CMSC724: Literature Survey
Applying Compression Techniques in MapReduce

Greg Benjamin, Samet Ayhan, Kishan Sudusinghe

University of Maryland, College Park

1 Inspiration: Floratou et al.

2 Background

2.1 MapReduce

2.2 Column-stores

2.3 Parallel Databases (?)

3 MapReduce Performance

3.1 Serialization and Data Format

3.2 Data Layout

3.3 Indexing

3.4 Data Compression

4 Other Concerns

4.1 Energy Efficiency

4.2 anything else i'm forgetting (?)

References

[1] Avrilia Floratou, Jignesh M. Patel, Eugene J. Shekita, and Sandeep Tata. 2011. Column-oriented storage techniques for MapReduce. Proc. VLDB Endow. 4, 7 (April 2011), 419-429.