

Accepted Manuscript

T3-Scheduler: A topology and Traffic aware Two-level Scheduler for stream processing systems in a heterogeneous cluster

Leila Eskandari, Jason Mair, Zhiyi Huang, David Eyers

PII: S0167-739X(17)32643-2
DOI: <https://doi.org/10.1016/j.future.2018.07.011>
Reference: FUTURE 4329

To appear in: *Future Generation Computer Systems*

Received date : 17 November 2017

Revised date : 5 July 2018

Accepted date : 10 July 2018

Please cite this article as: L. Eskandari, J. Mair, Z. Huang, D. Eyers, T3-Scheduler: A topology and Traffic aware Two-level Scheduler for stream processing systems in a heterogeneous cluster, *Future Generation Computer Systems* (2018), <https://doi.org/10.1016/j.future.2018.07.011>

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.



T3-Scheduler: A Topology and Traffic Aware Two-Level Scheduler for Stream Processing Systems in a Heterogeneous Cluster

Leila Eskandari*, Jason Mair, Zhiyi Huang, David Eyers

*Department of Computer Science
University of Otago
Dunedin
New Zealand*

Email: {leila,jkmair,hzy,dme}@cs.otago.ac.nz

Abstract

To efficiently handle a large volume of data, scheduling algorithms in stream processing systems need to minimise the data movement between communicating tasks to improve system throughput. However, finding an optimal scheduling algorithm for these systems is NP-hard. In this paper, we propose a heuristic scheduling algorithm—T3-Scheduler—for a heterogeneous fog or cloud cluster that can efficiently identify the tasks that communicate with each other and assign them to the same node, up to a specified level of utilisation for that node. Using three common micro-benchmarks and an evaluation using two real-world applications, we demonstrate that T3-Scheduler outperforms current state-of-the-art scheduling algorithms, such as Aniello et al.’s popular ‘Online scheduler’ and R-Storm, improving throughput by up to 32% for the two real-world applications.¹

Keywords:

Stream processing, Scheduling, Big Data, Heterogeneous Cluster

*Primary contact: Phone: +64 3 479 8498 Fax: +64 3 479 8529

¹This work is an extension of an Auto-DaSP workshop paper in [1]

Download English Version:

<https://daneshyari.com/en/article/6872852>

Download Persian Version:

<https://daneshyari.com/article/6872852>

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