DB Weekly

Database Weekly

Issue 120 — September 2, 2016

FEATURED

MyRocks: A Space- and Write-Optimized MySQL Storage Engine — Facebook's
Yoshinori Matsunobu looks at how deploying MyRocks (a key/value store forked from Google's LevelDB) within a MySQL installation enabled a 50% reduction in storage requirements.
FACEBOOK CODE

A Comparison of 10 Open Source Time Series Databases — The author looks at and
ranks DalmatinerDB, InfluxDB, Prometheus, Riak TS, OpenTSDB, KairosDB, Elasticsearch, Druid, Blueflood and Graphite.
STEVEN ACREMAN

Tracking The Ever-Shifting Big Data Bottleneck — *"Bottlenecks are a fact of life in IT. No*
matter how fast you build something, somebody will find a way to max it out." And it seems newer in-memory frameworks may make CPU shared caches the next bottleneck for big data systems.
ALEX WOODIE

Forget Technical Debt - Build Technical Wealth — First Round Review
recently sat down with Andrea Goulet, CEO of Corgibytes, to discuss how "software remodeling" can help you pay down tech debt and tackle even the most difficult legacy codebases.
CORGIBYTES SPONSORED

Build a Raspberry Pi Hadoop Cluster to Run Spark on YARN — The Raspberry Pi isn't the
most natural platform for Hadoop or Spark work but it can provide a fun way to learn and this is a thorough walkthrough.
DQYDJ

Building A Recommendation Engine with AWS Data Pipeline, Elastic MapReduce and Spark — From Google's advertisements to Amazon's product suggestions, recommendation engines
are everywhere.
HUBBA

Apache Spark at Scale: A 60 TB+ Production Use Case — Facebook is a heavy user of
analytics for data-driven decision making and this post looks at how they use Spark for doing real-time entity ranking at scale.
FACEBOOK CODE

JOBS

IN BRIEF

IN BRIEF

[Over 18,000 Redis Instances Targeted by Fake Ransomware](#)
DUO LABS `NEWS`

[JSON Support Now Generally Available in Azure SQL Database](#)
MICROSOFT `NEWS`

[Azure SQL Data Warehouse General Availability Expanding to 18 Regions Worldwide](#)
PREM PRAKASH `NEWS`

[MariaDB 10.1.17 and MariaDB Galera Cluster 10.0.27 Now Available](#)
DANIEL BARTHOLOMEW `NEWS`

[Installing DreamFactory from source on Ubuntu 14.0.4 LTS](#) — DreamFactory auto-generates a rich API platform from any database. Here's a quick tutorial for a common Linux flavor.
DREAMFACTORY `SPONSORED`

[When To Avoid JSONB In A PostgreSQL Schema](#)
DAN ROBINSON

[A (Managing) Elasticsearch Cheat Sheet](#)
FRED DE VILLAMIL

[Valuing User Subscriptions Over Time with SQL](#)
PERISCOPE

[A Practical Guide to PostgreSQL Optimizations](#)
OMAR BOHSALI

[How A Japanese Cucumber Farmer Is Using Deep Learning and TensorFlow](#) — Uses of machine learning and deep learning are only limited by our imaginations: A farmer can use deep learning to sort cucumbers. See how.
GOOGLE DEVELOPERS

[How Mail.Ru Uses Tarantool to Help Scale Its Anti-Spam Service](#) — Tarantool is a NoSQL database running in a Lua application server.
HIGH SCALABILITY

[How Big Data Can Optimize Global Shipping](#)
DATANAMI

[Using JSON in SQL Server 2016 and Azure SQL Database](#)
CHANNEL 9 `VIDEO`

[Your Relational Database Management System Is Underutilized](#) — If you're just storing data, you're missing out says the author.
JASON PORRITT

[Spark Comparison: AWS vs. Google Cloud Platform](#) — Assessing cost, performance, and run time of a typical Spark workload.
MICHAEL LI AND ARIEL MNDANGE-PFUPFU

[Email API from SendGrid](#) — Reliably deliver your emails with a quick and simple API or SMTP integration. Try for Free
SENDGRID `SPONSORED` `TOOLS`

Want to subscribe? Enter your address here

Subscribe now »

Easy to unsubscribe at any time. Your e-mail address is safe — here's our privacy policy.

Want to subscribe? Enter your address here

Subscribe now »

Easy to unsubscribe at any time. Your e-mail address is safe — here's our privacy policy.