



www.datastax.com

3975 Freedom Circle, 4th Floor
 Santa Clara, CA 95054, USA
 Tel: +1 650 389 6000
 Email: info@datastax.com

DataStax DSE Graph

The company

DataStax is a database vendor that was founded in 2011. Its primary offering is DataStax Enterprise (DSE), the leading database built using Apache Cassandra. In 2015, the company acquired Aurelius, the chief developers of the Titan graph database, and have since leveraged their expertise to develop DSE Graph, a graph database add-on for DSE. In addition, DataStax is the leading contributor to Apache TinkerPop, the graph computing framework upon which various graph databases (including DSE Graph) are based.

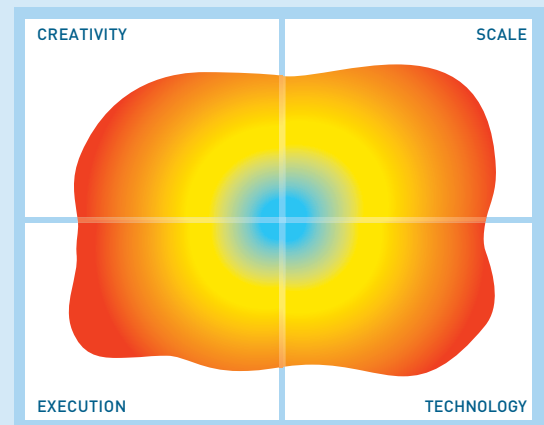
DataStax is headquartered in Santa Clara, CA, and has additional US offices in San Francisco, Austin, and Atlanta, as well as international offices in the UK, France, Germany, Japan and Australia. As of this writing, DataStax has more than 480 employees, 220 of which are dedicated engineering or product staff.

What is it?

DSE is a distributed database oriented towards (though not exclusive to) a hybrid-cloud architecture. It is built on top of Cassandra, but boasts numerous capabilities above and beyond what Cassandra alone offers, including native search and analytics, continuous availability, and significant increases to speed and performance.

DSE Graph is the graph database add-on for DSE. It is a property graph solution that is optimised for storing billions of items and relationships. It is suited for both transactional and analytical processing. In accordance with the latter, it also supports Spark-based analytics. It is available on-premises, in-cloud, or as part of a hybrid solution, and is additionally deployable as a Docker container.

The product originally existed as a bespoke version of the Titan database, optimised to run on the Cassandra database engine used within DSE. It has since been updated with a bevy of new features and capabilities, but retains basic compatibility with Titan, allowing for straightforward migration between the two. Moreover, as with Titan, DSE Graph is built to use



The image in this Mutable Quadrant is derived from 13 high level metrics, the more the image covers a section the better. Execution metrics relate to the company, Technology to the product, Creativity to both technical and business innovation and Scale covers the potential business and market impact.

Apache TinkerPop. Notably, DataStax is the primary contributor to TinkerPop, having contributed approximately 99% of TinkerPop's codebase.

What does it do?

DSE Graph is a property graph that is fully integrated with DSE. In fact, it relies on DSE (and Cassandra underneath it) as a data store. It also integrates with a number of built-in DSE capabilities, including DSE Search and DSE Analytics. In addition, it is highly scalable and performant, scaling up to billions of entities. In service to this, it leverages optimisation techniques such as query optimisation, data partitioning, and distributed query execution, among others.

Moreover, DSE Graph is designed for both transactional and analytical processing, and consequently features two processing engines – one transactional, one analytical – and allows for both OLTP and OLAP graph traversals. Moreover, for the purposes of OLAP, Gremlin (part of Tinkerpop), SQL and Spark APIs are supported, the latter including both batch and streaming. Furthermore, switching between engines (and therefore modes of traversal) is relatively simple, and can be done without altering the underlying data. This means that you can



| | |
|-------------|-------|
| Analytics | ★★★★ |
| Ease of Use | ★★★★★ |
| Features | ★★★★★ |
| Integration | ★★★★★ |

| | |
|-------------|-------|
| Language | ★★★★☆ |
| Operations | ★★★★★ |
| Performance | ★★★★★ |
| Scalability | ★★★★★ |

“Graph analytics is great for showing relationships between data points, and this can be very valuable in a healthcare scenario. By looking at data in different ways within the same platform, we can support more in-depth interactions with patients and improve healthcare outcomes.”
 Babylon Health

leverage transactional and analytic queries on a single set of data, as needed. In addition, analytical and transactional workloads are separated, and automatic workload management is available.

DSE Graph includes a variety of tools for managing all aspects of your graphs and graph clusters. This includes Lifecycle Manager and OpsCenter, which allow you to automate and visualise the creation of new graph clusters, respectively. However, the most important tool for interacting with DSE Graph might be the DataStax Studio, a visual, browser-based development environment for your graph. It supports Spark SQL, Gremlin, and CQL (Cassandra Query Language), and additionally comes with a built-in smart Gremlin editor, similar

to an RDBMS smart query editor. In fact, much of DataStax Studio is similar in feel to the visual development tools available in more conventional, relational environments. Moreover, to support the visualisation aspect of this tool, DataStax partners with a number of visualisation vendors, including Cambridge Intelligence, Tom Sawyer, Linkurious and Tableau (although the latter is a more general partnership, and not specific to DSE Graph).

Why should you care?

In general, the reason graph databases are worth caring about is that they perform well compared to more traditional databases for processing data that involves multiple, complex relationships. However, a graph database by itself can only do so much. In order to effectively address so-called graph problems, your graph must be embedded inside a full software stack that supports a wide range of capabilities, such as search, analytics, and so on. In other words, graph problems are bigger than just the graph database. This is where DataStax, with DSE and DSE Graph, excels, providing not only the graph database, but the full stack as well. Moreover, many of the benefits of DSE are carried over to DSE Graph – including continuous availability, hybrid-cloud deployment, scalability, and so on – and the two are well integrated, allowing DSE Graph to take advantage of a variety of capabilities that are available in DSE, including DSE Search and DSE Analytics. Moreover, DSE Graph in and of itself boasts some significant differentiators. This includes its dual processing engines, allowing you to easily switch between transactional and analytical processing, and DataStax Studio, a particularly impressive example of a visual development environment for graph.



Figure 1 – Monitoring the DataStax DSE environment

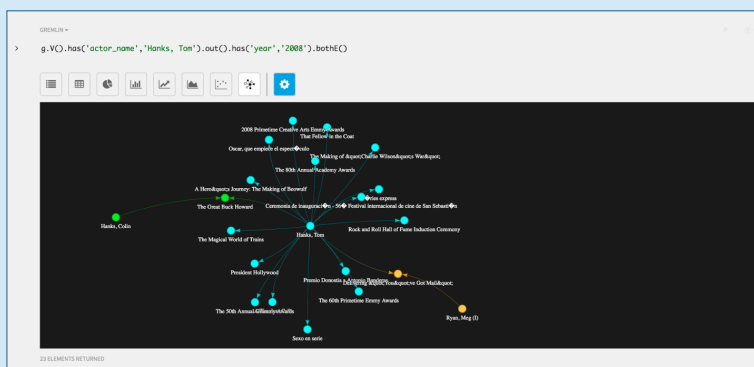


Figure 2 – DataStax Studio

The Bottom Line

Together with its parent platform, DSE Graph provides a complete and effective means of addressing graph problems, regardless of whether they are transactional or analytical in nature. If you already use DSE, DSE Graph makes for an excellent addition. If not, it provides a very good reason for doing so.

[FOR FURTHER INFORMATION AND RESEARCH CLICK HERE](#)