

ABSTRACT

The present application provides a data processing method and apparatus based on distributed storage, a device, and a medium. The method comprises: after first block data is written into a local cache and local storage, detecting a first block processing space of a storage buffer pool to obtain a
5 first space detection result; if the first space detection result indicates that the number of pieces of block data to be stored does not reach a buffer block processing threshold, writing the first block data into the storage buffer pool on the basis of the N pieces of block data to be stored; detecting a second block processing space of a storage processing pool to obtain a second space detection result; and if
10 the second space detection result indicates that the number of pieces of block data to be fed back does not reach a feedback block processing threshold, determining, in the storage buffer pool and on the basis of the number of pieces of block data to be fed back and the feedback block processing threshold, block data to be merged which is used for being subjected to overlapping detection with the block data to be fed back, and writing the data reserved after the overlapping detection into a distributed database. By means of the present application, the storage efficiency of data can be improved.