IT-Dienstleister mit einer ausgeprägten Expertise im Bereich relationaler Datenbanksysteme verfolgt die ORDIX AG stets aufkommende Trends und Technologien im gesamten Datenbankumfeld. Einen über die letzten Jahre anhaltenden Trend stellen NoSQL-Datenbanksysteme dar. Insbesondere im Kontext von Big Data geprägten Problemstellungen gewannen NoSQL-Datenbanksysteme, durch Eigenschaften wie der horizontalen Skalierung, an Bedeutung.

Ein wesentlicher Aspekt bei der Auswahl eines derartigen Systems, ist, dass die meisten Systeme für einen bestimmten Bereich von Anwendungsszenarien konzipiert wurden. Entsprechend erfolgten bereits bei der Architektur grundlegende Designentscheidungen, die das Verhalten der unterschiedlichen Systeme entsprechend charakterisieren. Neben Unterschieden bei der Performanz wirken sich diese wesentlich auf die Dauerhaftigkeit, Verfügbarkeit und Konsistenz von Daten aus.

Infolge der zunehmenden produktiven Verwendung und der kontinuierlichen Weiterentwicklung verfügen die Systeme mittlerweile über eine Vielzahl an Konfigurationsparametern, die eine individuelle Anpassung an die Bedürfnisse des Kunden erlauben.

Zielsetzung

Im Gegensatz zu einer Vielzahl an bereits durchgeführten Benchmarks von NoSQL-Datenbanksystemen, untersucht diese Arbeit nicht die Leistungsfähigkeit verschiedener Systeme im direkten Vergleich. Vielmehr wurden die Leistungsunterschiede verschiedenartiger Betriebskonfigurationen für einzelne NoSQL-Datenbanksysteme getrennt voneinander untersucht. Durch den Verzicht auf einen direkten Vergleich konnten die Systeme grundlegend entsprechend ihrer Fähigkeiten evaluiert werden, ohne dass dabei gleichzeitig zusätzliche Einstellungen vorgenommen werden mussten, die lediglich eine mehr oder weniger gute Vergleichbarkeit zwischen den heterogenen Systemen herstellen. Die somit praxisnäheren Ergebnisse sollen Mitarbeiter der ORDIX AG zukünftig bei der fundierten Technologieberatung Ihrer

Kunden unterstützen. Gegenstand der konkreten Untersuchungen sind die populären NoSQL-Datenbanksysteme Couchbase Server 3.0.1 und Apache Cassandra 2.1.2.

Vorgehensweise

Für das Benchmarking der Systeme wurde Thumbtack Technologies Variante des Yahoo! Cloud Serving Benchmarks (YCSB) ausgewählt und weiterentwickelt. Die durchgeführten Modifikationen umfassen bspw. funktionale Erweiterungen, wie den Ausbau der Multi-Client Unterstützung oder der bestehenden Warmup-Implementierung.

Für die Evaluierung der unterschiedlichen Benchmark-Konfigurationen wurden vier eigene Workload-Profile definiert. Hierbei handelt es sich um drei Operationsarten reine Workloads, welche eine fokussierte Betrachtung der Auswirkungen auf Read, Update und Insert-Operationen erlauben (Workload-R/U/I). Beim vierten Workload-M werden lesende und schreibende Operationen gleichermaßen berücksichtigt, wobei sich die schreibenden Operationen gleichwertig in Updates und Inserts unterteilen. Jeder Workload wurde insgesamt dreimal hintereinander gemessen, wobei jede Ausführung jeweils eine Laufzeit von zehn Minuten besaß. Diese unterteilte sich in eine fünfminütige Warmup-Phase und eine effektive Messphase von weiteren fünf Minuten. Sowohl die gewählte Anzahl an durchgeführten Wiederholungen, als auch die Laufzeit einer einzelnen Workload-Messung stellen einen Kompromiss zwischen einer höchstmöglichen Messstabilität und einer möglichst niedrigen Gesamtlaufzeit dar. Die Ergreifung dieser Maßnahme war notwendig, um eine Vielzahl unterschiedlicher Benchmark-Konfigurationen untersuchen zu können.

Ergebnisse

Die Evaluierung der Datenbanksysteme hat gezeigt, dass die unterschiedlichen Konfigurationsmöglichkeiten eine Verwendung der Systeme bei verschiedenartigen Anforderungen grundsätzlich ermöglicht. Je nach Konfiguration wirkt sich deren Anpassung jedoch unterschiedlich stark auf das Leistungsvermögen der Systeme aus. Als übergreifendes Resümee für die durchgeführte Evaluierung von Couchbase kann festgehalten

werden, dass insbesondere die vorherrschende In-Memory Fokussierung und der Verzicht auf eine garantierte Dauerhaftigkeit bei schreibenden Operationen (siehe Abb. 1) ein außerordentliches Leistungsverhalten ermöglicht. Sobald ein Anwendungsszenario jedoch abweichende Anforderungen erfordert, bricht der erreichbare Durchsatz an Operationen pro Sekunde essentiell ein.

Bezüglich der erfolgten Evaluierung von Cassandra wurde bei der Verarbeitung schreibender Anfragen ein ausgeprägtes Leistungsverhalten festgestellt. Erreicht wird dies unter anderem durch die asynchrone Persistierung des Commit Logs (siehe Abb. 2) und einer standardmäßig unterlassenen Existenzprüfung der betroffenen Keys bei schreibenden Operationen (siehe Abb. 3). Beachtliche Leistungseinbußen erfolgen dagegen durch die Anpassung der Write-Ahead-Logging-Konfiguration, so dass die Dauerhaftigkeit erfolgreicher Datenänderungen garantiert wird. Ebenso deutlich beeinträchtigt die Replikation mit zunehmender Konsistenzanforderung das Leistungsvermögen von Cassandra.

Neben der durchgeführten Erweiterung von Thumbtacks YCSB Variante wurden zudem einige grundlegende Fehler identifiziert und behoben. Hierzu zählen insbesondere die fehlerbehaftete Protokollierung der Messergebnisse während der Benchmark-Durchführung und die teilweise verfälschte Aggregation der Latenzwerte bei der nachträglichen Durchschnittsbildung. Darüber hinaus erwies sich ebenfalls die Implementierung der Benchmark-Terminierung als problematisch. Trotz einer maximal definierten Laufzeit wurde bei einzelnen Konfigurationen eine variierende Messdauer zwischen den eingesetzten Client-Threads festgestellt. Zur Vermeidung einer diesbezüglich bedingten Ergebnisverfälschung wurde die ursprüngliche Implementierung ersetzt, so dass keine Messwerte mehr nach Ablauf der maximal definierten Laufzeit protokolliert werden.