

Arvados - Feature #12444

Compute nodes monitor the tmpdir space over time

10/11/2017 08:41 PM - Bryan Cosca

| | |
|---|----------------------------------|
| Status: Resolved | Start date: 10/11/2017 |
| Priority: Normal | Due date: |
| Assigned To: Lucas Di Pentima | % Done: 100% |
| Category: | Estimated time: 0.00 hour |
| Target version: 2018-08-01 Sprint | |
| Description | |
| Subtasks: | |
| Task # 13842: Review 12444-tmpdir-monitoring | Resolved |
| Related issues: | |
| Related to Arvados - Feature #13913: Crunchstat-summary graphs tmpdir usage | Resolved 07/09/2019 |

Associated revisions

Revision f5d7521c - 07/25/2018 08:38 PM - Lucas Di Pentima

Merge branch '12444-tmpdir-monitoring'
Closes #12444

Arvados-DCO-1.1-Signed-off-by: Lucas Di Pentima <ldipentima@veritasgenetics.com>

History

#1 - 10/11/2017 08:42 PM - Bryan Cosca

- Tracker changed from Bug to Feature

shoot, I think this should be in a different backlog but I don't have the sufficient permissions to move it.

#2 - 12/05/2017 04:29 PM - Tom Morris

- Project changed from OPS to Arvados
- Target version set to To Be Groomed

#3 - 01/10/2018 08:41 PM - Lucas Di Pentima

- Enhance crunchstat so that available free space is periodically logged along with the already present mem, cpu & i/o stats.
- To avoid having to be calling the df command, golang provides a syscall package with an appropriate Statfs function.
- node-info logs already record the available space and i-nodes at the start, so this addition would complement that information.

#4 - 01/10/2018 09:37 PM - Tom Morris

- Target version changed from To Be Groomed to Arvados Future Sprints
- Story points set to 1.0

#5 - 01/11/2018 03:00 PM - Tom Clegg

Should log all three figures (available, used, total). Generally available+used<total because the "available only to root" portion is not counted as available. Reporting tools can decide the most useful way to report.

#6 - 07/18/2018 03:47 PM - Tom Morris

- Target version changed from Arvados Future Sprints to 2018-08-01 Sprint

#7 - 07/18/2018 03:47 PM - Tom Morris

- Assigned To set to Lucas Di Pentima

#8 - 07/23/2018 01:29 PM - Lucas Di Pentima

- Status changed from New to In Progress

#9 - 07/24/2018 02:35 PM - Lucas Di Pentima

Updates at [e4c31590b](#) - branch 12444-tmpdir-monitoring
Test run: <https://ci.curoverse.com/job/developer-run-tests/821/>

Adds tmpdir stats to crunchstat reporting: available, used & total with usage increments.

#10 - 07/24/2018 04:42 PM - Lucas Di Pentima

There's a sdk/python test failure, looking into it.

#11 - 07/24/2018 05:40 PM - Lucas Di Pentima

As suggested by Peter, rebasing on latest master fixed the issue.

Now at [b211e857d304f7fbe8787d2b65a307da841d047b](#)
Test run: <https://ci.curoverse.com/job/developer-run-tests-sdk-python-ruby/108/>

#12 - 07/24/2018 05:44 PM - Peter Amstutz

```
err := syscall.Statfs("/tmp", &s)
```

Shouldn't be hardcoded. By default it should use \$TMPDIR, but it would be better to pass it in. Crunch-run should be updated to pass in runner.parentTemp.

```
total:      s.Blocks * bs,  
used:      (s.Blocks - s.Bavail) * bs,  
available:  s.Bavail * bs,
```

Should be used: (s.Blocks - s.Bfree) * bs

(because Bavail < Bfree)

```
r.Logger.Printf("tmpdir available:%d used:%d total:%d%s\n",  
nextSample.available, nextSample.used, nextSample.total, delta)
```

I don't think this formatting is consistent with the other crunchstat lines (which have the number first, then the name), should be something like:

```
r.Logger.Printf("statfs %d available %d used %d total%s\n",  
nextSample.available, nextSample.used, nextSample.total, delta)
```

#13 - 07/24/2018 08:19 PM - Lucas Di Pentima

Updates at [85f6919fa](#)
Test run: <https://ci.curoverse.com/job/developer-run-tests/824/>

Addressed above comments and added a notice on the log to report which directory is being monitored.

#14 - 07/25/2018 07:51 PM - Peter Amstutz

Lucas Di Pentima wrote:

```
Updates at 85f6919fa  
Test run: https://ci.curoverse.com/job/developer-run-tests/824/
```

Addressed above comments and added a notice on the log to report which directory is being monitored.

This LGTM.

Could you file a follow-on ticket to update crunchstat-summary to graph disk usage?

#15 - 07/25/2018 09:59 PM - Lucas Di Pentima

- Status changed from In Progress to Resolved
- % Done changed from 0 to 100

Applied in changeset [arvados!f5d7521ca506d63f631c603938cac5f40663bcca](#).

#16 - 07/26/2018 11:09 AM - Lucas Di Pentima

- Related to Feature #13913: Crunchstat-summary graphs tmpdir usage added

#17 - 11/13/2018 09:02 PM - Tom Morris

- Release set to 13