

# Hadoop Distributed File System

Consider the following file. The file is stored in block zip format. Each block starts with the unique signature BZ2. The file is split into three parts so that it can be parsed in parallel. On this file, mark the first and last lines processed by processing each of the three splits.

<div>BZ2..... ..... ..... ..... .....BZ2..... ..... ..... .....BZ2..... .....</div>
<div>BZ2..... ..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... .....</div>
<div>..... ..... .....BZ2..... ..... .....BZ2..... ..... .....BZ2..... .....</div>