Arvados - Story #15107

[controller] Implement native Google login (configurable as an alternative to sso-provider)

04/17/2019 05:03 PM - Tom Clegg

Status:	Resolved	**Start date:**	10/31/2019
Priority:	Normal	**Due date:**	
Assigned To:	Tom Clegg	**% Done:**	100%
Category:	API	**Estimated time:**	0.00 hour
Target version:	2019-11-20 Sprint		

Description

See Native login implementation

Implement an OpenID Connect login mechanism that supports (at least) Google login.

Cluster configuration should make it possible to

- continue using sso-provider as before (default), or
- use the new OpenID Connect mechanism to sign in with Google.

This issue does **not** include:

- Offering the user a backend chooser
- Supporting both sso-provider and OpenID Connect at the same time
- Supporting multiple backends at the same time
- LDAP

Subtasks:

Task # 15512: Review 15107-google-login	**Resolved**
Task # 15825: Review 15107-rails-bad-redirect	**Resolved**
Task # 15830: Review 15107-alt-email	**Resolved**
Task # 15833: Review 15107-prefer-domain-for-username	**Resolved**

Related issues:

Related to Arvados - Story #15477: Use email address for Arvados account linking	**Duplicate**		
Related to Arvados - Story #15795: [API] Accept configured SystemRootToken wi...	**Resolved**	**11/23/2019**	
Related to Arvados - Bug #15867: LoginCluster redirect broken with EnableBeta...	**Resolved**		
Blocks Arvados Epics - Story #15322: Replace and delete sso-provider	**Resolved**	**03/11/2020**	**08/26/2020**

Associated revisions

Revision b30dca66 - 11/06/2019 06:59 PM - Tom Clegg

Merge branch '15107-google-login'

refs #15107

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tclegg@veritasgenetics.com>

Revision d97c9ecc - 11/14/2019 11:27 PM - Tom Clegg

Merge branch '15107-alt-email'

refs #15107

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tclegg@veritasgenetics.com>

Revision fa8e7a73 - 11/15/2019 03:57 AM - Tom Clegg

Merge branch '15107-rails-bad-redirect'

refs #15107

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tclegg@veritasgenetics.com>

Revision bdc8a763 - 11/18/2019 09:36 PM - Tom Clegg

Merge branch '15107-prefer-domain-for-username'

closes #15107

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tclegg@veritasgenetics.com>

History

#1 - 04/17/2019 08:08 PM - Tom Clegg

- Subject changed from [controller] Implement native login (to replace sso-provider) to [controller] Implement native Google login (configurable as an alternative to sso-provider)

- Description updated

- Category set to API

- Story points set to 3.0

#2 - 04/24/2019 04:41 PM - Tom Morris

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

#3 - 06/05/2019 03:27 PM - Tom Clegg

- Blocks Story #15322: Replace and delete sso-provider added

#4 - 06/05/2019 03:41 PM - Tom Clegg

- Assigned To set to Tom Clegg

- Target version changed from Arvados Future Sprints to 2019-06-19 Sprint

#5 - 06/05/2019 03:44 PM - Tom Clegg

- Assigned To deleted (Tom Clegg)

- Target version changed from 2019-06-19 Sprint to Arvados Future Sprints

#6 - 07/31/2019 03:35 PM - Tom Morris

- Assigned To set to Tom Clegg

- Target version changed from Arvados Future Sprints to 2019-08-14 Sprint

#7 - 07/31/2019 03:37 PM - Tom Clegg

- Related to Story #15477: Use email address for Arvados account linking added

#8 - 08/14/2019 03:16 PM - Tom Clegg

- Target version changed from 2019-08-14 Sprint to 2019-08-28 Sprint

#9 - 08/28/2019 02:41 PM - Tom Morris

- Target version changed from 2019-08-28 Sprint to 2019-09-11 Sprint

#10 - 09/11/2019 02:45 PM - Tom Clegg

- Target version changed from 2019-09-11 Sprint to 2019-09-25 Sprint

#11 - 09/25/2019 03:15 PM - Tom Clegg

- Target version changed from 2019-09-25 Sprint to 2019-10-09 Sprint

#12 - 10/01/2019 01:33 PM - Tom Clegg

- Status changed from New to In Progress

#13 - 10/09/2019 03:08 PM - Tom Clegg

- Target version changed from 2019-10-09 Sprint to 2019-10-23 Sprint

#14 - 10/23/2019 02:46 PM - Tom Clegg

- Target version changed from 2019-10-23 Sprint to 2019-11-06 Sprint

#16 - 10/31/2019 01:41 PM - Tom Clegg

15107-google-login @ deaf1d8f2f694b09562eddac055ccebba5a98517 -- https://ci.curoverse.com/view/Developer/job/developer-run-tests/1622/

(also tested on 4xphq using real Google credentials)

#17 - 11/01/2019 07:28 PM - Peter Amstutz

Getting this working locally (I've done it successfully with SSO so I know its possible).

First thing I noticed, ProviderAppID can't be empty but it also conflicts with GoogleClientID:

```
2019-11-01_18:33:40.56022 Login.ProviderAppID cannot be empty
2019-11-01_18:33:40.56045 /usr/src/arvados/services/api/lib/config_loader.rb:118:in `block in coercion_and_che
ck'
2019-11-01_18:33:40.56046 /usr/src/arvados/services/api/lib/config_loader.rb:80:in `each'
2019-11-01_18:33:40.56046 /usr/src/arvados/services/api/lib/config_loader.rb:80:in `coercion_and_check'
2019-11-01_18:33:40.56047 /usr/src/arvados/services/api/config/arvados_config.rb:232:in `<top (required)>'
```

(To work around it I made ProviderAppID non-essential).

Now in the callback phase I'm getting this error:

```
{"errors":["request failed: http://localhost:8004/auth/controller/callback: 401 Unauthorized: Invalid authoriz
ation header (req-1bz95m2kax6xz19va72o)"]}
```

After some poking I modified the error code to get this:

```
{"errors":["request failed: http://localhost:8004/auth/controller/callback: 401 Unauthorized: Invalid authoriz
ation header got 'Bearer' expected 'Bearer ' (req-ov89tfhbfya42uwb317e)"]}
```

It turns out arvbox isn't setting SystemRootToken. Whoops. That should be validated. We can check this in config/arvados_config.rb, but probably this belongs in the validation on the Go side.

```
arvcfg.declare_config "SystemRootToken", NonemptyString
```

In order to have feature parity with SSO Google login we need to support alternate emails and the customer-requested "username domain" feature. The Ruby code starts at https://dev.arvados.org/projects/arvados/repository/sso-provider/revisions/master/entry/app/controllers/users/omniauth_callbacks_controller.rb#L20

1. Needs the scope "user.emails.read" (might actually be called "https://www.googleapis.com/auth/user.emails.read")
2. Need to get the "me" record from the google people service (https://godoc.org/google.golang.org/api/people/v1) using the google token we received from the login callback
3. Go through the email addresses in the response to determine the primary & alternate email addresses for the account
4. "first_name" and "last_name" are also missing from auth_info, my new user is called "null null"
5. Support an optional "username domain", the email address that has that domain will be used for the preferred username

Still need to manually test that LoginCluster redirection works.

#18 - 11/05/2019 04:03 PM - Tom Clegg

15107-google-login @ ae562784e8d8d8bd501c0bd373739d0a2da8fc9f -- https://ci.curoverse.com/view/Developer/job/developer-run-tests/1623/

- RailsAPI doesn't error out if ProviderAppID/Secret are empty
- First/last name propagated to RailsAPI
- Test for propagation of user info to RailsAPI

Can add "alternate email addrs" and "preferred domain for choosing username" in a subsequent branch.

Making SystemRootToken mandatory sounds like an improvement. To avoid worsening the unforgiving sequence of install steps (and save a few db queries) we'll probably also want the RailsAPI auth middleware to recognize it by looking at config, instead of requiring the installer to create an api_client_authorizations row and then copy the token to config.

#19 - 11/05/2019 06:26 PM - Peter Amstutz

> Can add "alternate email addrs" and "preferred domain for choosing username" in a subsequent branch.

Same ticket or new ticket?

> Making SystemRootToken mandatory sounds like an improvement. To avoid worsening the unforgiving sequence of install steps (and save a few db queries) we'll probably also want the RailsAPI auth middleware to recognize it by looking at config, instead of requiring the installer to create an api_client_authorizations row and then copy the token to config.

Also agree. Same ticket or new ticket?

#20 - 11/05/2019 06:45 PM - Tom Clegg

Peter Amstutz wrote:

> Can add "alternate email addrs" and "preferred domain for choosing username" in a subsequent branch.

Same ticket or new ticket?

Same, this seems like part of implementing Google login.

> Making SystemRootToken mandatory

Also agree. Same ticket or new ticket?

Added #15795

#21 - 11/05/2019 06:54 PM - Tom Clegg

- Related to Story #15795: [API] Accept configured SystemRootToken without doing a database lookup added

#22 - 11/05/2019 09:40 PM - Peter Amstutz

```
arvcfg.declare_config "SystemRootToken", String, :SystemRootToken
```

- The 3rd argument is the corresponding key to import from legacy application.yml, but I don't think that ever existed?
- Instead of 'String' it can be 'NonemptyString' to prevent the API server from starting if it is empty, the description in #15795 doesn't say anything about requiring SystemRootToken have a valid value for services to start

Are you sure the "claims" struct we get from Google doesn't already have first_name and last_name separated?

... ran out of time will look at it some more tomorrow

#23 - 11/05/2019 10:02 PM - Tom Clegg

Peter Amstutz wrote:

> - The 3rd argument is the corresponding key to import from legacy application.yml, but I don't think that ever existed?

Removed.

> - Instead of 'String' it can be 'NonemptyString' to prevent the API server from starting if it is empty, the description in #15795 doesn't say anything about requiring SystemRootToken have a valid value for services to start

OK, noted on #15795. But if we made it mandatory now, wouldn't installation require you to add a bogus token, then run the "generate valid root token" rake task now that the config is valid, then replace the bogus token with the real one? (This is what I meant by "worsening the unforgiving sequence of install steps" above.)

> Are you sure the "claims" struct we get from Google doesn't already have first_name and last_name separated?

I guess not, but I don't see them at https://developers.google.com/identity/protocols/OpenIDConnect

#24 - 11/06/2019 04:13 PM - Tom Clegg

- Target version changed from 2019-11-06 Sprint to 2019-11-20 Sprint

#25 - 11/06/2019 05:04 PM - Peter Amstutz

Also needs to be added to documentation as a "beta" feature

#26 - 11/06/2019 06:27 PM - Peter Amstutz

Tom Clegg wrote:

> Are you sure the "claims" struct we get from Google doesn't already have first_name and last_name separated?

> I guess not, but I don't see them at https://developers.google.com/identity/protocols/OpenIDConnect

Ok, I looked at what Omniauth does, it makes a callback to the "https://www.googleapis.com/oauth2/v3/userinfo" endpoint to get a record that includes the separated given / family name. The "people/me" endpoint should also provide this information, so we can just switch to using that when we do that branch (next).

ae562784e8d8d8bd501c0bd373739d0a2da8fc9f LGTM

#27 - 11/06/2019 06:53 PM - Peter Amstutz

Ah, I never did a manual test that LoginCluster works. Note to self: do more manual testing when reviewing the next branch.

#29 - 11/10/2019 06:54 AM - Tom Clegg

15107-alt-email @ e0b0048703b78b6a6846e4b5e0f71ec4b1553aa0 -- https://ci.curoverse.com/view/Developer/job/developer-run-tests/1640/

#31 - 11/12/2019 06:44 PM - Peter Amstutz

When EnableBetaController14287 is turned on, but I'm not using the Google login, when visiting the '/login' route, the redirect to "/auth/joshid" seems to be rewritten by controller to the Rails server internal URL.

With EnableBetaController14287 turned off, the redirect to "/auth/joshid" (and from there to the SSO server) works as expected.

#32 - 11/12/2019 09:00 PM - Tom Clegg

If the bad redirect target was https://internal_rails_host:port (as opposed to http://internal_rails_host:port) then this should fix it. It turns out the "X-Forwarded-Proto: https" header -- in addition to reassuring Rails that it doesn't need to redirect plain requests to https -- also caused it to rewrite relative redirect targets as the bogus "https://host:port/target" (instead of "http://host:port/target"), and controller just passed it through as is instead of rewriting it, because it wasn't same-origin.

So it seems force_ssl really does force ssl (by either redirecting all reqs to bogus URLs or mangling relative redirect_to targets)... surprisingly enough.

~~15107-rails-bad-redirect @ 4f938f3a77ef2629d934dec56e25a314c682b6aa~~https://ci.curoverse.com/view/Developer/job/developer-run-tests/1644/

15107-rails-bad-redirect @ c00d9e1595d07e6941bb2fbfb8b4e57c3c4ba856 -- https://ci.curoverse.com/view/Developer/job/developer-run-tests/1645/

- stop sending the X-Forwarded-Proto header (don't mangle redirect_to targets)
- ~~turn off force_ssl (allow http reqs without an X-Forwarded-Proto header)~~
- leave force_ssl on (default) but disable the redirect-all-requests-to-https feature

#33 - 11/14/2019 03:02 PM - Tom Clegg

15107-alt-email @ ca6544470298ca1586b7de5ead8c5ff4894443fe -- https://ci.curoverse.com/view/Developer/job/developer-run-tests/1646/

- Retrieve additional email addresses and passes the verified ones through to RailsAPI
- ...but allow the admin to disable this via config (in case the People API can't be enabled immediately and they would rather sacrifice the feature for now than get stuck on it)
- Mention any ignored (non-verified) email addresses in logs, to help troubleshooting
- If the People API returns a "primary" name, use it (with the provided first/last split) instead of splitting the OIDC full name on whitespace

#35 - 11/14/2019 06:47 PM - Peter Amstutz

Tom Clegg wrote:

> If the bad redirect target was https://internal_rails_host:port (as opposed to http://internal_rails_host:port) then this should fix it. It turns out the "X-Forwarded-Proto: https" header -- in addition to reassuring Rails that it doesn't need to redirect plain requests to https -- also caused it to rewrite relative redirect targets as the bogus "https://host:port/target" (instead of "http://host:port/target"), and controller just passed it through as is instead of rewriting it, because it wasn't same-origin.
>
> So it seems force_ssl really does force ssl (by either redirecting all reqs to bogus URLs or mangling relative redirect_to targets)... surprisingly enough.
>
> ~~15107-rails-bad-redirect @ 4f938f3a77ef2629d934dec56e25a314c682b6aa~~
> https://ci.curoverse.com/view/Developer/job/developer-run-tests/1644/
>
> 15107-rails-bad-redirect @ c00d9e1595d07e6941bb2fbfb8b4e57c3c4ba856 --
> https://ci.curoverse.com/view/Developer/job/developer-run-tests/1645/
>
> - stop sending the X-Forwarded-Proto header (don't mangle redirect_to targets)
> - ~~turn off force_ssl (allow http reqs without an X-Forwarded-Proto header)~~
> - leave force_ssl on (default) but disable the redirect-all-requests-to-https feature

Tested this and it works for me, LGTM.

#36 - 11/14/2019 08:11 PM - Peter Amstutz

After banging my head against this for a while, I finally figured out that "go get ..." fetches the latest of everything, the correct command to use with modules is "go mod download".

I've pushed commit 3b9af4b0f to 15107-alt-email that fixes arvbox to use "go mod download".

#37 - 11/14/2019 08:36 PM - Peter Amstutz

Finally able to review the branch.

This is missing the feature of designating a specific email domain who's username will be used as the preferred arvados username. This is a customer-requested feature.

#38 - 11/15/2019 05:23 PM - Tom Clegg

15107-prefer-domain-for-username @ 943827578884b09a155443a9d2bb685a327070f9 -- https://ci.curoverse.com/view/Developer/job/developer-run-tests/1649/

- adds Users.PreferDomainForUsername config entry

#39 - 11/18/2019 09:30 PM - Peter Amstutz

Tom Clegg wrote:

> 15107-prefer-domain-for-username @ 943827578884b09a155443a9d2bb685a327070f9 -- https://ci.curoverse.com/view/Developer/job/developer-run-tests/1649/
>
> - adds Users.PreferDomainForUsername config entry

The test case didn't properly verify that was getting the username from preferdomainforusername.example.com, and also didn't check the "+" section would be handled properly. I went ahead and just tweaked the tests and pushed that commit and the rest LGTM:

542a72e8ea402a65d75a5251ba219341834fb2c9

#40 - 11/19/2019 07:24 AM - Tom Clegg

- Status changed from In Progress to Resolved

Applied in changeset arvados|bdc8a7630030494c63fb0426be4c15a93a9a37cb.

#41 - 11/22/2019 04:00 PM - Peter Amstutz

- Related to Bug #15867: LoginCluster redirect broken with EnableBetaController: true added

#42 - 01/21/2020 09:22 PM - Peter Amstutz

- Release set to 22