

Accepted Manuscript

Improving the robustness and performance of parallel joins over distributed systems

Long Cheng, Spyros Kotoulas, Tomas E. Ward, Georgios Theodoropoulos



PII: S0743-7315(17)30205-8
DOI: <http://dx.doi.org/10.1016/j.jpdc.2017.06.016>
Reference: YJPDC 3706

To appear in: *J. Parallel Distrib. Comput.*

Received date: 2 February 2017

Revised date: 9 June 2017

Accepted date: 18 June 2017

Please cite this article as: L. Cheng, S. Kotoulas, T.E. Ward, G. Theodoropoulos, Improving the robustness and performance of parallel joins over distributed systems, *J. Parallel Distrib. Comput.* (2017), <http://dx.doi.org/10.1016/j.jpdc.2017.06.016>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Improving the Robustness and Performance of Parallel Joins over Distributed Systems

- We introduce the state-of-the-art parallel joins and analyze their possible performance issues in the presence of data skews.
- We introduce the query-based join, a novel parallel join approach for handling data skew in distributed architectures.
- We propose an efficient and robust join algorithm referred to as PRPQ (partial redistribution & partial query) which is capable of higher performance than current techniques.
- Our experimental results demonstrate that PRPQ always outperform the state-of-art approach.

Download English Version:

<https://daneshyari.com/en/article/4951618>

Download Persian Version:

<https://daneshyari.com/article/4951618>

[Daneshyari.com](https://daneshyari.com)