IJARSCT

ISSN (Online) 2581-9429

Impact Factor: 7.301

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 8, April 2023

Optimizing the Performance of Data Warehouse by Query Cache Mechanism in Big Data

Anand Tilagul[1], Praveen S[2], Rachan S H[3], Shreyas M[4]

Assistant Professor, Department of Information Science and Engineering[1]

Students, Department of Information Science and Engineering[2,3,4]

SJC Institute of Technology Chikkaballapura, India

Abstract*:* *In today's world of Business Intelligence (BI), fast and efficient access to data from Data Warehouses (DW) is crucial. With the increasing amount of Big Data, caching has become one of the most effective techniques for improving data access performance. DWs are widely used by organizations for managing and using data in Decision Support Systems (DSS). To optimize the performance of fetching data from DWs, various methods have been employed, and one of them is the Query Cache method. Our proposed work focuses on a cache-based mechanism that improves the performance of DWs in two ways. First, it reduces the execution time by directly accessing records from cache memory, and second, it saves cache memory space by eliminating non-frequently used data. Our goal is to fill the cache memory with the most frequently used data. To achieve this objective, we utilize an aging-based Least Frequently Used (LFU) algorithm that considers the size and frequency of data simultaneously. This algorithm manages the priority and expiry age of the data in cache memory by taking into account both the size and frequency of data. LFU assigns priorities and counts the age of data placed in the cache memory. The cache block entry with the lowest age count and priority is eliminated first. Ultimately, our proposed cache mechanism efficiently utilizes cache memory and significantly improves the performance of data access between the main DW and the business user query.*

Keywords: Business Intelligence

REFERENCES

[1]. W. Moudani, M. Hussein, M. Moukhtar, and F. Mora- Camino, ‘‘An intelligent approach to improve the performance of a data warehouse cache based on association rules,’’ J. Inf. Optim. Sci., vol. 33, no. 6, pp. 601–621, Nov. 2012.

[2]. A. Simitsis, P. Vassiliadis, and T. Sellis, ‘‘Optimizing ETL processes in data warehouses,’’ in Proc. 21st Int. Conf. Data Eng. (ICDE), Tokyo, Japan, 2005, pp. 564–575.

[3]. D. Matani, K. Shah, and A. Mitra, ‘‘An O (1) algorithm for implementing the LFU cache eviction scheme,’’ Oct. 2021, arXiv:2110.11602.

[4]. S. Huang, Q. Wei, D. Feng, J. Chen, and C. Chen, ‘‘Improving flash-based disk cache with lazy adaptive replacement,’’ ACM Trans. Storage, vol. 12, no. 2, pp. 1–24, Mar. 2016.

[5]. M. S. A. Khaleel, S. E. F. Osman, and H. A. N. Sirour, ‘‘Proposed ALFUR using intelegent agent comparing with LFU, LRU, SIZE and PCCIA cache replacement techniques,’’ in Proc. Int. Conf. Commun., Control, Comput. Electron. Eng. (ICCCCEE), Khartoum, Sudan, Jan. 2017, pp. 1–6.

[6]. Z. Wang, K. Zeng, B. Huang, W. Chen, X. Cui, B. Wang,

[7]. J. Liu, L. Fan, D. Qu, Z. Hou, T. Guan, C. Li, and J. Zhou, ‘‘Grosbeak: A data warehouse supporting resource-aware incremental computing,’’ in Proc. ACM SIGMOD Int. Conf. Manage. Data, Jun. 2020, pp. 2797–2800.

DOI: 10.48175/IJARSCT-9627