

Reviewer Report

Title: Bioinformatics applications on Apache Spark

Version: Revision 1 **Date:** 07/23/2018

Reviewer name: **Tae-Hyuk Ahn**

Reviewer Comments to Author:

Minor 1: To separate different applications and approaches, authors used paragraph break (single line space). However, several paragraphs are excessively short. For example, page 2 line 53 and page 7 line 49. Please update the manuscript to have better structure of paragraphs.

Minor 2: Table 1 is mentioned at the last paragraph of Discussion section. Please cite the table in the beginning of manuscript (just after the Introduction section) for users to easily catch the features of many Spark-based bioinformatics applications.

Minor 3: Few suggestions for correction:

Page 2: Line 49: As in Figure 1 -> As shown in Figure 1

Page 2: Line 49: Spark application runs as independent processes on the cluster and are coordinated by the SparkContext in the driver program -> Each Spark application runs as independent process on the cluster coordinated by the SparkContext in the driver program

Page 3: Line 51: because of a node failure -> because of the node failure

Level of Interest

Please indicate how interesting you found the manuscript: Choose an item.

Quality of Written English

Please indicate the quality of language in the manuscript: Choose an item.

Declaration of Competing Interests

Please complete a declaration of competing interests, considering the following questions:

- Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
- Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
- Do you hold or are you currently applying for any patents relating to the content of the manuscript?
- Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?
- Do you have any other financial competing interests?
- Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (<http://creativecommons.org/licenses/by/4.0/>). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal

To further support our reviewers, we have joined with Publons, where you can gain additional credit to further highlight your hard work (see: <https://publons.com/journal/530/gigascience>). On publication of this paper, your review will be automatically added to Publons, you can then choose whether or not to claim your Publons credit. I understand this statement. Yes