

## Opinion: Four key considerations for evaluating graph warehouses



**Barry Zane**

Vice President of Engineering



Published by Information-Management.com on September 14, 2018

We are witnessing a surging interest in graph databases for data warehouses, and for good reason. The best of the graph databases offer the same analytic functionality as relational options while making life simpler and quicker. Graphs also provide analytic capabilities that relational technologies can't offer, such as ML and AI applications.

So, it's no surprise that a slew of vendors are now considering offering graph database options, making it imperative for prospects to realize one simple truth: graph use cases are not all the same. Some are more suited for Graph Online Analytical Processing (GOLAP), others for Graph Online Transaction Processing (GOLTP).

Organizations should consider issues of parallelism, read/write capabilities, query types (open-ended questions and the percent of data involved), and Business Intelligence features when distinguishing the needs of GOLAP versus GOLTP. These factors delineate whether you need transactional or analytics systems to maximize ROI in graph databases. Failing to seize that distinction could squander investments in this innovative technology on the wrong use cases.

### Parallelism

Parallel computing is the main architectural difference between OLTP systems and OLAP systems. While modern servers have dozens of CPU cores, clusters of servers can have thousands of CPU cores.



In a parallel system, every CPU core is working on a part of the data processed by the query, which makes every query run faster. In an OLTP system, adding CPUs does not make the individual queries faster because only one core is generally working on each query.

The loading of data leverages all the cores working together and the results can be thousands of times faster on parallel OLAP systems than traditional OLTP architectures.

## Read/Write Capabilities

Organizations should also evaluate graph databases in terms of how much reading and writing they'll require. GOLAP systems mostly read data for query purposes. Since data warehouses are usually batch jobs, their writing capabilities aren't as important as their ability to swiftly query data for answers. However, the performance of batch loading is crucial.

OLTP systems are constantly updating small portions of their transactional data via their writing capacity. For example, tollbooths are continually reading the license plates of vehicles and updating transactional data for passing motorists. Other examples include point-of-sale (POS) checkout systems, either for e-commerce or physical shopping locations. The same data from the tollbooths or consumer checkout is subsequently used by OLAP systems for establishing highway systems improvements, or pricing and marketing options for POS.

## Query Types

Another defining attribute of OLTP and OLAP systems is the type of query required. In general, OLTP systems are primed for answering narrow, well-defined questions. OLAP systems are designed for identifying patterns and trends. OLTP systems can readily pinpoint which products a customer has while OLAP systems can identify the most popular products your typical customer has.

In a relational database management system, one of the main analytic benefits is that users can ask several questions regarding relationships. For example, users can ask whether Dick and Jane are neighbors, co-workers, spouses, etc. However, unlike a graph system, users don't have the ability to ask an open-ended question such as 'how are Dick and Jane related.'

GOLAP systems excel by answering sophisticated open-ended questions against vast amounts of data. GOLAP systems allow you to ask these same complex questions, but against entire populations of Dick, Jane and everybody else, which makes them great for determining customer or product characteristics and finding other patterns. Open-ended queries often involve a greater percentage of available data than closed-ended ones do.

## BI Features

The differences in query types for OLAP and OLTP relate to the BI characteristics of analytic systems, which aren't necessarily feasible with transactional ones. GOLAP is ideal for complicated BI use cases in which users issue initial questions, then ask increasingly specific ones based on results. Often times, these queries involve aggregation before honing in on a specific area for more details.

With GOLAP systems, users can identify how many customers they have before categorizing that information according to region, daily trends, and daily trends for certain months. OLTP systems don't deliver that amount of breadth.

## Supporting Each Other

An examination of requirements for parallel processing, reading and writing, query types, and BI features denotes whether organizations need GOLAP or GOLTP systems for a given project. Still, these systems support each other in critical ways. The patterns and trends uncovered through GOLAP systems enable customer segmentation and micro-segmentation.

OLTP systems operationalize these results with real-time recommendations based on customer behavior revealed by GOLAP. E-commerce systems—in which the results of data warehouse analytics are used to issue recommendations for customers when checking out—illustrate a good example of this utility. In these instances, OLTP supports OLAP by providing a means of capitalizing on the latter's analytics, while OLAP supports OLTP by performing the computations for such low latency action.

These developments are exciting because they demonstrate how graph databases innately enhance some of the richest use cases of analytics in production today and how their impact can propel the industry forward.

<https://www.information-management.com/opinion/four-key-considerations-for-evaluating-graph-warehouses>



**1 Beacon Street, 15th Floor,  
Boston, MA 02108 USA**

**CambridgeSemantics.com**

**info@cambridgesemantics.com**