Accepted Manuscript

Resemblance and Mergence based Indexing for High Performance Data Deduplication

Panfeng Zhang, Ping Huang, Xubin He, Hua Wang, Ke Zhou

PII: S0164-1212(17)30038-9
DOI: 10.1016/j.jss.2017.02.039
Reference: JSS 9926

To appear in: *The Journal of Systems & Software*

Received date: 29 July 2016
Revised date: 20 December 2016
Accepted date: 23 February 2017

Please cite this article as: Panfeng Zhang, Ping Huang, Xubin He, Hua Wang, Ke Zhou, Resemblance and Mergence based Indexing for High Performance Data Deduplication, *The Journal of Systems & Software* (2017), doi: 10.1016/j.jss.2017.02.039

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.

ACCEPTED MANUSCRIPT

Highlights

- A new de-duplication scheme based on two-level index structure and dynamic bloom filter.
- A fast resemblance approach to index duplicate data based on “segments”.
- A novel resemblance mergence strategy that groups segments into bins.
- A new frequency based cleanup method to avoid storing low-frequent fingerprints.
- A thorough evaluation of our approach to demonstrate its effectiveness.

ACCEPTED MANUSCRIPT

Download English Version:

https://daneshyari.com/en/article/4956476

Download Persian Version:

https://daneshyari.com/article/4956476

Daneshyari.com