

Arvados - Feature #17609

arvados-client subcommand to run diagnostics on already installed cluster

04/29/2021 02:48 PM - Peter Amstutz

Status:	Resolved	Start date:	06/09/2021
Priority:	Normal	Due date:	
Assigned To:	Tom Clegg	% Done:	100%
Category:	Deployment	Estimated time:	0.00 hour
Target version:	2021-06-23 sprint		
Description			
This is the list of tests will do			
https://docs.google.com/spreadsheets/d/1--O03eo9-5gQYnP5eBti9a6E6ZYApM_lpnRsYZo9pqM/edit#gid=0 https://docs.google.com/spreadsheets/d/1--O03eo9-5gQYnP5eBti9a6E6ZYApM_lpnRsYZo9pqM/edit#gid=0			
Then once we have the list will include it to the arvados-client test			
<ul style="list-style-type: none">• Run the tests that can be run:<ul style="list-style-type: none">◦ If config.yml is available, check that◦ If cypress can be run, run browser-based tests• Warn about what can be run / cannot be run• put everything into a diagnostics project			
Ward's 3 electric rails:			
<ul style="list-style-type: none">• uploading through keepproxy• running workflows• properly configured keep-web<ul style="list-style-type: none">◦ uploading via webdav◦ downloading via webdav and s3			
Nico's tests:			
<ul style="list-style-type: none">• Fetching discovery document / public config• Check hostnames, ports, certificates of service ExternalURL are valid• Check nginx geo section			
Tom's modes:			
<ul style="list-style-type: none">• User option to run assuming it is inside (check that things treat you as inside)• User option to run assuming it is outside (check that things treat you as outside)			
Healthcheck:			
<ul style="list-style-type: none">• Use healthcheck endpoints, see if some tests can be part of healthcheck<ul style="list-style-type: none">◦ Any check that can be done as a healthcheck, probably should be• Needs management token• Use healthcheck aggregator			
<pre>\$ arvados-client diagnostics --inside Checking connectivity to https://api.arvados.example.com ...OK Checking TLS certificate on https://api.arvados.example.com ...FAIL</pre>			
Guidelines:			
<ul style="list-style-type: none">• run arvados-server check-config as early as possible.• verbose mode that communicates as much as possible about what each test is trying to do• be very explicit about failures			
Subtasks:			
Task # 17731: Review 17609-diagnostics-cmd			Resolved
Related issues:			

Associated revisions

Revision e627df27 - 06/18/2021 02:58 PM - Tom Clegg

Merge branch '17609-diagnostics-cmd'

closes #17609

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tom@curii.com>

History

#1 - 04/29/2021 02:59 PM - Peter Amstutz

- Subject changed from *Installed cluster diagnostic test* to *arvados-client subcommand to run diagnostics on already installed cluster*

#2 - 04/29/2021 03:06 PM - Nico César

- Assigned To set to Nico César

#3 - 04/29/2021 03:11 PM - Nico César

- Description updated

#4 - 05/11/2021 06:01 PM - Peter Amstutz

- Description updated

#5 - 05/11/2021 06:02 PM - Peter Amstutz

- Related to Story #16444: Improved error detection/reporting added

#6 - 05/11/2021 06:02 PM - Peter Amstutz

- Assigned To deleted (Nico César)

#7 - 05/12/2021 04:02 PM - Tom Clegg

- Assigned To set to Tom Clegg

#8 - 05/25/2021 02:42 PM - Tom Clegg

For discussion:

```
$ arvados-client diagnostics
INFO discovery document: ok, BlobSignatureTTL is 1209600
INFO exported config: ok, Collections.BlobSigning = true
INFO api call (get current user): ok, uuid = ce8i5-tpzed-ol81h55xqo4i23l
INFO http connection: https://keep.ce8i5.arvadosapi.com/: ok
INFO http connection: https://*.collections.ce8i5.arvadosapi.com/: ok
ERRO http connection: https://52.147.168.174/: error: Make-coffee "https://52.147.168.174/": x509: cannot validate certificate for 52.147.168.174 because it doesn't contain any IP SANs
WARN service url path seems unlikely to work: wss://ws.ce8i5.arvadosapi.com/websocket
INFO http connection: wss://ws.ce8i5.arvadosapi.com/websocket: ok
INFO http connection: https://workbench.ce8i5.arvadosapi.com/: ok
INFO http connection: https://workbench2.ce8i5.arvadosapi.com/: ok
INFO cors header: https://ce8i5.arvadosapi.com/: ok
INFO cors header: https://keep.ce8i5.arvadosapi.com/d41d8cd98f00b204e9800998ecf8427e+0: ok
ERRO cors header: https://52.147.168.174/: error: Get "https://52.147.168.174/": x509: cannot validate certificate for 52.147.168.174 because it doesn't contain any IP SANs
INFO api call (list projects): ok, using existing project, uuid = ce8i5-j7d0g-tqax6h79xj5884w
INFO api call (create collection): ok, uuid = ce8i5-4zz18-viqkal3gsqajo9s
ERRO webdav upload: error performing http request: Put "https://52.147.168.174/c=ce8i5-4zz18-viqkal3gsqajo9s/testfile": x509: cannot validate certificate for 52.147.168.174 because it doesn't contain any IP SANs
INFO webdav external url https://*.collections.ce8i5.arvadosapi.com/ looks ok
INFO webdav download https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/foo: ok
INFO webdav download https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/testfile: ok
ERRO webdav download https://52.147.168.174/c=d41d8cd98f00b204e9800998ecf8427e+0/_/foo: Get "https://52.147.168.174/c=d41d8cd98f00b204e9800998ecf8427e+0/_/foo": x509: cannot validate certificate for 52.147.168.174 because it doesn't contain any IP SANs
ERRO webdav download https://52.147.168.174/c=d41d8cd98f00b204e9800998ecf8427e+0/_/testfile: Get "https://52.147.168.174/c=d41d8cd98f00b204e9800998ecf8427e+0/_/testfile": x509: cannot validate certificate for 52.147.168.174 because it doesn't contain any IP SANs
ERRO webdav download https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/testfile: une
```

```
xpected response status: 404 Not Found
ERRO webdav download https://52.147.168.174/c=ce8i5-4zz18-viqkal3gsqajo9s/_/testfile: Get "https://52.147.168.174/c=ce8i5-4zz18-viqkal3gsqajo9s/_/testfile": x509: cannot validate certificate for 52.147.168.174 because it doesn't contain any IP SANs
```

tbd

- exit non-zero on warnings/errors?
- machine-readable output?
- highlight errors/warnings with color or other ascii art?

#9 - 05/25/2021 05:13 PM - Peter Amstutz

Nico: in order to run diagnostics, the user must specify the situation it is being run from:

- Inside the cluster, with config.yml available
- Inside the cluster, without config.yml
- Outside the cluster, without config.yml

#10 - 05/25/2021 05:31 PM - Peter Amstutz

Output needs to be very explicit:

- exactly what is each check doing
- if something failed, explain the implication of that failure

Can offer different verbose levels: only failures, all tests, tests + lots of extra debug info

Assign ids to individual tests, have ability to run specific tests

#11 - 05/26/2021 03:19 PM - Peter Amstutz

- Status changed from New to In Progress

#12 - 05/26/2021 03:19 PM - Peter Amstutz

- Target version changed from 2021-05-26 sprint to 2021-06-09 sprint

#13 - 06/01/2021 08:42 PM - Tom Clegg

For the first version, I'm aiming to include

- ability to catch a couple of configuration problems we see in the wild when troubleshooting (internal/external client detection, wildcard dns/tls, unreachable service URLs)
- enough log detail that if someone sends me their output via email/chat I'll have a decent chance of figuring out what's wrong with their setup
- run a container

There are some things I'm not trying to include (although they are desirable and will hopefully happen in future)

- explain the implications of any given thing being broken
- explain how to fix any given thing
- read cluster config from /etc/arvados/ (as an alternative to using env vars and exported config)

I don't have "run a container" yet. In the meantime here is some sample output:

```
$ arvados-client diagnostics -external-client
INFO 10 getting discovery document from https://ce8i5.arvadosapi.com/discovery/v1/apis/arvados/v1/rest
INFO 20 getting exported config from https://ce8i5.arvadosapi.com/arvados/v1/config
INFO 30 getting current user record
INFO 40 connecting to service endpoint https://keep.ce8i5.arvadosapi.com/
INFO 41 connecting to service endpoint https://*.collections.ce8i5.arvadosapi.com/
INFO 42 connecting to service endpoint https://download.ce8i5.arvadosapi.com/
INFO 43 connecting to service endpoint wss://ws.ce8i5.arvadosapi.com/websocket
INFO 44 connecting to service endpoint https://workbench.ce8i5.arvadosapi.com/
INFO 45 connecting to service endpoint https://workbench2.ce8i5.arvadosapi.com/
INFO 50 checking CORS headers at https://ce8i5.arvadosapi.com/
INFO 51 checking CORS headers at https://keep.ce8i5.arvadosapi.com/d41d8cd98f00b204e9800998ecf8427e+0
INFO 52 checking CORS headers at https://download.ce8i5.arvadosapi.com/
INFO 60 checking internal/external client detection
INFO 61 reading+writing via keep service at https://keep.ce8i5.arvadosapi.com:443/
INFO 80 finding/creating "scratch area for diagnostics" project
INFO 90 creating temporary collection
INFO 100 uploading file via webdav
INFO 110 checking WebDAV ExternalURL wildcard (https://*.collections.ce8i5.arvadosapi.com/)
INFO 120 downloading from webdav (https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/)
```

```
foo)
INFO 121 downloading from webdav (https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/
testfile)
INFO 122 downloading from webdav (https://download.ce8i5.arvadosapi.com/c=d41d8cd98f00b204e9800998ecf8427e+0/_
/foo)
INFO 123 downloading from webdav (https://download.ce8i5.arvadosapi.com/c=d41d8cd98f00b204e9800998ecf8427e+0/_
/testfile)
INFO 124 downloading from webdav (https://3ec3d27e95a51d659178f6350fd8d9bf-52.collections.ce8i5.arvadosapi.com
/testfile)
INFO 125 downloading from webdav (https://download.ce8i5.arvadosapi.com/c=ce8i5-4zz18-vn7r69bv85902ow/_/testfi
le)
INFO 130 getting list of virtual machines
INFO 140 getting workbench1 webshell page
INFO 150 connecting to webshell service
ERROR connecting to webshell service (10009ms): Post "https://webshell.ce8i5.arvadosapi.com/shell.ce8i5.arvado
sapi.com?": context deadline exceeded
INFO 9990 deleting temporary collection
```

--- cut here --- error summary ---

```
ERROR connecting to webshell service (10009ms): Post "https://webshell.ce8i5.arvadosapi.com/shell.ce8i5.arvado
sapi.com?": context deadline exceeded
exit status 1
```

#14 - 06/08/2021 09:49 PM - Tom Clegg

- Status changed from In Progress to New

Added "run a container" test. For now it's just an easy way to check whether it works. We will certainly want to add more features to illuminate *why* it fails, when it fails.

Example output:

```
$ arvados-client diagnostics -log-level=debug
INFO 10: getting discovery document from https://ce8i5.arvadosapi.com/discovery/v1/apis/arvados/v1/rest
DEBUG ... BlobSignatureTTL = 1209600
DEBUG 10: getting discovery document from https://ce8i5.arvadosapi.com/discovery/v1/apis/arvados/v1/rest (
1010 ms): ok
INFO 20: getting exported config from https://ce8i5.arvadosapi.com/arvados/v1/config
DEBUG ... Collections.BlobSigning = true
DEBUG 20: getting exported config from https://ce8i5.arvadosapi.com/arvados/v1/config (123 ms): ok
INFO 30: getting current user record
DEBUG ... user uuid = ce8i5-tpzed-ol81h55xqo4i231
DEBUG 30: getting current user record (152 ms): ok
INFO 40: connecting to service endpoint https://keep.ce8i5.arvadosapi.com/
DEBUG 40: connecting to service endpoint https://keep.ce8i5.arvadosapi.com/ (310 ms): ok
INFO 41: connecting to service endpoint https://*.collections.ce8i5.arvadosapi.com/
DEBUG 41: connecting to service endpoint https://*.collections.ce8i5.arvadosapi.com/ (382 ms): ok
INFO 42: connecting to service endpoint https://download.ce8i5.arvadosapi.com/
DEBUG 42: connecting to service endpoint https://download.ce8i5.arvadosapi.com/ (376 ms): ok
INFO 43: connecting to service endpoint wss://ws.ce8i5.arvadosapi.com/websocket
DEBUG 43: connecting to service endpoint wss://ws.ce8i5.arvadosapi.com/websocket (454 ms): ok
INFO 44: connecting to service endpoint https://workbench.ce8i5.arvadosapi.com/
DEBUG 44: connecting to service endpoint https://workbench.ce8i5.arvadosapi.com/ (1624 ms): ok
INFO 45: connecting to service endpoint https://workbench2.ce8i5.arvadosapi.com/
DEBUG 45: connecting to service endpoint https://workbench2.ce8i5.arvadosapi.com/ (367 ms): ok
INFO 50: checking CORS headers at https://ce8i5.arvadosapi.com/
DEBUG 50: checking CORS headers at https://ce8i5.arvadosapi.com/ (115 ms): ok
INFO 51: checking CORS headers at https://keep.ce8i5.arvadosapi.com/d41d8cd98f00b204e9800998ecf8427e+0
DEBUG 51: checking CORS headers at https://keep.ce8i5.arvadosapi.com/d41d8cd98f00b204e9800998ecf8427e+0 (1
06 ms): ok
INFO 52: checking CORS headers at https://download.ce8i5.arvadosapi.com/
DEBUG 52: checking CORS headers at https://download.ce8i5.arvadosapi.com/ (104 ms): ok
INFO 60: checking internal/external client detection
DEBUG ... controller returned only proxy services, this host is treated as "external"
DEBUG 60: checking internal/external client detection (162 ms): ok
INFO 61: reading+writing via keep service at https://keep.ce8i5.arvadosapi.com:443/
DEBUG 61: reading+writing via keep service at https://keep.ce8i5.arvadosapi.com:443/ (258 ms): ok
INFO 80: finding/creating "scratch area for diagnostics" project
DEBUG ... using existing project, uuid = ce8i5-j7d0g-tqax6h79xj5884w
DEBUG 80: finding/creating "scratch area for diagnostics" project (145 ms): ok
INFO 90: creating temporary collection
DEBUG ... ok, uuid = ce8i5-4zz18-3qrqhs7u0uy2vau
DEBUG 90: creating temporary collection (178 ms): ok
INFO 100: uploading file via webdav
```

```
DEBUG    ... ok, status 201 Created
DEBUG    ... ok, pdh 3ec3d27e95a51d659178f6350fd8d9bf+52
DEBUG    100: uploading file via webdav (397 ms): ok
INFO     110: checking WebDAV ExternalURL wildcard (https://*.collections.ce8i5.arvadosapi.com/)
DEBUG    110: checking WebDAV ExternalURL wildcard (https://*.collections.ce8i5.arvadosapi.com/) (0 ms): ok
INFO     120: downloading from webdav (https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/foo)
DEBUG    120: downloading from webdav (https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/foo) (104 ms): ok
INFO     121: downloading from webdav (https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/testfile)
DEBUG    121: downloading from webdav (https://d41d8cd98f00b204e9800998ecf8427e-0.collections.ce8i5.arvadosapi.com/testfile) (102 ms): ok
INFO     122: downloading from webdav (https://download.ce8i5.arvadosapi.com/c=d41d8cd98f00b204e9800998ecf8427e+0/_/foo)
DEBUG    122: downloading from webdav (https://download.ce8i5.arvadosapi.com/c=d41d8cd98f00b204e9800998ecf8427e+0/_/foo) (103 ms): ok
INFO     123: downloading from webdav (https://download.ce8i5.arvadosapi.com/c=d41d8cd98f00b204e9800998ecf8427e+0/_/testfile)
DEBUG    123: downloading from webdav (https://download.ce8i5.arvadosapi.com/c=d41d8cd98f00b204e9800998ecf8427e+0/_/testfile) (101 ms): ok
INFO     124: downloading from webdav (https://3ec3d27e95a51d659178f6350fd8d9bf-52.collections.ce8i5.arvadosapi.com/testfile)
DEBUG    124: downloading from webdav (https://3ec3d27e95a51d659178f6350fd8d9bf-52.collections.ce8i5.arvadosapi.com/testfile) (364 ms): ok
INFO     125: downloading from webdav (https://download.ce8i5.arvadosapi.com/c=ce8i5-4zz18-3qrqhs7u0uy2vau/_/testfile)
DEBUG    125: downloading from webdav (https://download.ce8i5.arvadosapi.com/c=ce8i5-4zz18-3qrqhs7u0uy2vau/_/testfile) (119 ms): ok
INFO     130: getting list of virtual machines
DEBUG    130: getting list of virtual machines (203 ms): ok
INFO     140: getting workbench1 webshell page
DEBUG    ... url https://workbench.ce8i5.arvadosapi.com/virtual_machines/ce8i5-2x53u-submlh2cc0lnkvg/webshell/testusername
DEBUG    140: getting workbench1 webshell page (1863 ms): ok
INFO     150: connecting to webshell service
DEBUG    ... url https://webshell.ce8i5.arvadosapi.com/shell.ce8i5.arvadosapi.com?
ERROR    150: connecting to webshell service (10000 ms): Post "https://webshell.ce8i5.arvadosapi.com/shell.ce8i5.arvadosapi.com?": context deadline exceeded
INFO     160: running a container
DEBUG    ... container request uid = ce8i5-xvhdp-ta72wgo7cj0j3ei
DEBUG    ... container uid = ce8i5-dz642-wx76ibvve58bfxf
INFO     ... container request submitted, waiting up to 10m for container to run
DEBUG    ... container state = Queued
DEBUG    ... container state = Locked
DEBUG    ... container state = Running
DEBUG    ... container request state = Final
DEBUG    ... container state = Complete
DEBUG    160: running a container (196973 ms): ok
INFO     9990: deleting temporary collection
DEBUG    9990: deleting temporary collection (179 ms): ok
```

--- cut here --- error summary ---

```
ERROR    150: connecting to webshell service (10000 ms): Post "https://webshell.ce8i5.arvadosapi.com/shell.ce8i5.arvadosapi.com?": context deadline exceeded
exit status 1
```

17609-diagnostics-cmd @ [08c8b9cc496627bc3fd3d87ae333fadce4797eaa](#)

#15 - 06/09/2021 02:05 PM - Peter Amstutz

- Status changed from New to In Progress

#16 - 06/09/2021 03:49 PM - Peter Amstutz

- Target version changed from 2021-06-09 sprint to 2021-06-23 sprint

#17 - 06/10/2021 02:48 PM - Ward Vandewege

Reviewing [08c8b9cc496627bc3fd3d87ae333fadce4797eaa](#):

- moving the NoPrefixFormatter to lib/cmd is nice. We should also remove the copy in lib/deduplicationreport.
- the NoPrefixFormatter that was used in lib/costanalyzer did not have a trailing newline. Since you switched to the one in lib/cmd the output for

costanalyzer is now a bit wonky. Can we have a flag to NoPrefixFormatter that indicates if a newline should be added or not? Or is there another/better way to do this in costanalyzer?

- the `runtests()` function is very long, but I'm not sure what can be done about that that wouldn't just add needless boilerplate/overhead.

LGTM otherwise!

#18 - 06/11/2021 02:34 PM - Tom Clegg

Ward Vandewege wrote:

- remove the copy in lib/deduplicationreport.

Done.

- the NoPrefixFormatter that was used in lib/costanalyzer did not have a trailing newline

Oh right, I forgot to follow up on that. Now removed the newlines from the format strings in the logger.Info() etc. calls in costanalyzer.

I changed two recently-added fmt.Print to logger.Debug since they seemed a bit repetitive:

```
-Considering ce8i5-xvhdp-jn8lm49n18abkix
-Processing ce8i5-xvhdp-jn8lm49n18abkix
Collecting child containers for container request ce8i5-xvhdp-jn8lm49n18abkix
(2021-05-04 23:28:33.846312 +0000 UTC)
```

Since logging is line-oriented, the "print a dot for each child" progress meter doesn't fit very well. Rather than open things up and pass stdout/stderr through to the right place for that purpose, I added a thing that prints "... 123 of 456" every 5 seconds. Do you think that's a reasonable alternative?

17609-diagnostics-cmd @ [056b3d2368b151a626fbf79025d9989a4d29a018](#) -- <https://ci.arvados.org/view/Developer/job/developer-run-tests/2519/>

- the `runtests()` function is very long, but I'm not sure what can be done about that that wouldn't just add needless boilerplate/overhead.

Yeah, I think we will need to split this up sooner or later. I'm thinking one big section to do the initial staging setup (current user, scratch project/collection) and then something more modular for an ever-growing set of test funcs that are independent of one another. Not sure whether I should do it right now, or put it on the todo list. Something like this might work, with minimal repetition:

```
type testenv struct {
    project   arvados.Group
    collection arvados.Collection
}

var _ = addtest(1234, "testing some stuff", func(env *testenv) error { ... })
```

#19 - 06/18/2021 02:26 PM - Ward Vandewege

Tom Clegg wrote:

Ward Vandewege wrote:

- remove the copy in lib/deduplicationreport.

Done.

- the NoPrefixFormatter that was used in lib/costanalyzer did not have a trailing newline

Oh right, I forgot to follow up on that. Now removed the newlines from the format strings in the logger.Info() etc. calls in costanalyzer.

I changed two recently-added fmt.Print to logger.Debug since they seemed a bit repetitive:

Thanks that's good.

[...]

Since logging is line-oriented, the "print a dot for each child" progress meter doesn't fit very well. Rather than open things up and pass stdout/stderr through to the right place for that purpose, I added a thing that prints "... 123 of 456" every 5 seconds. Do you think that's a reasonable alternative?

Yes, that even provides a bit more information (the 'total' is known up front) so it's a nice improvement.

17609-diagnostics-cmd @ [056b3d2368b151a626bf79025d9989a4d29a018](https://ci.arvados.org/view/Developer/job/developer-run-tests/2519/) --
<https://ci.arvados.org/view/Developer/job/developer-run-tests/2519/>

- the `runtests()` function is very long, but I'm not sure what can be done about that that wouldn't just add needless boilerplate/overhead.>

Yeah, I think we will need to split this up sooner or later. I'm thinking one big section to do the initial staging setup (current user, scratch project/collection) and then something more modular for an ever-growing set of test funcs that are independent of one another. Not sure whether I should do it right now, or put it on the todo list. Something like this might work, with minimal repetition:

[...]

That sounds better yeah, up to you if you want to fit that in before this merge.

LGTM.

#20 - 06/18/2021 03:28 PM - Tom Clegg

- % Done changed from 0 to 100

- Status changed from In Progress to Resolved

Applied in changeset [arvados!e627df2797dae0d6fa95da61f1a58bb9fafa8240](https://ci.arvados.org/view/Developer/job/developer-run-tests/2519/).

#21 - 11/16/2021 04:26 PM - Peter Amstutz

- Release set to 42