

SERVICE STEERING METHOD, APPARATUS, AND SYSTEM

ABSTRACT

A service steering method, an apparatus, and a system are provided. The system includes a cloud management platform node, a plurality of resource pools, and a steering node. Each resource pool includes a plurality of hosts configured to run services for a plurality of tenants, and at least one instance is run on each host. An instance cluster corresponding to a target service of a target tenant is created in each resource pool, and an instance in the instance cluster is used to run the target service. In the method, the cloud management platform node obtains available resource information of each resource pool associated with the target service, where the available resource information of each resource pool indicates a quantity of instances that are of an instance cluster and that can be provided for the target service from the resource pool in a time period. The cloud management platform node generates, based on the available resource information of each resource pool, a steering plan of the target service in the time period. The steering node allocates service traffic of the target service in the time period to the plurality of resource pools according to the steering plan.

Figure 4