**Apache Hadoop Performance-Tuning Methodologies and Best Practices**

**Summary**

This paper introduces hardware and software tuning techniques including BIOS, OS, JVM and Hadoop configuration parameter tuning. For Hadoop configuration turning, the author shares some parameters that have been observed to improve the performance. The author finds that using LZO codec has shown better performance (same opinion in 2 papers). For JVM tuning, author tries command-line flags like AggressiveOpts, UseCompressedStrings, and UseStringCache, finding around 2% degradation in performance and 3% improvement by tuning GC flags on Cluster A. For OS configuration tuning, author finds using EXT4 filesystem is a better option than EXT3 filesystem. Also, using noatime FS attribute, CFQ scheduler also improves the performance. For BIOS configuration tuning, the author introduces Native command queuing (NCQ) feature that helps to improve I/O performance.

Parameter List:

A screenshot of a cell phone

Description automatically generated