I setup a cluster and want to specify the primary osds through crush rule. Here is the test script

```bash
#!/bin/sh
OSDMAPTOOL=./bin/osdmaptool
CRUSHTOOL=./bin/crushtool
$OSDMAPTOOL --createsimple 240 osdmap --clobber --with-default-pool --mark-up-in --osd_pool_default_size=4
$CRUSHTOOL --build --num_osds 240 host straw2 20 rack straw2 3 datacenter straw2 2 root straw2 2 -o crush
$CRUSHTOOL --d crush --o crush.txt

# set crush rule to choose DC-A first, so primary osd are in DC-A
sed -i 's/step take root0/step take datacenter0/g' crush.txt
sed -i 's/step chooseleaf firstn 0 type host/step chooseleaf firstn 2 type host/g' crush.txt
sed -i 's/step emit/step emit\nstep take datacenter1\nstep chooseleaf firstn 2 type host\nemit/g' crush.txt

$CRUSHTOOL --c crush.txt --o crush
$OSDMAPTOOL osdmap --import-crush crush crush

# use upmap
$OSDMAPTOOL osdmap --upmap-deviation 0.01 --upmap-max 10000 --upmap-pool rbd --upmap-result.sh --debug_osd=20 --debug_crush=20 --upmap-save /dev/null > /tmp/upmap.txt 2>&1

$OSDMAPTOOL osdmap --test-map-pgs-dump-all --pool 1 > dc-a.log

exit 0
```

The result should be that no primary pg on osd.120~osd.239, but we got this in dc-a.log

```
osd.160 257 0 0 1 1
osd.161 256 1 1 1 1
osd.162 256 1 1 1 1
osd.163 256 1 1 1 1
```
In the dc-a.log there are many pgs choose 3 hosts in one DC and 1 host in another, like:

```
1.1585 raw ([57,115,202,145], p57) up ([161,115,202,145], p161) acting ([161,115,202,145], p161)
```

that is not the expected result, we want every pg got 2 hosts in each DC after upmap.

### Related issues:
- Copied to RADOS - Backport #38858: mimic: upmap broken the crush rule  
  - Resolved
- Copied to RADOS - Backport #38859: luminous: upmap broken the crush rule  
  - Resolved
- Copied to RADOS - Backport #38860: nautilus: upmap broken the crush rule  
  - Resolved

### History

**#1 - 03/20/2019 08:27 AM - huang jun**

Here is the crush rule

```plaintext
# rules
rule replicated_rule {
  id 0
  type replicated
  min_size 1
  max_size 10
  step take datacenter0
  step chooseleaf firstn 2 type host
  step emit
  step take datacenter1
  step chooseleaf firstn 2 type host
  step emit
}
```

**#2 - 03/22/2019 07:37 AM - Kefu Chai**

- Status changed from New to Pending Backport
- Backport set to mimic,nautilus
- Pull request ID set to 27068

**#3 - 03/22/2019 07:39 AM - xie xingguo**

- Backport changed from mimic,nautilus to luminous,mimic,nautilus

**#4 - 03/22/2019 01:01 PM - Nathan Cutler**

- Copied to Backport #38858: mimic: upmap broken the crush rule added

06/06/2021
#5 - 03/22/2019 01:01 PM - Nathan Cutler
- Copied to Backport #38859: luminous: upmap broken the crush rule added

#6 - 03/22/2019 01:01 PM - Nathan Cutler
- Copied to Backport #38860: nautilus: upmap broken the crush rule added

#7 - 04/10/2019 10:18 PM - Nathan Cutler
- Status changed from Pending Backport to Resolved