

OLAP Performance Analysis in Neo4j/Cypher Deployed on OpenStack: Challenges and Results for Medium-Sized TPC-H Data Sets

Thursday 18 September 2025 15:15 (15 minutes)

Among NoSQL databases, Neo4j has become a leading solution for managing graph-structured data, particularly known for its flexibility in handling complex relationships. However, the performance characteristics of Neo4j in analytical (OLAP) workloads deployed within cloud-based environments, such as OpenStack, remain underexplored. In this paper, we present an analysis of OLAP performance for Neo4j clusters deployed on a private OpenStack cloud. Using the TPC-H benchmark dataset, transformed specifically into a graph schema, we generated and executed a comprehensive suite of Cypher queries designed to test Neo4j's capabilities under varying workloads and configurations. Our experiments reveal that increasing data volume reduces performance and success rates of query executions. Additionally, expanding the cluster size through additional nodes provided limited performance gains, highlighting diminishing returns from horizontal scaling for medium-sized analytical workloads.

Author: ESANU, Codrin-Stefan (Alexandru Ioan Cuza University of Iasi, Cegeka Romania)

Co-authors: Prof. FOTACHE, Marin (Alexandru Ioan Cuza University of Iasi); CLUCI, Marius-Iulian (Alexandru Ioan Cuza University of Iasi; S.C. HVM SRL Iași); Dr PINZARU, Ciprian (UAIC); RUSU, Octavian (Alexandru Ioan Cuza University, Iasi, Romania)

Presenter: ESANU, Codrin-Stefan (Alexandru Ioan Cuza University of Iasi, Cegeka Romania)

Session Classification: Cloud Computing and Network Virtualisation

Track Classification: Grid, Cloud & High Performance Computing in Science