Storage - Bug #8144

check_mk disk monitoring for XFS is not good enough

03/23/2015 07:30 AM - Gerard Bernabeu Altayo

Status:	New		Start date:	03/23/2015
Priority:	Normal		**Due date:**	
Assignee:	Chih-Hao Huang		**% Done:**	0%
Category:			**Estimated time:**	3.00 hours
Target version:			**Spent time:**	0.00 hour

Description

When a system has broken mountpoints check_mk does not alert:

```
[root@cmsstor164 ~]# df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda3       213G 3.4G 199G   2% /
tmpfs           7.8G    0 7.8G   0% /dev/shm
/dev/sda1       976M  88M 838M  10% /boot
/dev/sdd         13T  12T 955G  93% /storage/data3
/dev/sdb         13T  12T 917G  93% /storage/data1
/dev/sdc         13T  12T 895G  94% /storage/data2
cmssrmtemp:/pnfs 1.0E  20P 1005P  2% /pnfs
[root@cmsstor164 ~]# ll /storage/data1
ls: cannot access /storage/data1: Input/output error
[root@cmsstor164 ~]# ll /storage/data2
ls: cannot access /storage/data2: Input/output error
[root@cmsstor164 ~]# ll /storage/data3
ls: cannot access /storage/data3: Input/output error
[root@cmsstor164 ~]# mount
/dev/sda3 on / type ext4 (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
devpts on /dev/pts type devpts (rw,gid=5,mode=620)
tmpfs on /dev/shm type tmpfs (rw)
/dev/sda1 on /boot type ext4 (rw)
/dev/sdd on /storage/data3 type xfs (rw,nobarrier,inode64)
/dev/sdb on /storage/data1 type xfs (rw,nobarrier,inode64)
/dev/sdc on /storage/data2 type xfs (rw,nobarrier,inode64)
none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw)
sunrpc on /var/lib/nfs/rpc_pipefs type rpc_pipefs (rw)
cmssrmtemp:/pnfs on /pnfs type nfs
(rw,noatime,nodiratime,nfsvers=3,intr,hard,rsize=65536,wsize=65536,noacl,addr=131.225.204.188)
[root@cmsstor164 ~]#
```

/var/log/messages is full of errors:

```
Mar 23 07:16:56 cmsstor164 kernel: XFS (sdb): xfs_log_force: error 5 returned.
Mar 23 07:16:56 cmsstor164 kernel: XFS (sdd): xfs_log_force: error 5 returned.
Mar 23 07:16:57 cmsstor164 kernel: XFS (sdc): xfs_log_force: error 5 returned.
```

Also dCache logs are showing errors:

```
[root@cmsstor164 ~]# tail /var/log/dcache/w-cmsstor164-tape-disk1Domain.log
at java.util.concurrent.FutureTask.run(FutureTask.java:166) ~[na:1.7.0_25]
at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.access$201(ScheduledThreadPoolExecutor.java:178)
~[na:1.7.0_25]
at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.run(ScheduledThreadPoolExecutor.java:292)
~[na:1.7.0_25]
... 3 common frames omitted
Caused by: java.io.IOException: Input/output error
at java.io.RandomAccessFile.writeBytes(Native Method) ~[na:1.7.0_25]
at java.io.RandomAccessFile.write(RandomAccessFile.java:499) ~[na:1.7.0_25]
at com.sleepycat.je.log.FileManager.writeToFile(FileManager.java:1467) ~[je-4.1.21.jar:na]
```

at com.sleepycat.je.log.FileManager.writeLogBuffer(FileManager.java:1347) ~[je-4.1.21.jar:na]
... 27 common frames omitted

Possible fixes:

1. Add puppet log error parser (like I used to have, if pool is not online=issue)

[root@cmsstor164 ~]# grep -i mode /var/log/dcache/w-cmsstor164-tape-disk1Domain.log | tail -2
23 Mar 2015 07:19:38 (w-cmsstor164-tape-disk1) [] Pool mode changed to disabled(fetch,store,stage,p2p-client,p2p-server): Pool disabled: I/O test failed
23 Mar 2015 07:20:38 (w-cmsstor164-tape-disk1) [] Pool mode changed to disabled(fetch,store,stage,p2p-client,p2p-server): Pool disabled: I/O test failed
[root@cmsstor164 ~]#

2. Do a check_mk sensor on the mounts (a touch or an ls?), dCache is already doing that in (1).

Option 1 is better (although is dCache specific), probably should do both to have a solution for EOS as well.

History

#1 - 05/26/2015 06:29 PM - Gerard Bernabeu Altayo

- File check_dcache_pic_wrapper-1.sh added

- Assignee changed from Gerard Bernabeu Altayo to Natalia Ratnikova

Hi Natalia,

I'd like you to work on this task.

About 'possible fixes' I just realized I meant 'dCache log parser', not puppet log parser.

I'm attaching the nagios sensor that PIC runs, it needs to be slightly addapted to our infrastructure AND the output needs to change a bit to be check_mk compliant.

Most questions about check_mk sensor formatting should be resolved from https://mathias-kettner.de/checkmk_localchecks.html

The idea is that check_mk will run this sensor on each dCache system, to start let's focus in the pools. You don't need full integration to start working on it; you can copy the script to a pool and start editing it and running it locally. Let me know if you have any further questions.

Thanks,
Gerard

#2 - 05/28/2015 05:36 PM - Natalia Ratnikova

Hi Gerard,

I noticed this ticket today, and i am confused:
The title is about fixing bug in disk mounts monitoring. And the attached script for PIC is for checking dcache processes and ports.
The mounts task has a high priority, and it is one month old. Is that something we need to fix urgently? Or you just want me to start work on improving monitoring in general, and learning how to write check_mk sensors?

Thanks, Natalia.

#3 - 05/29/2015 02:01 PM - Gerard Bernabeu Altayo

- Priority changed from High to Normal

We discussed this yesterday in person, but to keep proper track of it:

1. It was clarified in person already. Like the ticket says, that script will check dCache pool's status that originate from a FS check, which will get the XFS FS checked. The script will also detect other conditions, is a 'high level' sensor that will help monitor dCache better.

2. Priority=high was an artifact, I did not expect it to have any impact, sorry for the confusion. I just set it to Normal. Priorities are set in our in-person meetings; let me reinforce that this is lower priority than the cmsstor409/410 related tickets.

Gerard

#4 - 08/13/2015 02:16 PM - Gerard Bernabeu Altayo

- Assignee changed from Natalia Ratnikova to Chih-Hao Huang

Chih-Hao is working on dCache pool monitoring, moving this task to him

Files

check_dcache_pic_wrapper-1.sh	8.16 KB	05/26/2015	Gerard Bernabeu Altayo