

[« Prev](#)[Next »](#)

Database Weekly

Issue 185 — December 22, 2017

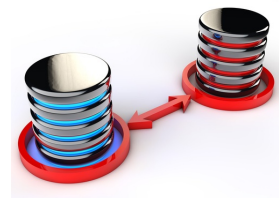
In this final issue of 2017, we're looking back at what new database systems were released, the best case studies/stories, and the biggest releases of 2017. We'll be back on January 12, 2018 - thanks for your support and we hope you have a happy holiday season! :-)

Roundups

DB-Engines Popularity Ranking of Database Systems — Since December 2016, PostgreSQL, Elasticsearch, MariaDB, and Azure Cosmos DB have shown the biggest gains. Oracle, MySQL, SQL Server have had the biggest falls relatively, but still top the chart.
DB-ENGINES

Developers' Most Loved, Dreaded, Wanted and Used Databases — Insights from 64,000 developers who took Stack Overflow's latest survey show Redis, Postgres, and Mongo as the 'most loved' databases, Oracle the most dreaded, but MySQL and SQL Server the most used overall.
STACK OVERFLOW

MySQL Replication Tutorial For Disaster Recovery — This blog post is a step by step tutorial on how to set up MySQL Replication between AWS regions. This is an essential part of our disaster recovery plan at Engine Yard. A previous [blog post](#) gives a higher level overview on disaster recovery.
ENGINE YARD SPONSORED



A Comparison of Advanced, Modern Cloud Databases — A non-exhaustive primer of modern cloud database solutions like Aurora, Cosmos, and Spanner.
BRANDUR LEACH

A Look at the Graph Database Landscape — Graph databases are the fastest growing category in all of data management, according to DB-Engines.com.
DATANAM

Big Releases

PostgreSQL 10 Released — The popular open source database includes native logical replication, declarative table partitioning, and improved query parallelism. [More on what's new here.](#)
POSTGRES.ORG

MongoDB 3.6 Released: Security, Robustness and JSON Schema — A new version of the popular NoSQL database: better hardened against network outages with 'Retryable Writes', as well as against ransomware by only binding to localhost by default. It now also supports JSON Schema for data validation.

MONGODB

See What's New in MongoDB 3.6 — Read the guide or try 3.6 right away.

MONGODB

SPONSORED



Redis 4.0 Released — The popular data structure server took a step forward with several key improvements including a new replication engine and official support for modules.

SALVATORE SANFILIPPO

TOOLS

SQL Server 2017 Released: What's New? — A look at the year's big SQL Server release with new features from Python support to adaptive query optimization and a built-in graph database.

MICROSOFT

Elasticsearch 6.0.0 Released — The popular full-text search-oriented database gains zero downtime upgrades, faster restarts, faster query times, and more. The new version is based on Lucene 7.

CLINTON GORMLEY

Apache Kafka Goes 1.0 — Billing the popular event streaming platform 'enterprise capable', its creators reflect on its history and feature set.

CONFLUENT

BigchainDB 1.0: A Scalable Blockchain Database — BigchainDB is a decentralized database with blockchain characteristics, at scale. GitHub repo.

TIM DAUBENSCHÜTZ

CockroachDB 1.0: A Production-Ready Go-Based SQL Database — One of the "NewSQL" generation of databases, CockroachDB is an open source, distributed SQL database designed for high availability.

SPENCER KIMBALL

TOOLS

MapD Open Sources Its GPU-Powered Database — Mark Litwintschik has also written a handy guide to compiling MapD from source.

MAPD BLOG

Case Studies and Stories

Moving Yelp's Core Search to Elasticsearch — A post mortem of Yelp's successful migration to Elasticsearch from a custom system built on top of Lucene.

YELP ENGINEERING

\$20 Free On A New Linode Account — Linux cloud hosting starting at 1GB of RAM for \$5/mo. Get \$20 credit on a new account.

LINODE CLOUD HOSTING

SPONSORED

Architecture of Giants: Data Stacks at Facebook, Netflix, Airbnb, and Pinterest —

Simple event data infrastructure diagrams from several fast-scaling companies.

MICHELLE WETZLER

How Discord Stores Billions of Messages with Cassandra — Discord is a popular chat system for gamers. They started out with MongoDB but here they explain why and how they moved to Cassandra, and how they dealt with garbage collection issues.

STANISLAV VISHNEVSKIY

A Look at Instapaper's Outage Cause and Recovery Process — Bookmarking service Instapaper experienced a 30 hour outage. Here's the story of how their MySQL database failed, why, and how it was resolved.

BRIAN DONOHUE

PostgreSQL At 10TB and Beyond — Chris Travers discusses what happens when managing over 10 terabytes of data in PostgreSQL. Does it scale and what kinds of problems need to be resolved?

CHRIS TRAVERS **VIDEO**

Building a New Database Management System in Academia — Andy Pavlo of Carnegie Mellon University explains why his group had to build a new database system (Peloton) rather than use an existing one.

ANDY PAVLO

Writing a Time Series Database from Scratch — Prometheus is an open source monitoring tool that includes a custom time series database. This is a deep dive into fleshing out a new architecture to address the existing database engine's shortcomings.

FABIAN REINARTZ

Elasticsearch Cluster Lifecycle at eBay — A look at some of what's involved with streamlining the rollout and management of Elasticsearch clusters.

SUDEEP KUMAR

Publishing with Apache Kafka at The New York Times — How Apache Kafka and its Streams API are used for storing and processing all the articles published by the NYT.

BOERGE SVINGEN

New Database Systems

Timescale: An Open Source Time-Series Database — SQL made scalable for time-series data. It's Postgres compatible and optimized for fast ingest & complex queries.

TIMESCALE

Microsoft Unveiled Azure Cosmos DB, A New Multi-Model DB — An Azure-based globally-distributed multi-model database built for low latency, elastic scalability, and high availability.

MICROSOFT **TOOLS**

Cloud Spanner: A Global Database Service from Google — A globally-distributed relational database service with ACID transactions and SQL semantics. WIRED's writeup gives the bigger picture of why this was a big deal.

GOOGLE

YugaByte: A New Open Source, Cloud-Native Database — From the team that built Facebook's internal NoSQL platform comes a new database for mission-critical applications which supports both the Cassandra and Redis APIs. [GitHub repo](#).

YUGABYTE BLOG

RocksDB: A Persistent Key-Value Store for Flash and RAM Storage — From Facebook came a library forming the core of a fast memory-based key-value store.

FACEBOOK [TOOLS](#)

Peloton: The Self-Driving Database Management System — An in-memory, DRAM/NVM optimized, relational database management system designed with autonomous operation and optimization in mind.

CARNEGIE MELLON UNIVERSITY DATABASE GROUP [TOOLS](#)

JanusGraph: An Open-Source, Distributed Graph Database — A highly scalable transactional graph database optimized for storing and querying large graphs with billions of vertices and edges distributed across a multi-machine cluster.

JANUSGRAPH [TOOLS](#)

AgensGraph: A Transactional Graph Database based on Postgres — An open source, multi-model database which supports both relational and graph-based data models, even supporting both SQL and openCypher within the same query.

BITNINE

GryadkaJS: A Paxos-Based Replicated Key/Value Layer On Top of Redis — Gryadka is a minimalistic Paxos-based master-master replicated consistent key/value layer on top of multiple instances of Redis.

DENIS RYSTSOV [TOOLS](#)

Tarantool: The Good, The Bad and The Ugly — A practical demonstration of [Tarantool](#), an open source in-memory NoSQL database and Lua app server.

VADIM POPOV

PumpkinDB: An Event Sourcing DB Engine That Doesn't Overwrite Data — A compact event sourcing database, featuring on-disk storage, flexible approach to event structure & encoding, sophisticated event indexing and querying. Open source, written in Rust.

PUMPKINDB.ORG [TOOLS](#)

GeoMesa: An Open-Source Spatio-Temporal Database Layer — GeoMesa provides spatio-temporal indexing on top of Accumulo, Bigtable, and Cassandra, as well as near real-time stream processing and spatial semantics on top of Kafka.

THE GEOMESA PROJECT [TOOLS](#)

Badger: A High-Perf Key/Value Storage Engine — An embeddable, persistent, simple and fast key-value store, written natively in Go. [GitHub repo](#).

DGRAPH [CODE](#)

TigerGraph Emerges with Native Parallel Graph Database — A startup named TigerGraph has emerged from stealth with a new native parallel graph database its founder thinks can shake up the analytics market.

ALEX WOODIE

ProfaneDB: A Protocol Buffers Database — Based on RocksDB and can be used with any language that supports gRPC.

GIORGIO AZZINNARO [CODE](#)

TiDB 1.0: A Distributed NewSQL Database for Analytics — An open source distributed Hybrid Transactional/Analytical Processing (HTAP) database written in the Go language.

PINGCAP

TileDB: A Multi-Dimensional Array Data Management System — A database that started life at MIT and Intel that's designed for storing massive dense and sparse multi-dimensional array data.

TILEDB, INC. [TOOLS](#)

Amazon Aurora Serverless: Databases on Demand — Aurora is AWS's MySQL- and Postgres-compatible scalable database service and they're now working on a pay-as-you-go variant for highly variable workloads.

AMAZON WEB SERVICES

Hello Memgraph: A Real-Time Transactional Graph Database — Designed for the 'artificial and machine intelligence' era, Memgraph is focused on speed and scale and is available in an early access 'Community Edition'.

DOMINIK TOMCEVIC [TOOLS](#)

Amazon Neptune: A Fully Managed Graph Database Service — A big week for graph databases, it seems, as Amazon announces a new graph database service for AWS users.

AMAZON WEB SERVICES

[« Prev](#)

[Next »](#)

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](#).