Overhyped Graph Database Valuations

by, David Floyer

February 17th, 2021

Graph Database

IT executives can tolerate some diversity of databases in small exploratory projects. However, if these projects become stable and larger in scope, they should be migrated to a converged database. This will allow IT and lines of business a much greater ability to access all data directly, leading to much faster real-time or near-real-time applications with greater scope to automate business processes.
**Graph Database**

A Graph Database is a well-established NoSQL database type, allowing programmers to quickly express and query data networks. There are nodes and connections between the nodes, as shown in Figure 1 below.

![Figure 1 - Graph Database Illustration of a Network of Nodes and Connections](image)

*Source: Neo4j, 2013*

The nodes a varied, and the connections are varied. They are handy for applications such as fraud detection and recommendation engines. Programmers can find and compare connections quickly. The amount of data held is small and can often be held in memory. This leads to great performance over millions and billions of entries. Apart from TigerGraph, other well-known graph databases are Datastax Enterprise Graph and Neo4J.

**Graph-Database Limitations**

Graph databases do not do everything. They are one of many analytic databases that specialize in network-type data. Other databases specialize in document data (MongoDB), time-series (InfluxDB, Prometheus), log databases (Splunk), wide-column stores (Cassandra, HBase, key-value DB (Aerospike, Amazon DynamoDB), search (Elasticsearch), object-relational databases (PostgreSQL), columnar databases (Oracle), hierarchical (IBM IMS), blockchain DB (with immutable tables), etc., etc. NoSQL databases were all the fashion for a while (e.g., Couchbase), but almost all are SQL compliant these days because it requires much less effort for programmers to code.

**Single Purpose vs. Converged Database**

One of the problems with specialized databases is that it leads to complexity in sharing the data with other databases. The data has to be transformed, and this takes elapsed time and often loses context. If you have 16 database types, there are 120 different complex link procedures IT has to provide, manage, and maintain.
An important trend Wikibon is following is converged databases, where all the data is held in a common format. Microsoft Cosmos DB (new) and Oracle Database (well established with the best functionality) are examples of fully functional graph databases. This means that data can be shared and used without any transformations by other databases.

The data should also be retrieved by any person you permit to, inside or outside the company. Snowflake is really good at this and uses the term “Data-mesh” to describe this.

**Conclusions**

Wikibon does not understand how any single type of database can be worth billions. Microsoft and Oracle have adequate converged in-memory graph databases that be accessed directly and quickly from other analytic and OLTP databases. We don’t understand how TigerGraph will carve out a niche and differentiate itself long-term, even if it is faster than the speed of light.

**Action Item**

IT executives can tolerate some diversity of databases in small exploratory projects. However, if these projects become stable and larger in scope, they should be migrated to a converged database. This will allow IT and lines of business a much greater ability to access all data directly, leading to much faster real-time or near-real-time applications with greater scope to automate business processes.
David Floyer spent more than 20 years at IBM, holding positions in research, sales, marketing, systems analysis and running IT operations for IBM France. He worked directly with IBM’s largest European customers, including BMW, Credit Suisse, Deutsche Bank and Lloyd’s Bank. Floyer was a Research Vice President at International Data Corporation (IDC) and is a recognized expert in IT strategy, economic value justification, systems architecture, performance, clustering and systems software.