

DATENBANKANWENDUNGEN EFFIZIENTER MACHEN DAS MySQL-CLUSTER-PROJEKT AM URZ

Die meisten der vom URZ betriebenen Web-Anwendungen sind datenbankbasiert – darunter für den Lehrbetrieb unabdingbare Anwendungen wie LSF oder Moodle – und verlangen einem Datenbankverwaltungssystem einiges ab: Ein Höchstmaß an Verfügbarkeit, Schnelligkeit, Datenintegrität und Effizienz in der Administration ist gefragt. Die am häufigsten eingesetzten Datenbankverwaltungssysteme, die „hinter“ den Web-Anwendungen laufen, sind PostgreSQL und MySQL. Letzteres ist mit mehr als 6 Millionen Installationen zur Zeit das populärste System weltweit und wird auch vom URZ eingesetzt. Ziel des vor einigen Monaten am URZ gestarteten Projekts ‚MySQL-Cluster‘ ist es nun, eine Cluster-Technologie finden, die für die datenbankbasierten Webanwendungen des URZ am besten passt, und diese dann auch in Betrieb zu nehmen.

Es gibt drei große Herausforderungen an Datenbankverwaltungssysteme: Erstens muss ein unterbrechungsfreies Aktualisieren der Software möglich sein – ‚Rolling Upgrades‘ und ‚Rolling Restarts‘ –, zweitens eine sehr gute Performance, um die Web-Anwendungen in einer akzeptablen Arbeitsgeschwindigkeit bedienen zu können und drittens muss die Möglichkeit zu einer schnellen und möglichst unterbrechungsfreien Erweiterung bei steigender Nutzung der Web-Anwendungen bestehen.

Um diese Herausforderungen zu bewältigen sind in den letzten Jahren Cluster-Technologien entwickelt worden, die jeweils über spezifische Vor- und Nachteile verfügen. In unserem ‚MySQL-Cluster‘-Projekt haben wir uns verschiedene Lösungen angeschaut.

Vergleich der Cluster-Technologien: Warum die Auswahl auf Galera fiel

Ein Cluster ist ein Verbund mehrerer Server, die zum einen untereinander und zum anderen in ihren Web-Anwendungen vernetzt sind. Bei der Auswahl einer Cluster-Technologie für MySQL stehen zur Zeit vier Kandidaten zur Auswahl:

- Master-Master-Replication Manager
- Lösungen mit Shared Storage
- MySQL-NDB-Cluster
- Galera

Der *Master-Master-Replication-Manager* basiert auf einem Verbund von Servern, von denen einer der Master ist, die anderen sind die sogenannten Slaves. Die Daten werden vom Master asynchron auf die Slaves repliziert: Dieses Verfahren eignet sich vor allem für Anwendungen mit sehr vielen Lesezugriffen, denn sowohl der Master als auch Slaves können Daten liefern, die Replikation sorgt für Datenredundanz. Bei einem Ausfall des Masters sind allerdings die Daten auf den Slaves nicht immer synchron, und beim Neustart des Masters kommen neuere Daten des alten Masters nicht mehr bei den Slaves an.

Bei den *Lösungen mit Shared Storage*, etwa SAN (Storage Area Network) oder DRBD (Distributed Replicated Block Device), gewährleistet das Storage-System die Hochverfügbarkeit. Deshalb muss MySQL so konfiguriert sein, dass jedes Commit (= ...) sicher in der Speicherschicht ankommt – es reicht nicht, dass es im Cache des Dateisystems steht. Anders als bei der Replikation

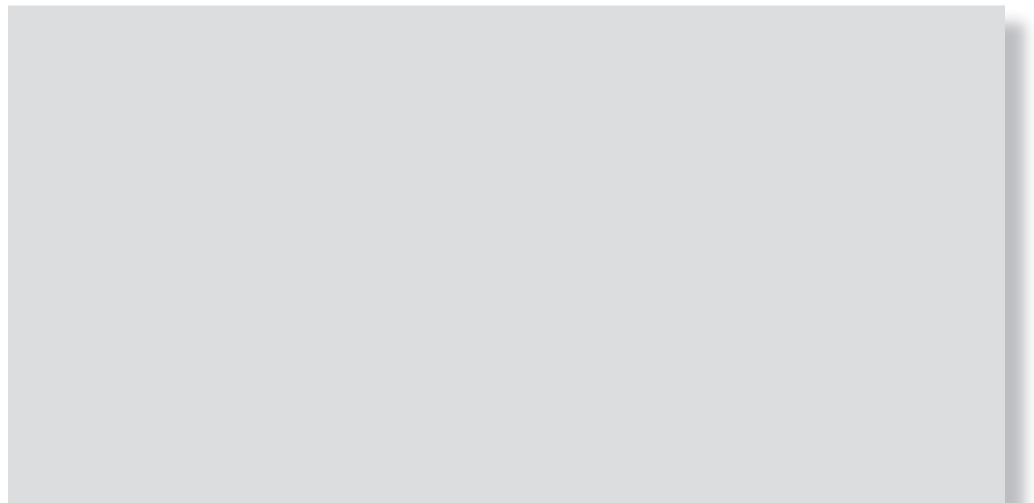
sind Slaves nicht einmal zum Lesen nutzbar. Kommt es bei einem Datenbank- oder Serverabsturz zu einem Failover, führt der startende MySQL-Daemon erst ein Recovery der Datenbank durch. Das dauert und ist nicht zwangsläufig erfolgreich. Bei einem Fehler ist kein Failover möglich, und der ganze Cluster steht.

Der *MySQL-NDB-Cluster* besteht aus Storage- und API-Nodes (Node = Knoten). Die Storage-Nodes bilden sogenannte Node Groups, die ihre Daten gegenseitig spiegeln.

Es findet eine zweistufige synchrone Replikation statt: Transaktionen werden zuerst auf den Nodes bis zum Commit ausgeführt. Die Knoten melden den Vollzug an den Master. Liegen die Vollzugsmeldungen von allen Nodes vor, schickt der Master den Nodes die eigentliche Aufforderung zum Commit.

Den Ausfall eines Node fängt der Partner auf, und beim erneuten Start synchronisiert ein Knoten seine Daten online vom Partner. Auch das Hinzufügen neuer Pärchen kann online geschehen. Statt „online“ ist häufig

Das jüngst am URZ installierte Galera-MySQL-Cluster hat folgenden Aufbau (Screenshot: U. Krämer):



ERFAHRUNGSBERICHT AUS DEM E-LEARNING-CENTER: INSIDE A „MOODLE MOOC“

Wie geht es in einem MOOC zu? Was lernt man dort? Und vor allem, was hat man zu erwarten, wenn man an einem kanadischen MOOC mitmacht, dessen Veranstalter aus dem Umfeld der konnektivistischen cMOOC-Linie stammen? Mit diesen Fragen im Hintergrund nahm Ingrid Dethloff vom E-Learning-Center der Universität Heidelberg, im Juni 2013 an dem vierwöchigen „Moodle MOOC on WizIQ“ teil. Sie berichtet:

Von Beginn an herrschte eine lockere freundliche Atmosphäre in diesem international besuchten MOOC mit der Arbeitssprache Englisch und Teilnehmern aus den unterschiedlichsten Ländern, Kulturen und Zeitzonen. In den insgesamt sieben Live-Sessions (der maximal erreichte Wert waren rund 200 Teilnehmer) gab es Impulse von bekannten Experten wie zum Beispiel dem Moodle-Entwickler Martin Dougiamas oder den MOOC-Pionieren Dave Cormier und Stephen Downes. Viele der aktiven Teil-

nehmer kamen aus dem Hochschul- oder Schul-Umfeld („Educators“) und hatten bereits große Vorerfahrungen mit Moodle und Online-Lernen.

Verschiedene Plattformen

Für den MOOC wurden von den Veranstaltern drei Plattformen zur Verfügung gestellt: WizIQ für die Live-Sessions und Ankündigungen, eine Moodle-Plattform mit den wöchentlichen Aufgaben (jeweils ein Kurs für Anfänger, erfahrene Dozenten und Moodle-Manager) und eine Moodle 2.5-Plattform zum Ausprobieren. Die wirklich zahlreichen Moodle-Foren und Diskussionen waren für mich dann auch der Kern des Kurses. Der Kursplan war aus meiner Sicht ausgesprochen ambitioniert – sowohl den Zeitplan betreffend als auch den inhaltlichen Umfang, der weit über die Thematik Moodle hinausging. Die erste Woche war der Oberfläche von Moodle gewidmet, dann ging es schon weiter mit „Creating Engaging Learning Ac-

tivities“ über die Thematik Online-Lernen bis hin zur Unterrichtsgestaltung in einem Webkonferenz-System (WizIQ).

Exzellente Betreuung

MOOCs sind ja dahingehend offen, dass man sich noch später im Kursverlauf einschreiben kann (hier waren es rund 2000 Teilnehmer) und anhand selbst definierter Lernziele auswählt, was man im Kurs mitmacht. Wenn man das dann mit einem engen Zeitplan und gar Gruppenaufgaben kombiniert, wird es sehr schwierig. Das war auch in diesem Fall so, wurde aber vielfach aufgefangen durch die unglaublich engagierte und motivierende Betreuung der Kursleiterin Dr. Nellie Deutsch. Aufgrund der vergleichsweise offen formulierten Aufgabenstellungen gingen die Teilnehmerbeiträge (meist waren dies „Digital Artifacts“ in Form von selbst erstellten YouTube Videos) in unterschiedliche Richtungen, was das gegenseitige Anschauen und Feedback-Geben sehr interessant machte.

Fazit

Wer in einem cMOOC mitmacht, sollte nicht erwarten, einen durchgestylten Kurs mit Videos vorzufinden, in denen feste Lerninhalte komprimiert vermittelt werden und man alleine seinen Weg durch Selbsttests und Quizzes gehen kann (Bei dieser Zielsetzung wäre dann ein xMOOC zu empfehlen, z.B. via Plattformen wie Coursera und edX). In einem cMOOC dagegen geht es um die Bereitschaft, mit anderen zu diskutieren und den Kursverlauf selbst mitzugestalten. Das kann zeitintensiv und mitunter etwas verwirrend sein – eigene Flexibilität und Frustrationstoleranz können daher nicht schaden. Meine Teilnahme am „Moodle MOOC on WizIQ“ mit durchaus 10 Stunden pro Woche an Abend- und Wochenend-Aktivitäten fürs „Certificate of Completion“ habe ich nicht bereut: Es war ein großartiges (Lern-)Erlebnis.

Ingrid Dethloff, ELC