

ABSTRACT

Embodiments of the present application provide an I/O unloading method and system in a cloud environment, a device, and a storage medium. In the present application, a CPU of a host machine is additionally equipped with a queue
5 component, and it is innovatively proposed to use the queue component in an I/O processing process. Based on this, a virtual I/O device and a device queue corresponding to the virtual I/O device can be provided for an instance on the host machine through the queue component; in the CPU, scheduling is performed on an I/O request occurring between the instance and the virtual I/O device by using the
10 device queue; and an original I/O unloading card only needs to be responsible for listening to the I/O request in the device queue and transferring data related to the I/O request. In this way, the device queue in the I/O processing process can be moved up to the CPU, so that it can be linked with each component in the CPU in the process of the I/O scheduling to obtain a real-time load of each core in the CPU, which can
15 realize I/O scheduling more reasonably without blind scheduling, thereby effectively improving the I/O processing efficiency.

Fig. 1