

DISTRIBUTED STORAGE AND REPLICATION SYSTEM AND METHOD

ABSTRACT

The present invention discloses a distributed storage and replication system and a
5 method. The system includes at least one metadata control (MDC) module, multiple
IO routing modules, and multiple object-based storage device (OSD) nodes, where the
MDC module is adapted to configure, for each OSD node, at least one logical
partition (partition) corresponding to physical storage resources managed by each
OSD node; the IO routing module is adapted to route a received IO request to an OSD
10 node; and the OSD node is adapted to execute, according to the IO request, storage of
data corresponding to the IO request, where the MDC determines that an OSD node in
the system is a faulty OSD node, determines a partition on the faulty OSD node,
updates a partition view of a partition group that includes the partition on the faulty
OSD node, and sends an updating notification to a primary OSD node in the updated
15 partition view; and the primary OSD node is adapted to process, according to the
updated partition view after receiving the updating notification sent by the MDC
module, replication of the data corresponding to the IO request. According to
embodiments of the present invention, processing performance, fault tolerance, and
availability of consistency replication are improved.

FIGURE 2C