CRUSH Bucket Types

The CRUSH hierarchy organizes OSDs into a tree of different containers, called buckets. For a large installation, you can create a specific hierarchy to describe your storage infrastructure: data centers with rows of racks, racks, hosts, and OSD devices on those hosts. By creating a CRUSH map rule, you can cause Ceph to place an object's replicas on OSDs on separate servers, on servers in different racks, or even on servers in different data centers.

To summarize, buckets are the containers or branches in the CRUSH hierarchy. Devices are OSDs, and are leaves in the CRUSH hierarchy.

Some of the most important bucket attributes are:

- The ID of the bucket. These IDs are negative numbers to distinguish them from storage device IDs.
- The name of the bucket.
- The type of the bucket. The default map defines several types that you can retrieve with the `ceph osd crush dump` command.

  Bucket types include `root`, `region`, `datacenter`, `room`, `pod`, `pdu`, `row`, `rack`, `chassis`, and `host`, but you can also add your own types. The bucket at the root of the hierarchy is of the `root` type.
- The algorithm that Ceph uses to select items inside the bucket when mapping PG replicas to OSDs. Several algorithms are available: `uniform`, `list`, `tree`, and `straw2`. Each algorithm represents a trade-off between performance and reorganization efficiency. The default algorithm is `straw2`.

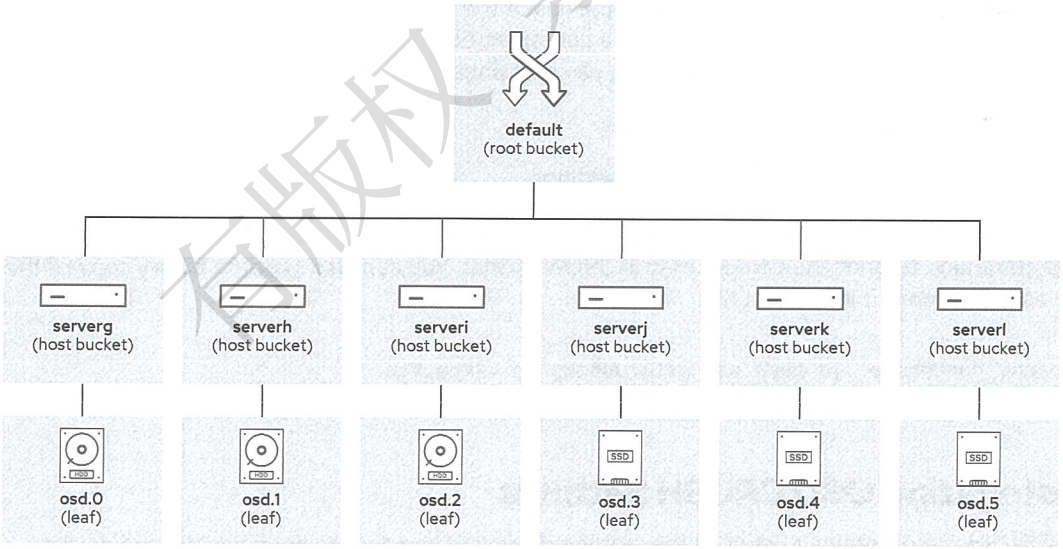

Figure 5.1: CRUSH map default hierarchy example

Customizing Failure and Performance Domains

The CRUSH map is the central configuration mechanism for the CRUSH algorithm. You can edit this map to influence data placement and customize the CRUSH algorithm.

版权禁止外传