OSDMAPTOOL – CEPH OSD CLUSTER MAP MANIPULATION TOOL

SYNOPSIS

osdmaptool *mapfilename* [–print] [–createsimple *numosd* [–pgbits *bitsperosd*]] [–clobber]

DESCRIPTION

osdmaptool is a utility that lets you create, view, and manipulate OSD cluster maps from the Ceph distributed storage system. Notably, it lets you extract the embedded CRUSH map or import a new CRUSH map.

OPTIONS

--print
   will simply make the tool print a plaintext dump of the map, after any modifications are made.

--clobber
   will allow osdmaptool to overwrite mapfilename if changes are made.

--import-crush `mapfile`
   will load the CRUSH map from mapfile and embed it in the OSD map.

--export-crush `mapfile`
   will extract the CRUSH map from the OSD map and write it to mapfile.

--createsimple `numosd [--pgbits bitsperosd]`
   will create a relatively generic OSD map with the numosd devices. If –pgbits is specified, the initial placement group counts will be set with bitsperosd bits per OSD. That is, the pg_num map attribute will be set to numosd shifted by bitsperosd.

--test-map-pgs `[--pool poolid]`
   will print out the mappings from placement groups to OSDs.

--test-map-pgs-dump `[--pool poolid]`
   will print out the summary of all placement groups and the mappings from them to the mapped OSDs.

EXAMPLE

To create a simple map with 16 devices:

```
osdmaptool --createsimple 16 osdmap --clobber
```

To view the result:

```
osdmaptool --print osdmap
```

To view the mappings of placement groups for pool 0:

```
osdmaptool --test-map-pgs-dump rbd --pool 0

pool 0 pg_num 8
0.0     [0,2,1] 0
0.1     [2,0,1] 2
0.2     [0,1,2] 0
0.3     [2,0,1] 2
0.4     [0,2,1] 0
0.5     [0,2,1] 0
0.6     [0,1,2] 0
```

```
0.7      [1,0,2] 1
#osd    count   first   primary c wt     wt
osd.0   8       5       5        1       1
osd.1   8       1       1        1       1
osd.2   8       2       2        1       1
 in 3
 avg 8 stddev 0 (0x) (expected 2.3094 0.288675x))
 min osd.0 8
 max osd.0 8
size 0  0
size 1  0
size 2  0
size 3  8
```

In which,

1. pool 0 has 8 placement groups. And two tables follow:
2. A table for placement groups. Each row presents a placement group. With columns of:
   - placement group id,
   - acting set, and
   - primary OSD.

3. A table for all OSDs. Each row presents an OSD. With columns of:
   - count of placement groups being mapped to this OSD,
   - count of placement groups where this OSD is the first one in their acting sets,
   - count of placement groups where this OSD is the primary of them,
   - the CRUSH weight of this OSD, and
   - the weight of this OSD.

4. Looking at the number of placement groups held by 3 OSDs. We have
   - avarge, stddev, stddev/average, expected stddev, expected stddev / average
   - min and max

5. The number of placement groups mapping to n OSDs. In this case, all 8 placement groups are mapping to 3 different OSDs.

In a less-balanced cluster, we could have following output for the statistics of placement group distribution, whose standard deviation is 1.41421:

```
#osd    count   first   primary c wt     wt
osd.0   8       5       5        1       1
osd.1   8       1       1        1       1
osd.2   8       2       2        1       1

#osd    count   first   primary c wt      wt
osd.0   33      9       9        0.0145874     1
osd.1   34      14      14       0.0145874     1
osd.2   31      7       7        0.0145874     1
osd.3   31      13      13       0.0145874     1
osd.4   30      14      14       0.0145874     1
osd.5   33      7       7        0.0145874     1
 in 6
 avg 32 stddev 1.41421 (0.0441942x) (expected 5.16398 0.161374x))
 min osd.4 30
 max osd.1 34
size 00
size 10
size 20
size 364
```

AVAILABILITY

osdmaptool is part of Ceph, a massively scalable, open-source, distributed storage system. Please refer to the Ceph documentation at http://ceph.com/docs for more information.

SEE ALSO

ceph(8), crushtool(8),