

Effective Traffic Management With OrientDB

Traffic management systems aim to improve the flow of traffic and reduce congestion in areas such as freeways, busy city streets and urban areas. Through the use of tools, roadside sensors, traffic lights, and imposed rules such as speed limits, regular traffic patterns that lead to congestion may be analyzed, efficient routes established and problematic areas identified. In order to process this data however, a fast, efficient and intuitive data processing technology is needed, capable of analyzing real-time data and its seemingly unpredictable behaviours to suggest alternate routes, predict movements and map traffic. Streets and roads are in constant flux, cars break down, accidents happen and an endless list of events (data) can affect traffic patterns. All this generates large amount of data which must be processed quickly, and in a secure environment, as it is sensitive by nature and would not be safe in the wrong hands.

But how is all this data handled? What is the most efficient method to store and process all this data?

Relational tables are simple and operational, though they are ineffective when data must be processed quickly due to slow reads and writes. Though document databases do away with tables and with their schema-less approach may allow for the creation of complex data structures, they are ineffective at forming relationships between events, have single points per unit and fixed time windows. Graph databases have become an increasingly popular choice when creating dynamic structures due to their fast traversal of index free nodes and edges, but they are not without disadvantages. Though events can be connected in multiple ways, writes are usually slow and reads not as fast as in document databases.

What if there was a way to take all the advantages of document and graph databases to maximize efficiency?

This paper will focus on OrientDB's multi-model approach which combines graph and document models, creating a unique method. Though polyglot persistence ensures that separate systems communicate with each other, this is not a faultless mechanism. Separate systems require separate skill sets and an added layer of scrutiny which is hard to maintain. Divided mechanisms such as these have an impact on performance, eventually leading to slower results. In an attempt to optimize the speed at which data is processed, these systems often sacrifice relevant data for first or second level links.

This was true until now.

OrientDB harnesses the speed of a document model with the flexibility of graphs and as a result of combining these models, documents may also contain pointers to other elements since OrientDB treats every edge and node as a JSON document.

Industry

Automotive/Location
Services/Security

Challenge

Exploit and process large amounts of traffic data to intuitively analyze and act upon real-time information to find alternate routes, predict movements and safely store information.

Approach

Harness the advantages of graph and document database models. Discover relevant relationships between events, markers and time to provide effective results store real-time information and process queries quickly and efficiently.

Solution

Use OrientDB's combined Graph and Document models as first class citizens to Store transactional data as graph vertices and using a SQL based query language.

Result

Create a single database for several key challenges to traffic management, resulting in increased speed and more relevant predictions.

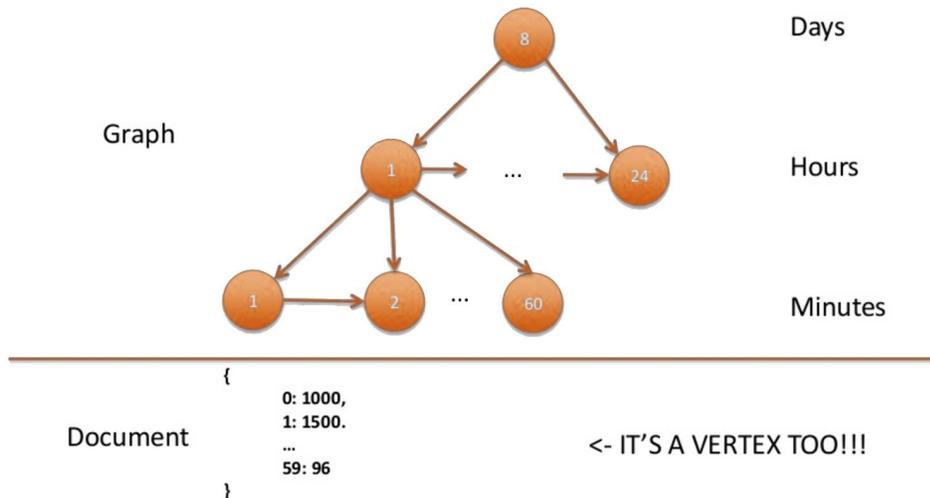


Fig 1.1 OrientDB Treats every Edge and Vertex as a JSON Document

Of course, security is paramount, especially when dealing with sensitive, real time data and as this paper will demonstrate, OrientDB excels at manipulating complex models in secure environments. OrientDB's multi-model approach will also enable systems to discover new time related dimensions as well as tackle complex forecasting.

For the purpose of this paper, let's analyze a complex scenario in which a car's movements must be tracked. In this scenario, we have the following conditions:

- A car is moving along a city grid
- Cameras or people track its movements and provide coordinates as well as other unstructured data
- You would like to predict where this car will go next

Time Series

In order to track a vehicle's movements, we must make use of coordinates and points on a map to form a time series. A time series is a sequence of data points, typically consisting of successive measurements made over a time interval. It tracks event sequences such as a set of events with a timestamp, a set of relationships (happened before/after) and cause and effect relationships. Time as a dimension can be direct (ex. I'm a friend of John's since) or calculated, such as speed (distance over time).

However, in the case of traffic management, time is also a restraint as it takes time to process complex queries. We do not have time since events happen in real time, must be written instantly and subsequently processed. The problem that arises here is being both fast and effective. Data arrives and must be written fast and reads must happen in a short time despite large amounts of data.

About OrientDB

"With the release of OrientDB 2.1 and breakthrough innovation in OrientDB 2.2, we are building the industry's first distributed document-graph database." said Luca Garulli, CEO and Founder of OrientDB. "When I started to work on OrientDB back in 2010, I could barely imagine the pace of growth we're seeing today thanks to our customers, users and advocates." Ranked second most popular graph db by DB Engines, sixth document store and seventh key value store, it's within the top ten in all three categories and steadily climbing the charts. These combined models in one engine makes it a powerful, easy to use tool that doesn't limit its users to adopt a single DB model.

Contact Info



enquiries@orientdb.com

So, in order to trace the movements of a vehicle, predict where it might move to next or suggest alternate routes, effective manipulation of complex operations such as aggregation, prediction and analysis must happen instantly.

Let us begin tracking a car's movements. In order to do this, a map is made of points and distances. Different streets may have different speed limits. Remember that distance over speed equals time.

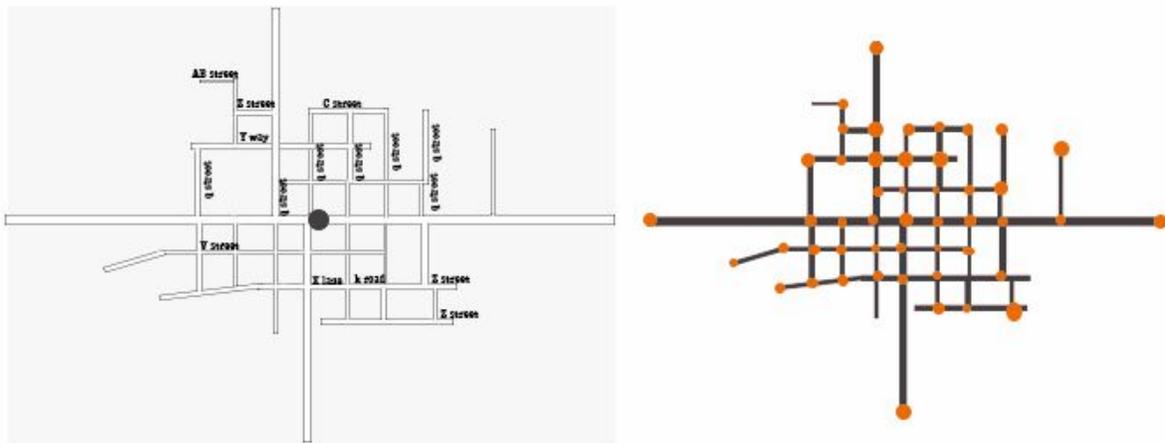


Fig. 2.1 Map and its point

A time series of the car's movements can be created and an increasingly complex event sequence formed.

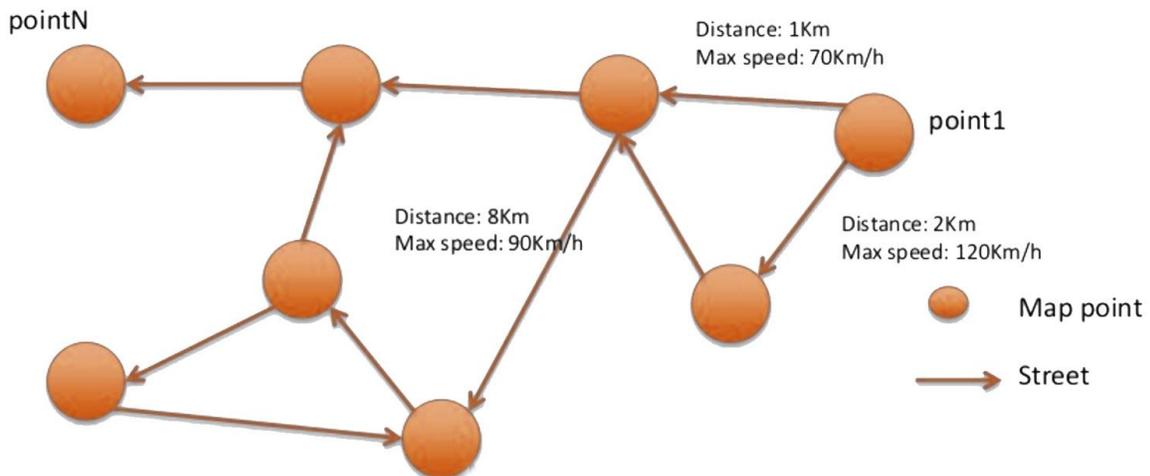


Fig. 2.2 Time series in an OrientDB graph.

Since OrientDB treats all vertices and edges as documents, complex data can be stored inside these and used in order to predict subsequent movements and calculated speeds or distances. This data can then be used to calculate the shortest distance from a given point to the car being tracked.

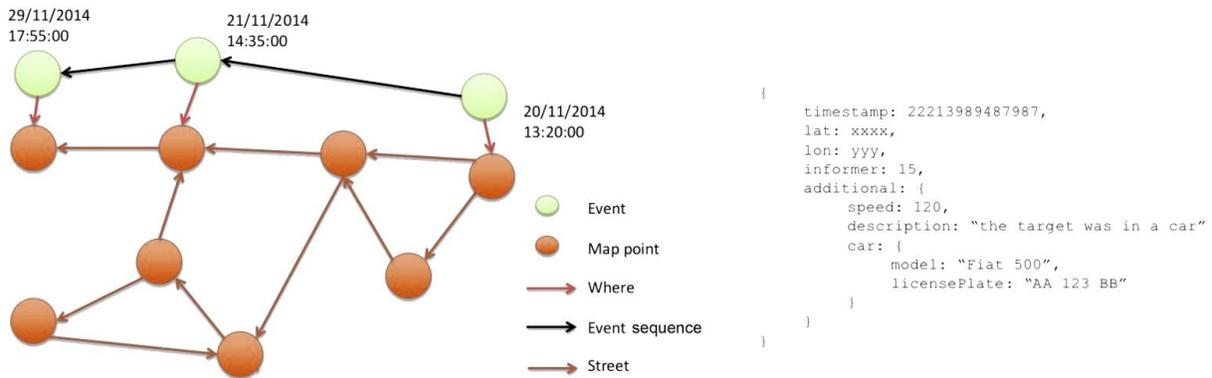


Fig. 2.3 OrientDB graph with Document properties

What about Security?

At OrientDB, we know that data is sensitive. In addition to the built-in authentication of users that happens inside the database engine, the latest OrientDB (2.2) security system supports multiple types of external authenticators. Each authenticator implements various methods to support authentication and authorization amongst other things. The OrientDB authenticator (enterprise edition only) provides Kerberos authentication and full browser SPNEGO support. The array of built-in and easily configurable authentication systems provided in OrientDB are enabled by default for each database, ensuring that your applications are safe.

In addition, OrientDB 2.2 now provides a customizable password validator that, when enabled, validates user chosen passwords based on minimum length and regular expressions. Nevertheless, the built-in password validator may be replaced with a custom validator component.

Importing of LDAP users is also available as an OrientDB enterprise feature. Each database can be configured to import from multiple LDAP domains. Each domain may specify multiple LDAP servers and the users and their corresponding roles can be specified per-domain using a standard LDAP filter.

"[OrientDB] has some obvious advantages in document-oriented environments, which makes it a potential competitor to RDF databases (OrientDB is ACID compliant) as well as graph databases. It supports Gremlin and Blueprints but, more importantly, it uses an extended form of SQL for query processing. The product is schema-free and uses sharding for distributing data across a cluster."

- Bloor Market Update

A Simple Choice

OrientDB is a 2nd Generation Open Source distributed graph-document database where every vertex and edge is a JSON Document. It supports schema-less, schema-full and schema-mixed modes, has a strong security profiling system based on user and roles and supports SQL amongst the query languages. It is incredibly fast, storing 220,000 records per second on common



WHITE PAPER

Time Series

hardware and allows one to traverse parts of or entire trees and graphs of records in a few milliseconds. Its multi-model approach lends itself to easily and efficiently link database types to maximize performance.

Ranked 2nd most popular graph DB by DB engines, 6th document store and 7th key value store, it's within the top 10 in all 3 model categories and steadily climbing the charts. Combining various models into one engine makes it a powerful, easy to use tool that doesn't limit its users to adopt a single DB model.

OrientDB comes in two editions. The Community Edition is open source and free for any use (Apache 2 license). For those companies looking for Professional Support, Consultancy and additional advanced enterprise features, OrientDB Enterprise Edition is available upon request. Fortune 500 companies, government entities and startups all use this technology to build large-scale innovative applications. Clients include the United Nations, Ericsson, Sky, CenturyLink and Sonatype. OrientDB is the winner of the 2015 Infoworld Bossie award and has been covered by multiple media outlets.

Replacing your DBMS, once it no longer meets requirements, can be a huge cost in time and resources. Is your database powerful, scalable and flexible enough to grow with you? OrientDB also reduces the need to support multiple products to achieve your goals. See for yourself, give it a try! If you would like to receive more information about OrientDB and its capabilities, send us a message at enquiries@orientdb.com.