

Online-Schema-Updates - Qualität und Quantität

Mario Rieß
ePages GmbH
Jena

Schlüsselworte

keine geplanten Ausfallzeiten, Schemaänderungen, Performance, MySQL 5.6 Online DDL

Einleitung

Die Anforderung eines Großkunden heißt: Es gibt keine geplanten Ausfallzeiten zur Software-Wartung. Mithin müssen auch alle DB-Änderungen, auch Schema-Änderungen, im laufenden Betrieb durchgeführt werden. Dabei sollen sowohl das Lesen als auch Schreiben bei möglichst geringen Einschränkungen der Performance gewährleistet sein. Das komplette Kopieren von Tabellen soll vermieden werden.

Theorie & Lösungsauswahl

Von Version zu Version der Software müssen Schema-Änderungen mit DDL Befehlen wie „ADD INDEX“ oder „ADD COLUMN“ ausgeführt werden. Währenddessen muss der volle Leistungsumfang des Systems gewährleistet sein. D.h. Lesen und Schreiben der DB bei guter Performance. Aufgrund der Größenordnung der Datenbestände soll das Ganze möglichst ohne Tabellenkopien vonstatten gehen.

Als Lösungsmöglichkeiten wurden Online-DDL-, Hochverfügbarkeits- und Replikationslösungen untersucht.

Die MySQL 5.6 Online DDL stellte sich dabei als am besten geeignet heraus.

Qualität - Funktionsumfang & Grenzen

Auch MySQL 5.6 deckt nicht den kompletten DDL-Befehlssatz gemäß der o.g. Anforderungen ab. Z.B. erlaubt der Befehl „ADD FLUTET INDEX“ kein gleichzeitiges Schreiben. Auch beim Ändern des Datentyps gibt es Einschränkungen. Diese können teilweise durch einen speziellen Patch-Ablauf aufgefangen werden.

Bei den Tests stellte sich heraus, dass man auf die Angabe der Optionen ALGORITHM (COPY, INPLACE) und LOCK (SHARED, NONE) verzichten kann, ohne Nachteile zu haben. Das System findet die optimale Variante ohne messbaren Zeitverlust.

Quantität

Bei ePages gibt es, bedingt durch die Scrum-Methode, alle zwei Wochen ein Release. Alle Performance-Betrachtungen wurden zunächst an einer kleinen DB getestet und danach an einer Real-DB durchgeführt, wobei es hauptsächlich um drei Tabellen mit 9 Mio., 1,4 Mio. und 0,6 Mio. Einträgen ging.

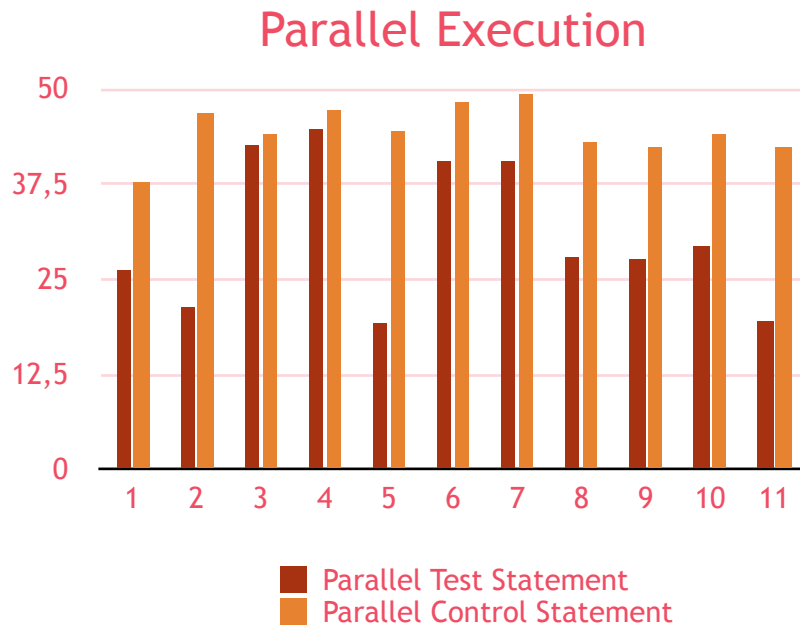


Abb. 1: Ergebnisse der Tests mit der kleinen DB - UPDATE-Operationen

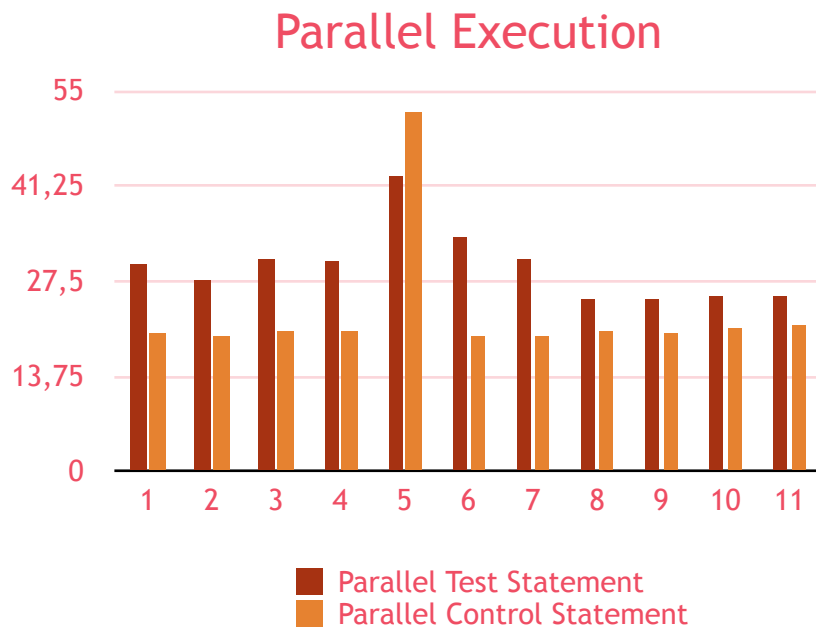


Abb. 2: Ergebnisse der Tests mit der kleinen DB - SELECT-Operationen

Zusammenfassend kann man die Ergebnisse so bewerten:

- Parallele Ausführung verlängert schreibende Befehle
- Folgende Befehle scheinen schreibende Befehle warten zu lassen, also sequenzielle Abarbeitung:
 - Add/Drop FullText Index
 - Add/Drop Foreign Key
 - Change DataType
- Parallele Ausführung hat weitestgehend keinen negativen Einfluss auf lesende Befehle
- Ausnahme bei komplexen SELECT Statements: Add/Drop Foreign Key

Im größeren Rahmen gab es drei verschiedene Test-Szenarien:



Abb. 3: Test-Szenarien - großer Rahmen

Ein erster Blick auf die Ergebnisse zeigt dieses Bild:

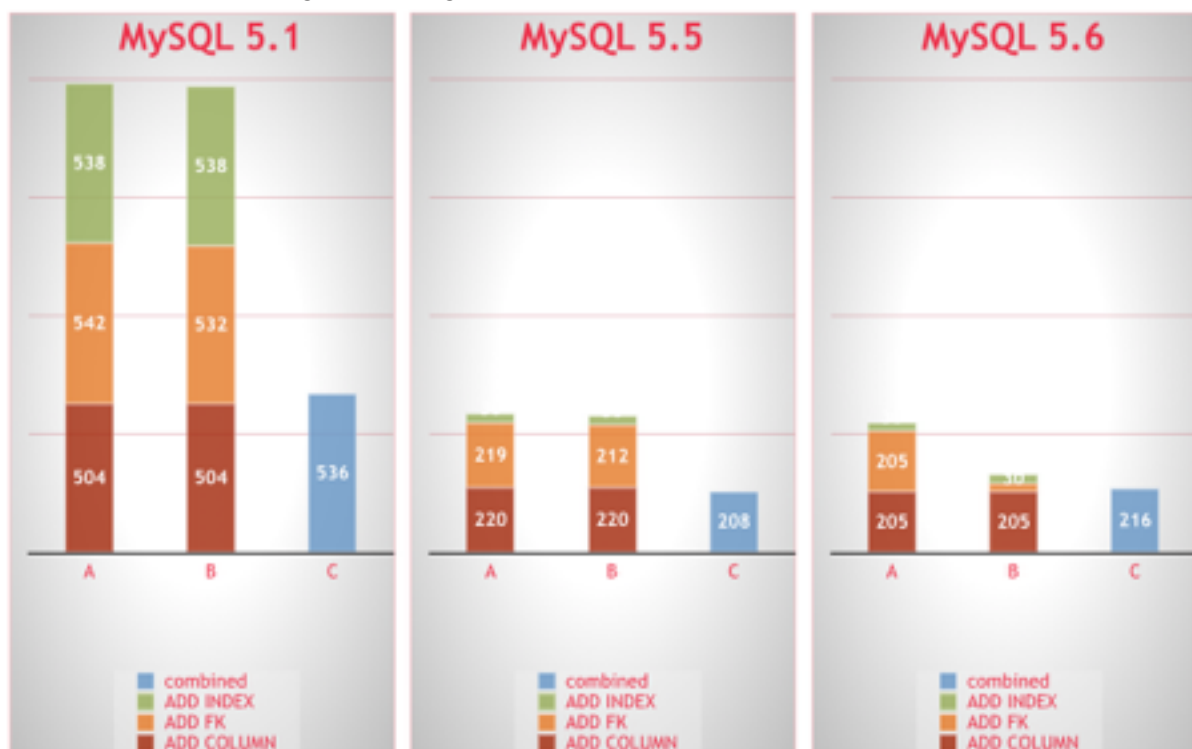


Abb. 4: Test-Ergebnisse - Ausführungszeiten DDL-Befehle

Daraus ergibt sich zunächst noch kein signifikanter Unterschied zwischen MySQL 5.5 und 5.6. Warum also auf jeden Fall 5.6?

Dies wird beim Blick auf aufgetretene Fehler, Log-Einträge sowie die Laufzeiten der Nutzer-Tests und des Patches klar (siehe Abb. 5).

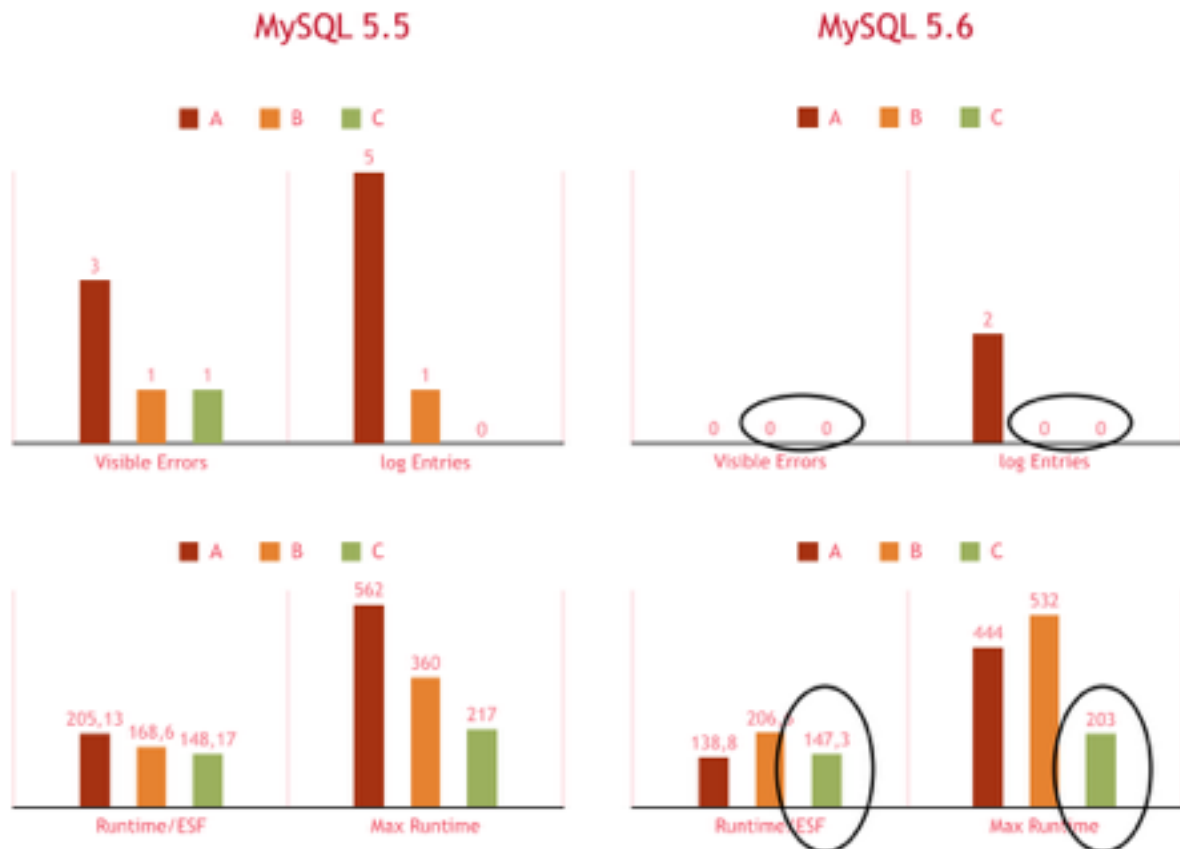


Abb. 5: Test-Ergebnisse - Fehleranzahl, Log-Einträge und Laufzeiten Nutzer-Tests und Patch

Im Einzelnen bedeutet dies:

- Minimale Unterschiede in den Laufzeiten von 5.5 zu 5.6 sowohl komplett als auch Änderungen der DB einzeln
- Analyse der ESF-Tests und Logs: eindeutige Verbesserung des Fehlerverhaltens von 5.5 zu 5.6
- In Bezug auf Fehlerverhalten sind B und C in 5.6 die besten Methoden
- In Bezug auf Laufzeit der ESF Tests ist Variante C in 5.6 die beste Alternative

Fazit

Mit MySQL 5.6 ist ein großer Schritt nach vorn gelungen. Sowohl der unterstützte Befehlsumfang als auch die Performance verbesserten sich wesentlich.

Für manche Operationen sind stufenweise Patches erforderlich. Vorbereitende DB-Änderungen werden in einem Zwischen-Patch ausgeführt, die eigentlichen Features nachträglich eingeschaltet. Zur Verbesserung der Ausführungszeiten sollten die DDL-Anweisungen zusammengefasst werden.

Kontaktadresse:

Mario Rieß
ePages GmbH
Heinrich-Heine-Straße 1
D-07749 Jena

Telefon: +49 (0) 40-350 188-127
Fax: +49 (0) 40-350 188-111
E-Mail: mriess@epages.com
Internet: www.epages.com