

Online Reconfiguration of Clustered NoSQL Databases for Time-Varying Workloads

Sophia automatically reconfigures NoSQL databases for optimal performance by predicting future workloads and performing a cost-benefit analysis before initiating any changes.

Innovators at Purdue University developed a system known as Sophia to reconfigure a NoSQL Database Management System for highest throughput as the application characteristics change.

Sophia increases the economic lifecycle of databases through a Cost Benefit Analysis (CBA) coupled with a Reconfiguration Plan, implements intelligent incremental reconfigurations, and predicts future workloads up to an hour ahead using a Workload Predictor, thereby optimizing throughput. In addition, the Sophia program will initiate reconfiguration only when it estimates the benefit over a future time window is higher than the cost of the reconfiguration. The invention is efficient for cloud computing as well as creating and managing on-premises databases.

Benefit

Advantages:

- Adaptable for changing workloads, including no change
- Compatible with a variety of NoSQL databases
- Keeps data available to users during reconfiguration
- Performs Cost Benefit Analysis to determine optimal reconfiguration

Potential Applications:

- Cloud computing
- Distributed NoSQL databases
- Data analytics for scientific research

TRL: 5

Technology ID

2019-BAGC-68646

Category

Artificial Intelligence & Machine
Learning/Reinforcement &
Federated Learning

Authors

Saurabh Bagchi
Somali Chaterji
Ashraf Mahgoub
Paul Curtis Wood

Further information

Matt Halladay
MRHalladay@prf.org

Erinn Frank
EEFrank@prf.org

View online



Intellectual Property:

Provisional-Patent, 2019-07-05, United States

Utility-Gov. Funding, 2020-04-16, United States

CON-Gov. Funding, 2022-09-06, United States

Keywords: NoSQL database reconfiguration, online tuning, NoSQL throughput optimization, dynamic workload tuning, Cost Benefit Analysis database, clustered NoSQL databases, NoSQL performance optimization, Sophia Purdue University, database economic lifecycle, intelligent incremental reconfigurations, Cache, Cloud Computing, Cost Efficient, Data, Data Storage, Database & Information Management, Efficiency, Information Interfaces and Presentation, Information Management System, Information Technology, Network Analyzer, Networks, Optimization, RAM, Tracking, User Interfaces