Errata for "Cerebro: A Data System for Optimized Deep Learning Model Selection"

Supun Nakandala, Yuhao Zhang, and Arun Kumar
University of California, San Diego
{snakanda, yuz870, arunkk}@eng.ucsd.edu

We discovered that there was an inconsistency in the communication cost formulation for the decentralized fine-grained training method in **Table 2** of our paper [1]. We used Horovod as the archetype for decentralized fine-grained approaches, and its correct communication cost is higher than what we had reported. So, we amend the communication cost of decentralized fine-grained to $\mathbf{2km(p-1)|S|\left\lceil \frac{|D|}{bp} \right\rceil}$, instead of $kmp|S|\left\lceil \frac{|D|}{bp} \right\rceil$.

With this correction, **Table 2** of our paper should be corrected as follows, which uses the same notation.

Table 1: Communication cost analysis of MOP and other approaches. *Full replication. †Remote reads. ‡Parameters for the example: $k = 20$, $|S| = 20$, $p = 10$, $m = 1$GB, $\langle D\rangle = 1$TB, and $|D|/b = 100$K.

	Comm. Cost	Example‡
Model Hopper Parallelism	$kmp\|S\| + m\|S\|$	4 TB
Task Parallelism (FR*)	$p\langle D\rangle + m\|S\|$	10 TB
Task Parallelism (RR†)	$k\|S\|\langle D\rangle + m\|S\|$	400 TB
Bulk Synchronous Parallelism	$2kmp\|S\|$	8 TB
Centralized Fine-grained	$2kmp\|S\|\left\lceil \frac{\|D\|}{bp} \right\rceil$	80 PB
Decentralized Fine-grained	$\mathbf{2km(p-1)\|S\|\left\lceil \frac{\|D\|}{bp} \right\rceil}$	**72 PB**

Also, the last two paragraphs of Section 2 that refers to the above table should be corrected as follows:

> All PS-style approaches have *high communication* due to their centralized all-to-one communications, which is proportional to the number of mini-batches and orders of magnitude higher than BSP, e.g., **10,000x** in Table 2.
>
> Decentralized Fine-grained. The best example is Horovod. It adopts HPC-style techniques to enable synchronous all-reduce SGD. While this approach is bandwidth optimal, communication latency is still proportional to the number of workers, and the synchronization barrier can become a bottleneck. The total communication overhead is also proportional to the number of mini-batches and orders of magnitude higher than BSP, e.g., **9,000x** in Table 2.

The above amendments are purely in the conceptual exposition and do not affect any technical findings, empirical results, or conclusions in the paper.

1. REFERENCES

[1] S. Nakandala, Y. Zhang, and A. Kumar. Cerebro: A data system for optimized deep learning model selection. *Proc. VLDB Endow.*, 13(12):21592173, July 2020.

Proceedings of the VLDB Endowment, Vol. 13, No. 11
ISSN 2150-8097.
DOI: https://doi.org/10.14778/3407790.3407816