

Corrections to “Toward Load-Balanced Redundancy Transitioning for Erasure-Coded Storage”

Keyun Cheng, Huancheng Puyang, Xiaolu Li, Patrick P. C. Lee, Yuchong Hu, Jie Li, and Ting-Yi Wu

Last modified: April 24, 2025

Issue: Incorrect reduction of the stripe group construction step to an NP-hard problem.

In the last paragraph of Section 3, we incorrectly reduce the stripe group construction step into an NP-hard problem, where we transform the stripe group construction step into a k -means clustering problem. The correct reduction is to transform a known NP-hard problem (i.e., k -means clustering) to our problem (i.e., stripe group construction). We modify the paragraph as follows:

Note that the stripe group construction step can be viewed as a variant of k -means cluster problem [30], which is known to be NP-hard. The k -means clustering problem partitions a set of data points in a space into multiple disjoint clusters, where each cluster is represented by the mean of the data points in the cluster, and the goal is to minimize a cost function, defined as the sum of the squared distances between each data point and the mean of its assigned cluster. To formulate stripe group construction based on k -means clustering, we represent each of the M input stripes as a data point in some N -dimensional space corresponding to the N nodes, and represent a stripe group as a cluster comprising λ data points. Intuitively, for load-balanced redundancy transitioning, our goal is to partition the data points into $\frac{M}{\lambda}$ disjoint clusters, where our goal is to minimize the sum of a cost function, defined based on the MTL. The subsequent parity block generation and stripe re-distribution steps can also affect the MTL of the resulting redundancy transitioning solution, thereby making the redundancy transitioning problem intractable.