Answers (/community/s/group/0F90L0000001NDBSA2/) — Li_wei__ (/community/s/profile/0050L0000093E3HQAU) (Customer) asked a question.
February 28, 2012 at 5:46 PM (/community/s/question/0D50L00006BIvL4SAL/mapr-filesystem-io-performance-issue)

MapR filesystem IO performance issue

I have just install M5 on 2 nodes ,each nodes hava this hardware configuration:

2*6 core

48G mem

2*10Gb NIC

14*15000rpm sas disk, 2 for RedhatLinux OS ,the rest 12 is for MapR,distsetup like :

#disksetup -W 8 /home/disks

more /home/disks

/dev/sdb

/dev/sdc

/dev/sdd

/dev/sde

/dev/sdf

/dev/sdg

/dev/sdh

/dev/sdi

/dev/sdj

/dev/sdk

/dev/sdl

/dev/sdm

question is :

after created a volume on these two nodes, I do a dd test through NFS ,the result is bad:

#mount vip:/mapr /hd #vip is virtual IP for 2 NFS service which be started up in 2 two nodes;

#cd /hd/MapR/mt   #mt is volume mount point

#time dd if=/dev/zero of=2.test bs=32k count=4000000

4000000+0 records in

4000000+0 records out

131072000000 bytes (131 GB) copied, 264.22 seconds, 496 MB/s

!!!just 496MB/s for two nodes, I had tested DD by EXT3 for each node (do 2 raid5 group for 12 disks in BIOS),and get a 1000MB/s for each node.

what's wrong with it ,how to improve MapR filesystem performance ?

| Mt (/community/s/topic/0TO0L000004VOnVWAW/mt) | Time (/community/s/topic/0TO0L000004VOnWWAW/time) | Vip (/community/s/topic/0TO0L000004VOnXWAW/vip) | +3 more |

⌃ Upvote   💬 Answer   ⬈ Share   8 answers   37 views

Ravindra Pindikura (/community/s/profile/0050L0000093CQ7QAM) (MAPR TECHNOLOGIES)   ▼
7 years ago

 - Do you have nfs service running on only 1 node or both nodes ?

  - If not, are you running dd on the same node that is running the nfs service ?

- When you run multiple dd, does it help if you fire them on both the nodes in parallel (different files) ?

Upvote   Reply

Li_wei_ (/community/s/profile/0050L0000093E3HQAU) (Customer)
7 years ago

configuration for each node:

2 cpu 6 core,48G mem,2NIC with 10Gb ,14 sas disk with 15000rpm

Upvote   Reply

Ravindra Pindikura (/community/s/profile/0050L0000093CQ7QAM) (MAPR TECHNOLOGIES)
7 years ago

can you run a simple streaming test over hadoop ? There's a program we distribute called RWSpeedTest sitting inside maprfs-test-0.1.jar inside the /opt/mapr/hadoop/hadoop-0.20.0/lib directory ... it simply writes (or reads) a single file. To run it

1. set compression off in the dir where you will create the test file(s):

    % hadoop fs -mkdir /a/b/c

    % hadoop mfs -setcompression off /a/b/c

2. run the test -- it will print the MB/s obtained.

    % hadoop jar lib/maprfs-test-01.jar com.mapr.fs.RWSpeedTest

     usage: RWSpeedTest filename [-]megabytes uri

    % hadoop jar lib/maprfs-test-01.jar com.mapr.fs.RWSpeedTest /a/b/c/testfile 1000 maprfs:///

You can run several copies of the RWSpeedTest in parallel to see what is aggregate b/w you get (run each one on a diff file)

Upvote   Reply

Li_wei_ (/community/s/profile/0050L0000093E3HQAU) (Customer)
7 years ago

configuration for each node:

2*6 core,48G mem,2*10Gb NIC,14*15000rpm sas disk

Upvote   Reply

Li_wei_ (/community/s/profile/0050L0000093E3HQAU) (Customer)
7 years ago

I use hdparm to test disk IO performance ,result can reach 196MB/s,

so it look like disk IO is ok ,but I found average is 5M~8M/s for earch disk by nmon monitor when I do dd test .

So bottlenech is not in disk , what's the bottleneck ,NFS ?

----

hdparm -t /dev/sdc

/dev/sdc:

 Timing buffered disk reads: 590 MB in 3.00 seconds = 196.40 MB/sec

[root@node1 ~]# hdparm -t /dev/sdm

/dev/sdm:

 Timing buffered disk reads: 604 MB in 3.01 seconds = 200.95 MB/sec

[root@node1 ~]#

Upvote   Reply

Nabeel Moidu (/community/s/profile/005E0000000IT9aIAG) (MapR Technologies)
7 years ago

What's the replication factor for your volumes?

Have you tried parallel writes to the volume ?

Upvote   Reply

Li_wei_ (/community/s/profile/0050L0000093E3HQAU) (Customer)
7 years ago

replication factor =1;

I found result had no apparent diffeerence for 1 or 3.

When I run 2 DD into same volume ,run time will double and 230M/s for each dd process:

#time dd if=/dev/zero of=2.test bs=32k count=4000000 &

#time dd if=/dev/zero of=3.test bs=32k count=4000000 &

Upvote   Reply

Li_wei_ (/community/s/profile/0050L0000093E3HQAU) (Customer)
7 years ago

I have just install M5 on 2 nodes ,each nodes hava this hardware configuration:

2*6 core

48G mem

2*10Gb NIC

14*15000rpm sas disk, 2 for RedhatLinux OS ,the rest 12 is for MapR,distsetup like :

#disksetup -W 8 /home/disks

more /home/disks

/dev/sdb

/dev/sdc

/dev/sdd

/dev/sde

/dev/sdf

/dev/sdg

/dev/sdh

/dev/sdi

/dev/sdj

/dev/sdk

/dev/sdl

/dev/sdm

question is :

after created a volume on these two nodes, I do a dd test through NFS ,the result is bad:

#mount vip:/mapr /hd #vip is virtual IP for 2 NFS service which be started up in 2 two nodes;

#cd /hd/MapR/mt   #mt is volume mount point

#time dd if=/dev/zero of=2.test bs=32k count=4000000

4000000+0 records in

4000000+0 records out

131072000000 bytes (131 GB) copied, 264.22 seconds, 496 MB/s

!!!just 496MB/s for two nodes, I had tested DD by EXT3 for each node (do 2 raid5 group for 12 disks in BIOS),and get a 1000MB/s for each node.

what's wrong with it ,how to improve MapR filesystem performance ?

Upvote   Reply

Write an answer...