Master Thesis

Sébastien Vaucher
sebastien.vaucher@unine.ch

13 January 2016

Contents

1 State of the art

1.1 Articles

1.1.1 XORing Elephants: Novel Erasure Codes for Big Data [1]

The article presents a new family of erasure codes called Locally Repairable Codes (LRCs). These codes enable local repair of faulty data. It means that to repair a file, it is not needed to load all of its blocks, as is common with e.g. Reed-Solomon. The authors implemented their algorithm in Hadoop HDFS and deployed a test to Facebook clusters. The repair process of LRC uses half the disk and network bandwidth compared to Reed-Solomon, at the expense of $14\,\%$ more storage usage.

Bibliography

[1] M. Sathiamoorthy *et al.*, "Xoring elephants: novel erasure codes for big data", in *Proceedings of the VLDB Endowment*, VLDB Endowment, vol. 6, 12 Jan. 2016, pp. 325–336.