

Beyond TPC-DS, a benchmark for Big Data OLAP systems (BDOLAP-Bench)*

Roberto Tardío¹[0000-0002-7913-1534], Alejandro Maté^{2,3}[0000-0001-7770-3693],
and Juan Trujillo³[0000-0003-0139-6724]

¹ StrateBI Business Solutions Ltd, Madrid, Spain

`roberto.tardio@stratebi.com`

² University of Alicante, Alicante, Spain

`{amate,jtrujillo}@dlsi.ua.es`

Keywords: Big Data OLAP, Benchmarking, Data modeling, Kylin, Druid

Published in: Future Generation Computer Systems, Vol. 132, pp. 136–151, 2022

Impact Factor: JCR 7.307 - Q1 - Position: 10/110 - Area: COMPUTER SCIENCE, THEORY & METHODS

DOI: <https://doi.org/10.1016/j.future.2022.02.015>

Abstract. Online Analytical Processing (OLAP) systems with Big Data support allow storing tables of up to tens of billions of rows or terabytes of data. At the same time, these tools allow the execution of analytical queries with interactive response times, thus making them suitable for the implementation of Business Intelligence applications. However, since there can be significant differences in query and data loading performance between current Big Data OLAP tools, it is worthwhile to evaluate and compare them using a benchmark. But we identified that none of the existing approaches are really suitable for this type of system. To address this, in this research we propose a new benchmark specifically designed for Big Data OLAP systems and based on the widely adopted TPC-DS benchmark. To overcome TPC-DS inadequacy, we propose (i) a set of transformations to support the implementation of its sales data mart on any current Big Data OLAP system, (ii) a choice of 16 genuine OLAP queries, and (iii) an improved data maintenance performance metric. Moreover, we validated our benchmark through its implementation on four representative systems.

* Este trabajo ha sido co-financiado por el proyecto AETHER-UA (PID2020-112540RB-C43) del Ministerio de Ciencia e Innovación y el proyecto BALLADEER (PROMETEO/2021/088), financiado por la Conselleria d'Innovació, Universitats, Ciència i Societat Digital.