

Towards Adaptive SQL Query Optimization in Distributed Stream Processing

ABSTRACT

Distributed stream processing is widely adopted for real-time data analysis and management. SQL is becoming a common language for robust streaming analysis due to the introduction of time-varying relations and event time semantics. However, query optimization in state-of-the-art stream processing engines (SPEs) remains limited: runtime adjustments to execution plans are not applied. This fact restricts the optimization capabilities because SPEs lack the statistical data properties before query execution begins. Moreover, streaming queries are often long-lived, and these properties can be changed over time.

Adaptive optimization, used in databases for queries with insufficient existing data statistics, can fit the streaming scenario. In this work, we explore the main challenges that SPEs face during the adjustment of adaptive optimization: retrieving and predicting statistical data properties, execution graph migration, misfit of SPEs programming interfaces, etc. We demonstrate the feasibility of the proposed approach within an extension of the Nexmark streaming benchmark and outline our further work on this topic.

1 INTRODUCTION

2 A RUNNING EXAMPLE

2.1 Query

2.2 Implementation

2.3 Idea

3 CHALLENGES

3.1 Fetching and predicting data statistics

3.2 Statistics integration in query processing

3.3 In-flight execution graph migration

4 PRELIMINARY EXPERIMENTS

5 DISCUSSION

6 RELATED WORK

7 CONCLUSION

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference'17, July 2017, Washington, DC, USA

© 2021 ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnn>