

: *What are post-relational databases?*

by Daniel Power

Editor, DSSResources.com

Post-relational or non-relational databases implement a conceptual schema using a non-relational data model. The term post-relational refers to databases intended to support an alternative logical data model to the transaction processing oriented relational model. Non-relational databases were developed both before and after relational databases like Oracle and DB2 were programmed and marketed in the late 1970s.

A database is developed from a diagram of concepts called a conceptual schema. A conceptual schema is a high-level description of a organization's information needs. A conceptual schema may be mapped to various logical data models. When data is modeled and stored in a format that differs substantially from the relational model of relations and primary and foreign keys, then the database can be characterized as a post-relational database. Examples include multidimensional databases, column databases, object-relational databases, and flat file databases.

Post-relational databases can store both "data in motion" and "data at rest", also called static data. Data in motion is collected and analyzed in real-time. For example, data collected to analyze the quality of products during the manufacturing process is data in motion. An example of data at rest would be historical data about all aspects of the customer relationship, including sales transactions, social media data, and customer service interactions, cf., Hurwitz, Nugent, Halper and Kaufman (2013). The customer service data could then be analyzed by a business analyst to better understand customers' buying patterns and preferences.

The post-relational databases developed since 2004 have become known as NoSQL databases and they include key-value data stores, network and graph databases and document-oriented databases. NoSQL (sometimes expanded to "not only SQL") is a broad class of database management systems that differ from the classic model of the relational database management system (RDBMS) in some significant ways. A NoSQL database provides a mechanism for storage and retrieval of data that uses less structured consistency models than traditional relational databases. The most popular NoSQL database systems include: MongoDB, Riak, Oracle NoSQL Database, Redis, CouchDB, Apache Cassandra and HBase. XML databases are a type of structured document-oriented database that allows querying based on XML document attributes.

The development of post-relational databases means data can now be stored in a wider variety of databases and formats. SQL relational databases remain very important for storing related, structured data at rest. NoSQL databases, improved computing technologies, distributed databases and Cloud computing and storage have made it possible to store and retrieve high volume, high

: *What are post-relational databases?*

variety and high velocity data, so-called “big data”.

A NoSQL database stores and retrieves data with different consistency requirements than traditional relational databases. Motivations for using this approach include simplicity of design, horizontal scaling and finer control over availability. NoSQL databases are often highly optimized key–value stores intended for simple retrieval and appending operations, with the goal being significant performance benefits in terms of latency and throughput. NoSQL databases are finding significant and growing industry use in big data and real-time web applications.

NoSQL Databases cannot provide full ACID (Atomicity, Consistency, Isolation, Durability) guarantees for transactions. Usually eventual consistency is guaranteed or transactions are limited to single data items. This means that given a sufficiently long period of time over which no changes are sent, all updates can be expected to propagate eventually through the system. NoSQL DB provide BASE guarantees: 1) Basic Availability, 2) Soft-state, and 3) Eventual consistency.

Many systems fall under the NoSQL category. NoSQL databases are next generation databases that are in general non-relational, distributed, open-source and horizontally scalable, cf., <http://nosql-database.org/>. Development of the graph database Neo4j was started in 2000. CouchDB research was started in 2005 and provides a document database inspired by Lotus Notes. The project moves to the Apache Foundation in 2008. Development for Google BigTable was started in 2004 and the research paper by Chang et al (2008) describing the system was published a few years later.. Facebook open sourced the Cassandra DB project in 2008. HBase is a BigTable clone for the Hadoop project.

Computerworld magazine (June, 2009) reported in an article about the NoSQL meet-up in San Francisco that “NoSQLers came to share how they had overthrown the tyranny of slow, expensive relational databases in favor of more efficient and cheaper ways of managing data.” The article explained that especially Web 2.0 startups have begun their business without Oracle and even without MySQL which formerly was popular among startups.

Post-relational databases use multiple technologies. There is one type of relational database, a SQL database, with minor variations, but there are many different types of NoSQL databases including key-value stores, document databases, wide-column stores, and graph databases. SQL databases were developed in 1970s to deal with first wave of data storage applications. NoSQL DB developed in 2000s to deal with limitations of SQL databases, particularly concerning scale, replication and unstructured data storage. These next generation databases are non-relational, distributed, open-source and horizontally scalable.

: *What are post-relational databases?*

Post-relational or NoSQL databases help store large volumes of varied data like free form text and photos in a distributed, Cloud storage and retrieval-only environment. Relational databases will continue to handle transaction processing and data warehouse queries. These relational data stores will use normalized or denormalized schemas and parallel processing to provide much of the business intelligence decision support.

References

Chang, F., J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach
M. Burrows, T. Chandra, A. Fikes, R. E. Gruber, "Bigtable: A Distributed Storage System for Structured Data," ACM Transactions on Computer Systems (TOCS), Volume 26 Issue 2, June 2008.

Computerworld, "No to SQL? Anti-database movement gains steam," June 2009, at URL http://www.computerworld.com/s/article/9135086/No_to_SQL_Anti_database_movement_gains_steam

Hurwitz, J., A. Nugent, F. Halper and M. Kaufman, Big Data For Dummies, 2013.

Strauch, C., "NoSQL Databases," at URL <http://www.christof-strauch.de/nosql dbs.pdf>

Author: Daniel Power

Last update: 2016-08-20 05:26