

FreeNAS - Feature #18428

Speed up snapshots listing

10/21/2016 07:14 AM - test test

Status: Resolved	Estimated time: 0.00 hour
Priority: Nice to have	
Assignee: William Grzybowski	
Category: OS	
Target version: 11.1-BETA1	
Severity: New	Needs Merging: Yes
Reason for Closing:	Needs Automation: No
Reason for Blocked:	Support Suite Ticket: n/a
Needs QA: No	Hardware Configuration:
Needs Doc: Yes	

Description

Due to

```
/sbin/zfs list -t snapshot -H -o name
```

being used instead of

```
/sbin/zfs list -t snapshot -H -o name -s name
```

it takes more than five times as long to complete.

I don't know if the script autosnap.py or others that are involved depend on the order given without -s name or if sorting after name is sufficient.

Could someone please have a look? We are talking about improvements like 5s vs 0.7s.

Thanks!

Associated revisions

Revision 08c3658d - 10/21/2016 10:39 AM - William Grzybowski

feat(snap): sort by name

Ticket: #18428

Revision 1d78ea76 - 12/08/2016 10:07 AM - William Grzybowski

Revert "feat(snap): sort by name"

This reverts commit 08c3658d0996387dce90ae2b6e8086dfa4a11b7c.

Ticket: #18428

(cherry picked from commit ff7c3c70560b973ced8d3147b1d8a572c2413063)

Revision 351f5a81 - 12/08/2016 10:07 AM - William Grzybowski

Revert "feat(snap): sort by name"

This reverts commit 08c3658d0996387dce90ae2b6e8086dfa4a11b7c.

Ticket: #18428
(cherry picked from commit ff7c3c70560b973ced8d3147b1d8a572c2413063)

Revision 42ea2db1 - 12/08/2016 10:15 AM - William Grzybowski

Revert "feat(snap): sort by name"

This reverts commit 08c3658d0996387dce90ae2b6e8086dfa4a11b7c.

Ticket: #18428
(cherry picked from commit ff7c3c70560b973ced8d3147b1d8a572c2413063)

Revision 584345ad - 12/09/2016 11:57 AM - William Grzybowski

Revert "feat(snap): sort by name"

This reverts commit 08c3658d0996387dce90ae2b6e8086dfa4a11b7c.

Ticket: #18428

Revision 46c49403 - 07/10/2017 11:01 AM - William Grzybowski

fix(snap): use zfs list -s name to speed up listing

Ticket: #18428

Revision a5d8cc66 - 07/21/2017 05:44 PM - William Grzybowski

fix(snap): use zfs list -s name to speed up listing

Ticket: #18428

Revision 5d16dfdf - 10/13/2017 11:56 AM - Dru Lavigne

Mention ZFS improvements.

Ticket: #13630

Ticket: #18428

History

#1 - 10/21/2016 07:20 AM - Bonnie Follweiler

- Assignee set to Kris Moore

#2 - 10/21/2016 07:57 AM - Kris Moore

- Assignee changed from Kris Moore to William Grzybowski

- Priority changed from No priority to Nice to have

- Target version set to 9.10.2

#3 - 10/21/2016 10:30 AM - William Grzybowski

- Status changed from *Unscreened* to *Screened*

For autosnap I dont foresee a problem, it would probably be for autorepl only.

#4 - 10/21/2016 10:40 AM - William Grzybowski

- Status changed from *Screened* to *19*

#5 - 11/22/2016 06:43 AM - William Grzybowski

- Status changed from *19* to *Needs Developer Review*

#6 - 11/22/2016 07:06 AM - William Grzybowski

- Assignee changed from *William Grzybowski* to *Josh Paetzel*

#7 - 11/28/2016 10:48 AM - Josh Paetzel

- Status changed from *Needs Developer Review* to *Reviewed*

- Assignee changed from *Josh Paetzel* to *William Grzybowski*

#8 - 12/08/2016 10:15 AM - William Grzybowski

- Status changed from *Reviewed* to *Closed: Not To Be Fixed*

Turns out this mucks with the deletion order of snapshots, which causes zfs destroy error (because recursive delete of parent)

#9 - 04/24/2017 07:03 PM - Bill T

- File *busy time.png* added

It isn't just the runtime of zfs list that is improved with this change, on my raid-z with 3 low RPM WD Red drives, the busy time dropped from ~20% (week 13-14 in the attached graph) down to below 5% (week 15 onward).

The huge spike in week 15 is when I ran a scrub, which dropped from ~36 hours to ~14 hours.

I did this by changing autosnap.py on freenas-9.10.2-U2 to:

```
272: zfsproc = pipeopen("/sbin/zfs list -t snapshot -H -o name -s name", debug, logger=log)
```

That's a lot of busy time to be reclaimed if the delete case can be addressed

#10 - 04/25/2017 05:14 AM - William Grzybowski

Bill T wrote:

It isn't just the runtime of zfs list that is improved with this change, on my raid-z with 3 low RPM WD Red drives, the busy time dropped from ~20% (week 13-14 in the attached graph) down to below 5% (week 15 onward).

The huge spike in week 15 is when I ran a scrub, which dropped from ~36 hours to ~14 hours.

I did this by changing autosnap.py on freenas-9.10.2-U2 to:

```
272: zfsproc = pipeopen("/sbin/zfs list -t snapshot -H -o name -s name", debug, logger=log)
```

That's a lot of busy time to be reclaimed if the delete case can be addressed

Yes, however the replication currently relies on that, so nothing can be done at this time regarding that.

#11 - 04/25/2017 10:01 PM - Bill T

Is the only problem deleting parents before children? If yes, then we can make this faster and not break recursive delete by sorting in python after calling zfs list, i.e.:

```
zfsproc = pipeopen("/sbin/zfs list -t snapshot -H -o name -s name", debug, logger=log)
lines = zfsproc.communicate()[0].split('\n')
# Sort snapshots, deepest filesystem first.
lines = sorted(lines, cmp=lambda a, b: b.count('/')-a.count('/'))
```

#12 - 04/27/2017 03:52 AM - William Grzybowski

- Status changed from Closed: Not To Be Fixed to Screened

- Target version changed from 9.10.2 to 11.1

I am unsure about that, i think there is more in the play here, e.g. txg transaction and the sorts. Also its possible doing the sorting in python is gonna be slower than the built-in version in C.

#13 - 04/27/2017 07:26 PM - Bill T

I'm fairly certain the cost of doing the sort in python will be negligible compared to the system time spent in 'zfs list' when not using '-s name'. On my system:

```
$ time zfs list -t snapshot -H -o name 2>&1 | grep total
zfs list -t snapshot -H -o name 2>&1  1.09s user 2.62s system 99% cpu 3.716 total

$ time zfs list -t snapshot -H -o name -s name 2>&1 | grep total
zfs list -t snapshot -H -o name -s name 2>&1  0.07s user 0.28s system 99% cpu 0.350 total
```

That's more than an order of magnitude slower when using the existing method. To see how fast python sorts:

```
$ zfs list -t snapshot -H -o name -s name > /tmp/snapshots
$ wc -l /tmp/snapshots
  7732 /tmp/snapshots
$ python -m timeit "lines = open('/tmp/snapshots').readlines()" "sorted(lines, cmp=lambda a, b: b.count('/')-a.count('/'))"
100 loops, best of 3: 17.2 msec per loop
```

And more importantly, for a storage focused platform, you're not consuming nearly as much disk iops.

Is there some unit / integration test I could run to validate my proposed change? I'd be willing to make the sort as fancy as necessary to get whatever failed in comment [#8](#) resolved.

#14 - 04/28/2017 04:46 AM - William Grzybowski

Bill T wrote:

I'm fairly certain the cost of doing the sort in python will be negligible compared to the system time spent in 'zfs list' when not using '-s name'. On my system:

[...]

That's more than an order of magnitude slower when using the existing method. To see how fast python sorts:

[...]

And more importantly, for a storage focused platform, you're not consuming nearly as much disk iops.

Totally agree there but since last experiment failed I am skeptical about rushing in trying it again. Its been long enough again I don't remember exact details it failed, our fault in lacking documentation of the matter.

Is there some unit / integration test I could run to validate my proposed change? I'd be willing to make the sort as fancy as necessary to get whatever failed in comment [#8](#) resolved.

Unfortunately not, QA team is working on integration test for this feature but this is still all manual.

#15 - 07/10/2017 11:02 AM - William Grzybowski

- Status changed from Screened to Ready For Release

<https://github.com/freenas/freenas/commit/46c494033f8e14da7d25069c0283b49dc8e50bb6>

#16 - 09/09/2017 10:54 AM - Dru Lavigne

- Subject changed from Make autosnap.py run faster to Speed up snapshots listing

#17 - 09/26/2017 04:18 PM - Dru Lavigne

- Target version changed from 11.1 to 11.1-BETA1

#18 - 10/24/2017 04:48 AM - Dru Lavigne

- Status changed from Ready For Release to Resolved

#19 - 10/30/2017 10:53 AM - Bonnie Follweiler

- Needs QA changed from Yes to No
- QA Status Test Passes FreeNAS added
- QA Status deleted (Not Tested)

Sorting nearly 2000 snapshots takes a few seconds

Files

busy time.png	12.1 KB	04/25/2017	Bill T
---------------	---------	------------	--------