

# DataWalk vs. TigerGraph™

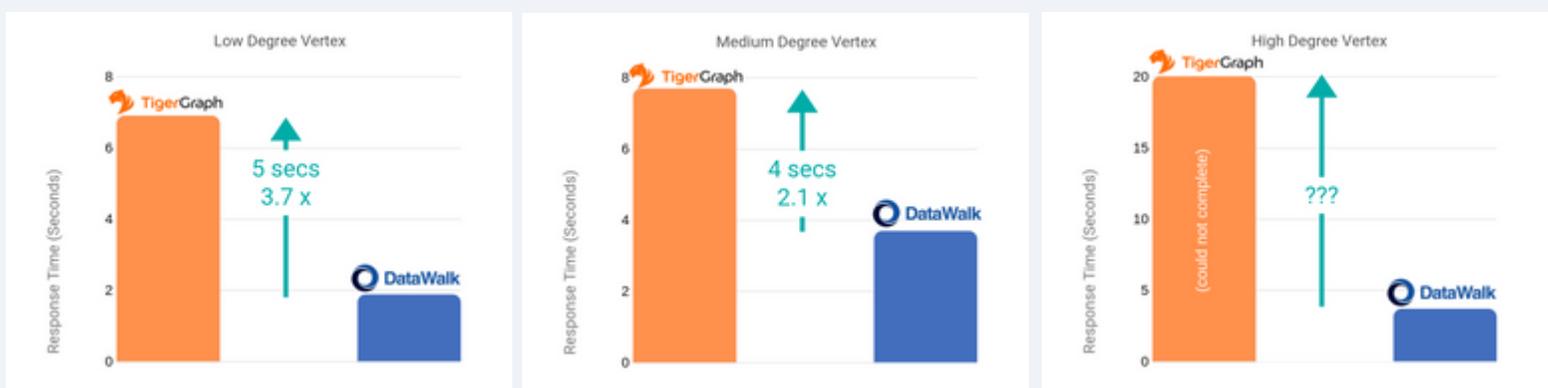
## Performance Benchmark: Find Paths Algorithm

### Executive Summary

Though DataWalk is a full-stack analytics platform and not a database, DataWalk does include facilities for graph analytics. In September 2019 DataWalk conducted performance benchmarks of the *find path* algorithm on both systems[1].

The first set of tests measured response time for a find path algorithm intended to reflect common real-world usage, where much is known about one vertex (object), but less is known about source vertices (objects) to which it is indirectly connected.

Tests were run for source vertices with low, medium, and high degree (i.e., few or many connections). In these tests all data (about 1B records) fits into available RAM on the database server, which is optimal for TigerGraph. Test results showed that DataWalk is 2.1 - 3.7X faster for low and medium degree vertices. DataWalk was also significantly faster for the high degree tests, though this could not be fully quantified as TigerGraph generated errors in 8 of the 10 tests.

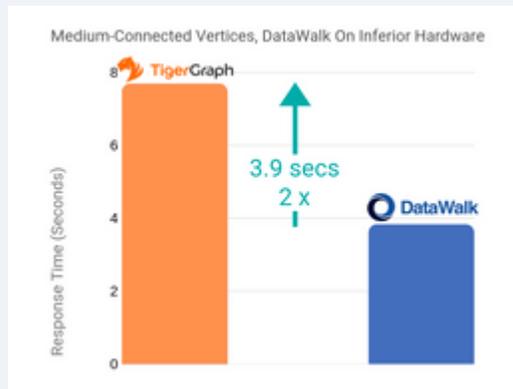
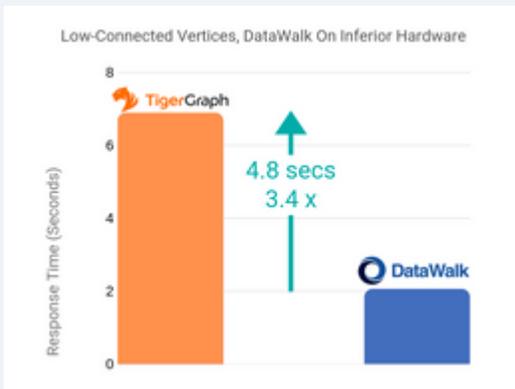
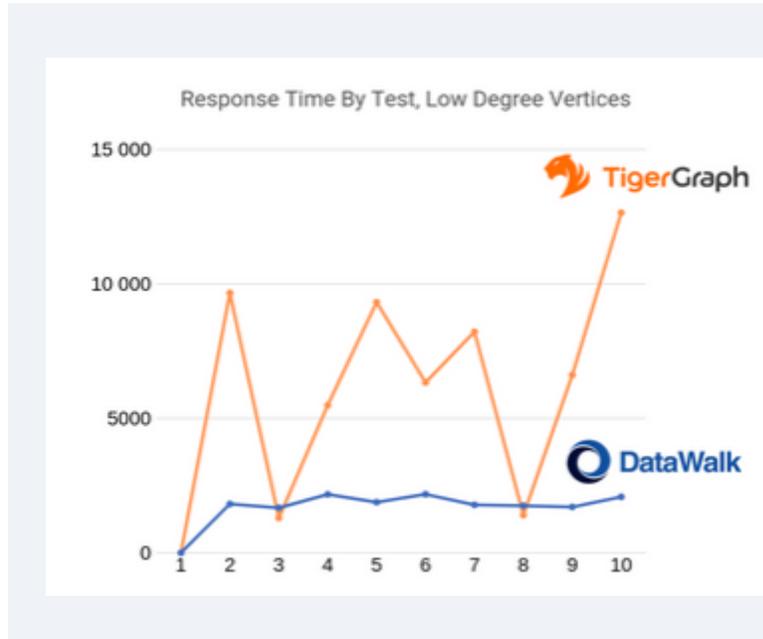


The second set of tests compared find paths performance between 50 random objects. Test results showed that DataWalk is dramatically faster (15.6X; an average of 38.2 seconds) if the graph does not fit into available RAM. This actually understates DataWalk's advantage, as TigerGraph timed-out on three of the 50 tests. When data did fit in available RAM, DataWalk was slower than TigerGraph by an average of 1.5 seconds:



Note that when the above tests were executed, the response time of DataWalk was relatively consistent, while response time of TigerGraph was highly variable. An example of this is shown at right, for the low degree vertex test.

During testing with a single user, it was also observed that TigerGraph typically required 60-80% memory utilization, while DataWalk memory utilization was roughly 2-4%. DataWalk's dramatically lower memory utilization and high levels of efficiency were also reflected when the above tests were re-run with DataWalk running on inferior hardware. Specifically, as shown below, when DataWalk was run on hardware that had half the number of vCPUs and half the amount of RAM, DataWalk still had a performance advantage of 2 - 3.4X, or roughly 4-5 seconds.



Also note that DataWalk can find paths between multiple objects and pre-defined groups, but as TigerGraph does not support this capability, a benchmark comparison was not possible.

For the full benchmark report, request a copy via [info@datawalk.com](mailto:info@datawalk.com).

DataWalk is a full-stack data analysis platform for revealing patterns, relationships, and anomalies for large-scale, multi-source data. DataWalk allows organizations to rapidly import and blend vast amounts of data from multiple sources into a singular data view, and then analyze data and connections via simple visual interfaces. For additional information, see [www.datawalk.com](http://www.datawalk.com).

TigerGraph is a trademark of TigerGraph Inc. DataWalk is a trademark of DataWalk Inc.

[1] Identical hardware was used unless otherwise indicated, though as DataWalk is a full stack system for data analysis, the DataWalk configuration also requires an additional server as an application server to support the broad array of functionality beyond the database.