

## Fast SQL Queries on large datasets

**Large Retailer Accelerates Queries and Reporting:** Running SQL queries on very large data sets on Hadoop for fast, ad-hoc explorative analytics and visualization can eliminate the need for a data warehouse.

*This case study details a project to speed up complex SQL queries in data sets ranging from 3-10B rows for a large retailer. Result: Queries taking 2 – 15 min earlier, now return results under 15 seconds*

### The Problem

Our client was facing severe performance issues when trying to run complex aggregations and multi-dimensional queries for analysis using Tableau on Hadoop. Their data set included several billion rows. They tried various products like commercial OLAP cubes as well as Apache HAWQ. Each solution had its drawbacks; we were able to find a better solution to query over 1B rows with most returning results in less than 10 seconds.

### The Solution

The solution was designed and integrated into Orzota BigForce – our big data management platform using open-source components and our accelerators. The analytics engine with full SQL support not only met the required query times, but also allowed frequent re-builds of the OLAP cubes (regular refreshes every 2-3 days was a firm business requirement). This combination allows for direct SQL based analytics on very

large data sets with extremely fast response times, eliminating the need for a separate data warehouse.

This architecture delivered great performance on interactive analytics functions, improved the productivity and efficiency of business analysts and eliminated the need for a Data Warehouse, thus allowing for significant cost savings.

### Benefits

Integrating the SQL analytics engine into the Hadoop environment can have a dramatic effect on your business, including:

- Fast SQL queries for interactive analytics using tools like Tableau
- Simultaneous analysis by multiple analysts solving different problems with great performance improving productivity
- Eliminate the requirement for a data warehouse or data mart.

## Orzota's Data Warehouse on Hadoop Services

If you have an Enterprise Data Warehouse, [Orzota](#) can help improve its efficiency. If you *don't* have an enterprise data warehouse, we can help eliminate its need and make its capabilities available on Hadoop, while bringing in more unstructured data for advanced analytics. For more information about this or other products and services, contact us at [info@orzota.com](mailto:info@orzota.com).