Arvados - Feature #14812

[Workbench1] Load configuration from cluster config file

02/06/2019 06:53 PM - Peter Amstutz

Status:	Resolved		**Start date:**	07/02/2019
Priority:	Normal		**Due date:**	
Assigned To:	Peter Amstutz		**% Done:**	100%
Category:			**Estimated time:**	0.00 hour
Target version:	2019-07-17 Sprint			

Description

At startup, Workbench should

1. execute arvados-server config-dump < /etc/arvados/config.yml (if that fails, error out)
2. load the resulting YAML document
3. use that document (instead of .../config/application.yml etc) as the application configuration

The main dev tasks are

- Loading the new config at startup
- Changing the config key uses to the new keys

Implementation notes:

- If .../config/application.yml exists, values there should take precedence. This should be done with the same code used in RailsAPI. "config dump" and "config diff" rake tasks should be provided, just like in RailsAPI.
- Install guide and upgrading/release notes need to be updated to ensure the "arvados-server" package is installed on the workbench host.
- Update Arvbox to install arvados-server.

Subtasks:	
Task # 15395: Review 14812-wb1-new-config	**Resolved**
Task # 15450: Fix packaging (ops)	**Resolved**

Related issues:		
Related to Arvados - Story #13648: [Epic] Use one cluster configuration file ...	**Resolved**	
Related to Arvados - Support #15449: Workbench and API server packages depend...	**Resolved**	
Blocked by Arvados - Story #13996: [API] Migrate RailsAPI to new cluster conf...	**Resolved**	**03/26/2019**

Associated revisions

Revision 73b441ef - 07/01/2019 09:50 PM - Peter Amstutz

Tweak arvbox build order, API server depends on arvados-server

refs #14812

Arvados-DCO-1.1-Signed-off-by: Peter Amstutz <pamstutz@veritasgenetics.com>

Revision 94972ae2 - 07/09/2019 02:00 PM - Peter Amstutz

Merge branch '14812-wb1-new-config' refs #14812

Arvados-DCO-1.1-Signed-off-by: Peter Amstutz <pamstutz@veritasgenetics.com>

Revision dd0104a2 - 07/09/2019 02:00 PM - Peter Amstutz

Arvbox won't create application.yml if arvados_config.rb exists.

refs #14812

Arvados-DCO-1.1-Signed-off-by: Peter Amstutz <pamstutz@veritasgenetics.com>

Revision f3293c48 - 07/09/2019 10:27 PM - Ward Vandewege

Our workbench and api server packages now have a runtime dependency on
arvados-server for configuration validation.

Our workbench package now also has a build-time dependency on arvados-server,
because it needs some configuration while precompiling assets.

closes #15449
refs #14812

Arvados-DCO-1.1-Signed-off-by: Ward Vandewege <wvandewege@veritasgenetics.com>

Revision cf43300e - 12/19/2019 04:27 PM - Ward Vandewege

A few fixes for workbench in the way it displays the ARVADOS_API_HOST
and ssh ProxyCommand.

refs #14812

Arvados-DCO-1.1-Signed-off-by: Ward Vandewege <ward@jhvc.com>

History

#1 - 02/06/2019 06:53 PM - Peter Amstutz

- Status changed from New to In Progress

#2 - 02/06/2019 06:53 PM - Peter Amstutz

- Status changed from In Progress to New

#3 - 02/06/2019 07:11 PM - Peter Amstutz

- Related to Story #13648: [Epic] Use one cluster configuration file for all components added

#4 - 03/20/2019 02:50 PM - Peter Amstutz

- Blocked by Story #13996: [API] Migrate RailsAPI to new cluster config file added

#5 - 03/20/2019 03:18 PM - Peter Amstutz

- Description updated

#6 - 06/05/2019 03:13 PM - Peter Amstutz

- Blocked by Feature #15000: [controller] publish safe config added

#7 - 06/05/2019 03:14 PM - Peter Amstutz

- Subject changed from [Workbench1] Use cluster config to [Workbench1] Use safe config from controller

- Description updated

#8 - 06/12/2019 04:37 PM - Tom Clegg

- Subject changed from [Workbench1] Use safe config from controller to [Workbench1] Load configuration from cluster config file

- Description updated

#9 - 06/12/2019 05:04 PM - Tom Clegg

- Description updated

#10 - 06/12/2019 05:22 PM - Tom Clegg

- Description updated

#11 - 06/12/2019 05:24 PM - Tom Morris

- Target version changed from To Be Groomed to Arvados Future Sprints

- Story points set to 2.0

#12 - 06/13/2019 07:49 PM - Tom Clegg

- Blocked by deleted (Feature #15000: [controller] publish safe config)

#13 - 06/19/2019 03:32 PM - Tom Morris

- Target version changed from Arvados Future Sprints to 2019-07-03 Sprint

#14 - 06/19/2019 03:33 PM - Peter Amstutz

- Assigned To set to Peter Amstutz

#15 - 06/25/2019 07:35 PM - Peter Amstutz

- Status changed from New to In Progress

#16 - 06/25/2019 08:45 PM - Tom Clegg

Instead of a full-on template approach, perhaps we can support two options for the keep-web URLs:

1. Leading *, like "https://*.collections.zzzzz.arvadosapi.com" or "https://*--collections..../", means replace the leading * with the relevant UUID/PDH
2. No leading *, like "https://download.zzzzz.arvadosapi.com", means append /c=uuid_or_pdh

For webshell, the options might be

1. Leading *, like "https://*.webshell.zzzzz.arvadosapi.com", means replace the leading * with the relevant VM name
2. No leading *, like "https://webshell.zzzzz.arvadosapi.com", is an error, or means no webshell support

#17 - 06/26/2019 03:29 PM - Peter Amstutz

Tom Clegg wrote:

> Instead of a full-on template approach, perhaps we can support two options for the keep-web URLs:
>
> 1. Leading *, like "https://*.collections.zzzzz.arvadosapi.com" or "https://*--collections..../", means replace the leading * with the relevant UUID/PDH
> 2. No leading *, like "https://download.zzzzz.arvadosapi.com", means append /c=uuid_or_pdh
>
> For webshell, the options might be
>
> 1. Leading *, like "https://*.webshell.zzzzz.arvadosapi.com", means replace the leading * with the relevant VM name
> 2. No leading *, like "https://webshell.zzzzz.arvadosapi.com", is an error, or means no webshell support

So basically, instead of search-and-replace for '%{uuid_or_pdh}' it would search-and-replace for '*' ? Changing this means both migrating the existing config and changing how Workbench interprets it to the new config, so there should probably be a better reason to change it than just aesthetics.

#18 - 06/26/2019 05:12 PM - Tom Clegg

Peter Amstutz wrote:

> So basically, instead of search-and-replace for '%{uuid_or_pdh}' it would search-and-replace for '*'?

I'm trying to get away from "tell us how to make a URL" in favor of "tell us which domain names / TLS certs you have, and we'll use that information to build URLs that work."

So, instead of the two typical examples being

- https://%{uuid_or_pdh}.collections.example.com
- https://collections.example.com/c=%{uuid_or_pdh}

they would be

- https://*.collections.example.com
- https://collections.example.com

In the second case, Arvados would know (without being told by the config file) how/when to append the .../c=UUID part.

#19 - 07/02/2019 06:28 PM - Peter Amstutz

14812-wb1-new-config @ cf98484b422c29dfa251da5a176e18096f09aa7a

https://ci.curoverse.com/view/Developer/job/developer-run-tests/1359/

- Migrate wb1 application.yml to new config structure
- Update wb1 code to new config structure.
- Update both wb1 and API server to get configuration by running "arvados-server config-defaults" and "arvados-server config-dump"
- Update config.go and defaults file for a few keys that were missed before
- Restructure a couple config items to accommodate the "no lists" design rule

#20 - 07/02/2019 06:40 PM - Tom Clegg

- Update both wb1 and API server to get configuration by running "arvados-server config-defaults" and "arvados-server config-dump"

You can get defaults (with a cluster ID of your choice) by doing something like this

```
echo 'Clusters: {aaaaa: {}}' | arvados-server config-dump -config -
```

I wonder if it would be better for apiserver/wb1 to use that?

There are probably other uses for a "arvados-server config-defaults" command but I wonder if they would be better served by outputting the literal text complete with comments, samples, etc.

#21 - 07/02/2019 07:03 PM - Peter Amstutz

Tom Clegg wrote:

> - Update both wb1 and API server to get configuration by running "arvados-server config-defaults" and "arvados-server config-dump"
>
> You can get defaults (with a cluster ID of your choice) by doing something like this
>
> [...]
>
> I wonder if it would be better for apiserver/wb1 to use that?

Ok, I had tried "echo '{}' | arvados-server config-dump -config -" and but got "config does not define any clusters". I didn't think to provide a bogus cluster id. Your trick works, so I'll see if I can use it.

> There are probably other uses for a "arvados-server config-defaults" command but I wonder if they would be better served by outputting the literal text complete with comments, samples, etc.

I originally wanted to just dump the literal text, but I needed the SAMPLE sections filtered out. Since I can get that with your empty cluster trick, I will change config-defaults to just print the text.

#22 - 07/02/2019 07:17 PM - Lucas Di Pentima

Made a quick grep of configuration usage and found:

- There's a mention of profiling_enabled config on apps/workbench/app/models/arvados_api_client.rb:64 that seems to not be migrated
- There's a mention of api_response_compression on apps/workbench/app/models/arvados_api_client.rb:104 that it is already migrated to Workbench.APIResponseCompression
- There's a mention of testing_override_login_url on apps/workbench/app/models/arvados_api_client.rb lines 239 and 240 but it seems is for testing purposes only, so I'm not sure if it belongs to the cluster config. Mentioning it just in case.
- There's a mention of enable_public_projects_page on apps/workbench/app/views/application/_breadcrumbs.html.erb:34 that's already migrated to Workbench.EnablePublicProjectsPage
- Side note: there's a mention of Rails.configuration.include_accept_encoding_header_in_api_requests on apps/workbench/test/controllers/application_controller_test.rb:452 … it seems to me that this test doesn't do anything useful, as I didn't find any mention of that configuration anywhere.

#23 - 07/02/2019 08:19 PM - Lucas Di Pentima

A couple more comments:

- The Cluster config scratch pad gdocs file (https://docs.google.com/document/d/13WO6fOWJpvCnPIUme_dY5Oudp06-BxM8g7BIaWzyqeg/edit) talks about TrustAllContent being below Collections for both keep-web & workbench. Is that still true? I ask because I'm also using it for the keep-web story.
- Commented debugging log on apps/workbench/app/views/jobs/_show_log.html.erb:83

#24 - 07/02/2019 08:40 PM - Peter Amstutz

Lucas Di Pentima wrote:

> Made a quick grep of configuration usage and found:
>
> - There's a mention of profiling_enabled config on apps/workbench/app/models/arvados_api_client.rb:64 that seems to not be migrated

Fixed.

> - There's a mention of api_response_compression on apps/workbench/app/models/arvados_api_client.rb:104 that it is already migrated to Workbench.APIResponseCompression

Fixed.

- There's a mention of testing_override_login_url on apps/workbench/app/models/arvados_api_client.rb lines 239 and 240 but it seems is for testing purposes only, so I'm not sure if it belongs to the cluster config. Mentioning it just in case.

I added that because arvados_login_base is always derived from Services.Controller.ExternalURL (as /login on controller) so it didn't make sense to have two configuration keys, but there was one test which relied on being able to override the login separately to provide a stub server.

- There's a mention of enable_public_projects_page on apps/workbench/app/views/application/_breadcrumbs.html.erb:34 that's already migrated to Workbench.EnablePublicProjectsPage

Fixed.

- Side note: there's a mention of Rails.configuration.include_accept_encoding_header_in_api_requests on apps/workbench/test/controllers/application_controller_test.rb:452 … it seems to me that this test doesn't do anything useful, as I didn't find any mention of that configuration anywhere.

This test was added in [#5556](#) and then the configuration key was changed to "api_response_compression" in #6087 but the test was never updated. I updated the test to use APIResponseCompression.

#25 - 07/03/2019 01:30 PM - Peter Amstutz

Lucas Di Pentima wrote:

> A couple more comments:
>
> - The Cluster config scratch pad gdocs file (https://docs.google.com/document/d/13WO6fOWJpvCnPIUme_dY5Oudp06-BxM8g7BIaWzyqeg/edit) talks about TrustAllContent being below Collections for both keep-web & workbench. Is that still true? I ask because I'm also using it for the keep-web story.

I moved it to Collections, thanks.

> - Commented debugging log on apps/workbench/app/views/jobs/_show_log.html.erb:83

Fixed.

https://ci.curoverse.com/view/Developer/job/developer-run-tests/1361/

#26 - 07/03/2019 01:33 PM - Peter Amstutz

- Target version changed from 2019-07-03 Sprint to 2019-07-17 Sprint

#27 - 07/03/2019 04:15 PM - Lucas Di Pentima

Just in case you didn't see it yet, tests failed with the following message at the workbench install phase: Collections.TrustAllContent expected Boolean but was NilClass

#28 - 07/03/2019 07:05 PM - Peter Amstutz

14812-wb1-new-config @ [fe06b864a537f5b6440a50bf85af9bc93d7aae80](#)

Fixed issue identified in review, fixed some other test issues.

https://ci.curoverse.com/view/Developer/job/developer-run-tests/1368/

#29 - 07/03/2019 07:42 PM - Lucas Di Pentima

Tried to start arvbox and I'm getting this on the workbench's log:

```
...
2019-07-03_19:41:11.16087 + bundle exec rake npm:install
2019-07-03_19:41:12.95978 Called 'load' without the :safe option -- defaulting to safe mode.
2019-07-03_19:41:12.95980 You can avoid this warning in the future by setting the SafeYAML::OPTIONS[:default_m
ode] option (to :safe or :unsafe).
2019-07-03_19:41:12.97944 time="2019-07-03T19:41:12.979368150Z" level=warning msg="deprecated or unknown confi
g entry: Clusters.x7k6z.NodeProfiles"
2019-07-03_19:41:12.97946 time="2019-07-03T19:41:12.979436278Z" level=warning msg="deprecated or unknown confi
g entry: Clusters.x7k6z.PostgreSQL.Connection.client_encoding"
2019-07-03_19:41:12.99627 rake aborted!
2019-07-03_19:41:12.99630 NoMethodError: private method `warn' called for nil:NilClass
```

```
2019-07-03_19:41:12.99639 /usr/src/arvados/apps/workbench/lib/config_validators.rb:11:in `validate_wb2_url_con
fig'
2019-07-03_19:41:12.99640 /usr/src/arvados/apps/workbench/config/arvados_config.rb:191:in `block in <top (requ
ired)>'
2019-07-03_19:41:12.99640 /var/lib/gems/ruby/2.3.0/gems/railties-5.0.7.2/lib/rails/railtie.rb:209:in `instance
_eval'
2019-07-03_19:41:12.99641 /var/lib/gems/ruby/2.3.0/gems/railties-5.0.7.2/lib/rails/railtie.rb:209:in `configur
e'
2019-07-03_19:41:12.99641 /var/lib/gems/ruby/2.3.0/gems/railties-5.0.7.2/lib/rails/railtie.rb:181:in `configur
e'
2019-07-03_19:41:12.99641 /usr/src/arvados/apps/workbench/config/arvados_config.rb:183:in `<top (required)>'
2019-07-03_19:41:12.99641 /usr/src/arvados/apps/workbench/config/application.rb:25:in `require_relative'
2019-07-03_19:41:12.99642 /usr/src/arvados/apps/workbench/config/application.rb:25:in `<class:Application>'
2019-07-03_19:41:12.99642 /usr/src/arvados/apps/workbench/config/application.rb:23:in `<module:ArvadosWorkbenc
h>'
2019-07-03_19:41:12.99642 /usr/src/arvados/apps/workbench/config/application.rb:22:in `<top (required)>'
2019-07-03_19:41:12.99642 /usr/src/arvados/apps/workbench/Rakefile:9:in `require'
2019-07-03_19:41:12.99642 /usr/src/arvados/apps/workbench/Rakefile:9:in `<top (required)>'
2019-07-03_19:41:12.99643 /var/lib/gems/ruby/2.3.0/gems/rake-12.3.2/exe/rake:27:in `<top (required)>'
...
```

#30 - 07/03/2019 08:42 PM - Lucas Di Pentima

Reviewing updates at [c12b6b6](#) now produces the following error:

```
...
2019-07-03_20:39:50.82684 rake aborted!
2019-07-03_20:39:50.82686 workbench2_url config is not an HTTP URL: wb2.example.org
2019-07-03_20:39:50.82687 /usr/src/arvados/apps/workbench/lib/config_validators.rb:11:in `validate_wb2_url_con
fig'
...
```

#31 - 07/08/2019 01:29 PM - Peter Amstutz

Lucas Di Pentima wrote:

> Reviewing updates at [c12b6b6](#) now produces the following error:
>
> [...]

It is supposed to do that. I think you wrote those validators originally?

#32 - 07/08/2019 01:57 PM - Peter Amstutz

14812-wb1-new-config @ [50b4bc6acb989af8902f329df284e30590b35ab1](#)

https://ci.curoverse.com/view/Developer/job/developer-run-tests/1371/

#33 - 07/08/2019 02:35 PM - Lucas Di Pentima

Yes, that error came from a custom application.yml file on apps/workbench/ that I use for testing purposes that got copied to the arvbox directory.
Pulled latest changes and deleted that file and tried again. Now I'm getting this error:

```
...
2019-07-08_14:15:20.71583 + RAILS_GROUPS=assets
2019-07-08_14:15:20.71584 + bundle exec rake npm:install
2019-07-08_14:15:22.55011 Called 'load' without the :safe option -- defaulting to safe mode.
2019-07-08_14:15:22.55013 You can avoid this warning in the future by setting the SafeYAML::OPTIONS[:default_m
ode] option (to :safe or :unsafe).
2019-07-08_14:15:22.56996 time="2019-07-08T14:15:22.569857621Z" level=warning msg="deprecated or unknown confi
g entry: Clusters.x7k6z.NodeProfiles"
2019-07-08_14:15:22.56998 time="2019-07-08T14:15:22.569916267Z" level=warning msg="deprecated or unknown confi
g entry: Clusters.x7k6z.PostgreSQL.Connection.client_encoding"
2019-07-08_14:15:22.58376 rake aborted!
2019-07-08_14:15:22.58377 ArgumentError: bad argument (expected URI object or URI string)
2019-07-08_14:15:22.58392 /usr/src/arvados/apps/workbench/config/arvados_config.rb:60:in `block in <top (requi
red)>'
2019-07-08_14:15:22.58393 /usr/src/arvados/apps/workbench/lib/config_loader.rb:71:in `block in migrate_config'
2019-07-08_14:15:22.58393 /usr/src/arvados/apps/workbench/lib/config_loader.rb:69:in `each'
2019-07-08_14:15:22.58393 /usr/src/arvados/apps/workbench/lib/config_loader.rb:69:in `migrate_config'
2019-07-08_14:15:22.58394 /usr/src/arvados/apps/workbench/config/arvados_config.rb:167:in `<top (required)>'
2019-07-08_14:15:22.58394 /usr/src/arvados/apps/workbench/config/application.rb:25:in `require_relative'
...
```

If you already tried this branch on arvbox, I don't want to keep stalling this merge because of local problems. My arvbox test attempt is being done on a freshly new instance (on its own newly created ~/.arvbox/14812/ directory), without any override config file.

#34 - 07/08/2019 03:28 PM - Lucas Di Pentima

Trying arvbox with 3100cf66fee222eed832bd813cc0b7430bed0f93

Getting this error:

```
...
2019-07-08_15:26:16.06279 + RAILS_GROUPS=assets
2019-07-08_15:26:16.06281 + bundle exec rake npm:install
2019-07-08_15:26:29.46300 Called 'load' without the :safe option -- defaulting to safe mode.
2019-07-08_15:26:29.46302 You can avoid this warning in the future by setting the SafeYAML::OPTIONS[:default_m
ode] option (to :safe or :unsafe).
2019-07-08_15:26:29.58167 time="2019-07-08T15:26:29.581537970Z" level=warning msg="deprecated or unknown confi
g entry: Clusters.x7k6z.NodeProfiles"
2019-07-08_15:26:29.58169 time="2019-07-08T15:26:29.581652890Z" level=warning msg="deprecated or unknown confi
g entry: Clusters.x7k6z.PostgreSQL.Connection.client_encoding"
2019-07-08_15:26:29.65491 rake aborted!
2019-07-08_15:26:29.65565 NoMethodError: undefined method `keep_web_url' for #<Rails::Application::Configurati
on:0x000055f3f81cc180>
2019-07-08_15:26:29.65567 /var/lib/gems/ruby/2.3.0/gems/railties-5.0.7.2/lib/rails/railtie/configuration.rb:95
:in `method_missing'
2019-07-08_15:26:29.65568 /usr/src/arvados/apps/workbench/config/load_config.rb:61:in `block in <top (required
)>'
2019-07-08_15:26:29.65568 /var/lib/gems/ruby/2.3.0/gems/railties-5.0.7.2/lib/rails/railtie.rb:209:in `instance
_eval'
2019-07-08_15:26:29.65568 /var/lib/gems/ruby/2.3.0/gems/railties-5.0.7.2/lib/rails/railtie.rb:209:in `configur
e'
2019-07-08_15:26:29.65568 /var/lib/gems/ruby/2.3.0/gems/railties-5.0.7.2/lib/rails/railtie.rb:181:in `configur
e'
2019-07-08_15:26:29.65568 /usr/src/arvados/apps/workbench/config/load_config.rb:20:in `<top (required)>'
...
```

#35 - 07/08/2019 06:28 PM - Lucas Di Pentima

Tried starting arvbox with 1a021e2c70aafe29222a2970ec97347f4f55e340 and it worked great!

What I think may be missing is to update the documentation, the workbench install guide and some upgrade instructions on the release notes, what do you think? (I think I forgot to check for this on the API server story)

#36 - 07/08/2019 08:14 PM - Peter Amstutz

Lucas Di Pentima wrote:

> Tried starting arvbox with 1a021e2c70aafe29222a2970ec97347f4f55e340 and it worked great!
>
> What I think may be missing is to update the documentation, the workbench install guide and some upgrade instructions on the release notes, what do you think? (I think I forgot to check for this on the API server story)

Updated documentation 14363bc8381266de4ad0b8db7f32f0859adaaece I updated the API server documentation as well because I forgot to do that previously.

https://ci.curoverse.com/view/Developer/job/developer-run-tests/1374/

#37 - 07/08/2019 09:25 PM - Lucas Di Pentima

Some minor comments on the documentation updates:

- Install API section:
  - Under "Set up the database" the documentation mentions the database.yml file and how to configure it.
  - Under "Collections.BlobSigningKey" header, there's still a mention of the application.yml file: "...Generate a random value and set it in application.yml."
- Install Workbench section:
  - Under "Configure Piwik" header, there's a missing word on "Piwik is optional, and can be ??? to gather usage…"
- Migrating Configuration section:
  - Under "Workbench" the database.yml file is mentioned but it really isn't used by workbench.

With that, it LGTM. Thanks!

#38 - 07/09/2019 10:15 PM - Ward Vandewege

- Related to Support #15449: Workbench and API server packages depend on arvados-server added

#39 - 07/16/2019 08:08 PM - Peter Amstutz

- Status changed from In Progress to Resolved

#40 - 01/22/2020 02:13 PM - Peter Amstutz

- Release set to 22