

## Arvados - Bug #5834

### [API] Be careful about memory usage when responding to collections#index queries

04/28/2015 03:32 PM - Ward Vandewege

<b>Status:</b>	Resolved	<b>Start date:</b>	04/30/2015
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assigned To:</b>	Brett Smith	<b>% Done:</b>	100%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	2015-05-20 sprint		

#### Description

API server logs:

```
#<NoMemoryError: failed to allocate memory>
/data/www/qr1hi.arvadosapi.com/releases/20150423194138/app/controllers/application_controller.rb:1
89:in `dump'
/data/www/qr1hi.arvadosapi.com/releases/20150423194138/app/controllers/application_controller.rb:1
89:in `send_json'
/data/www/qr1hi.arvadosapi.com/releases/20150423194138/app/controllers/application_controller.rb:4
58:in `render_list'
/data/www/qr1hi.arvadosapi.com/releases/20150423194138/app/controllers/application_controller.rb:7
6:in `index'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_co
ntroller/metal/implicit_render.rb:4:in `send_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/abstract_
controller/base.rb:167:in `process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_co
ntroller/metal/rendering.rb:10:in `process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/abstract_
controller/callbacks.rb:18:in `block in process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active
_support/callbacks.rb:579:in `_run__2616472115270357727__process_action__4453107480639721567__call
backs'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active
_support/callbacks.rb:405:in `__run_callback'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active
_support/callbacks.rb:385:in `_run_process_action_callbacks'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active
_support/callbacks.rb:81:in `run_callbacks'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/abstract_
controller/callbacks.rb:17:in `process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_co
ntroller/metal/rescue.rb:29:in `process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_co
ntroller/metal/instrumentation.rb:30:in `block in process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active
_support/notifications.rb:123:in `block in instrument'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active
_support/notifications/instrumenter.rb:20:in `instrument'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active
_support/notifications.rb:123:in `instrument'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_co
ntroller/metal/instrumentation.rb:29:in `process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_co
ntroller/metal/params_wrapper.rb:207:in `process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active_
record/railties/controller_runtime.rb:18:in `process_action'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/abstract_
controller/base.rb:121:in `process'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/abstract_
controller/rendering.rb:45:in `process'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_co
ntroller/metal.rb:203:in `dispatch'
```

```
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_controller/metal/rack_delegation.rb:14:in `dispatch'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_controller/metal.rb:246:in `block in action'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/routing/route_set.rb:73:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/routing/route_set.rb:73:in `dispatch'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/routing/route_set.rb:36:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/journey-1.0.4/lib/journey/router.rb:68:in `block in call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/journey-1.0.4/lib/journey/router.rb:56:in `each'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/journey-1.0.4/lib/journey/router.rb:56:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/routing/route_set.rb:608:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/omniauth-1.1.1/lib/omniauth/strategy.rb:177:in `call!'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/omniauth-1.1.1/lib/omniauth/strategy.rb:157:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/omniauth-1.1.1/lib/omniauth/builder.rb:48:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/best_standards_support.rb:17:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-1.4.5/lib/rack/etag.rb:23:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-1.4.5/lib/rack/conditional_get.rb:25:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/head.rb:14:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/params_parser.rb:21:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/flash.rb:242:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-1.4.5/lib/rack/session/abstract/id.rb:210:in `context'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-1.4.5/lib/rack/session/abstract/id.rb:205:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/cookies.rb:341:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active_record/query_cache.rb:64:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activerecord-3.2.17/lib/active_record/connection_adapters/abstract/connection_pool.rb:479:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/callbacks.rb:28:in `block in call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activesupport-3.2.17/lib/active_support/callbacks.rb:405:in `__run__1960694728971137051__call__1126081408567888176__callbacks'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activesupport-3.2.17/lib/active_support/callbacks.rb:405:in `__run_callback'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activesupport-3.2.17/lib/active_support/callbacks.rb:385:in `_run_call_callbacks'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activesupport-3.2.17/lib/active_support/callbacks.rb:81:in `run_callbacks'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/callbacks.rb:27:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/debug_exceptions.rb:16:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_dispatch/middleware/show_exceptions.rb:56:in `call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/railties-3.2.17/lib/rails/rack/logger.rb:32:in `call_app'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/railties-3.2.17/lib/rails/rack/logger.rb:16:in `block in call'  
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activesupport-3.2.17/lib/active
```

```

_support/tagged_logging.rb:22:in `tagged'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/railties-3.2.17/lib/rails/rack/
logger.rb:16:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_di
spatch/middleware/request_id.rb:22:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-1.4.5/lib/rack/methodoverr
ide.rb:21:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-1.4.5/lib/rack/runtime.rb:
17:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/activesupport-3.2.17/lib/active
_support/cache/strategy/local_cache.rb:72:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-1.4.5/lib/rack/lock.rb:15:
in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-cache-1.2/lib/rack/cache/c
ontext.rb:136:in `forward'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-cache-1.2/lib/rack/cache/c
ontext.rb:245:in `fetch'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-cache-1.2/lib/rack/cache/c
ontext.rb:185:in `lookup'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-cache-1.2/lib/rack/cache/c
ontext.rb:66:in `call!'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/rack-cache-1.2/lib/rack/cache/c
ontext.rb:51:in `call'
/data/www/qr1hi.arvadosapi.com/releases/20150423194138/app/middlewares/arvados_api_token.rb:59:in
`call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/actionpack-3.2.17/lib/action_di
spatch/middleware/remote_ip.rb:31:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/railties-3.2.17/lib/rails/engin
e.rb:484:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/railties-3.2.17/lib/rails/appli
cation.rb:231:in `call'
/var/www/qr1hi.arvadosapi.com/shared/vendor_bundle/ruby/2.1.0/gems/railties-3.2.17/lib/rails/railt
ie/configurable.rb:30:in `method_missing'
/usr/lib/ruby/vendor_ruby/phusion_passenger/rack/thread_handler_extension.rb:74:in `process_reques
t'
/usr/lib/ruby/vendor_ruby/phusion_passenger/request_handler/thread_handler.rb:141:in `accept_and_p
rocess_next_request'
/usr/lib/ruby/vendor_ruby/phusion_passenger/request_handler/thread_handler.rb:109:in `main_loop'
/usr/lib/ruby/vendor_ruby/phusion_passenger/request_handler.rb:455:in `block (3 levels) in start_t
hreads'

```

Error 1430180085+f49f1f8f: 422

Rendered text template (0.0ms)

Completed 422 Unprocessable Entity in 9907.9ms (Views: 0.5ms | ActiveRecord: 5932.8ms)

#### Subtasks:

Task # 5871: Review 5834-api-max-response-size-wip

**Resolved**

#### Related issues:

Related to Arvados - Bug #6074: [API] collections are being returned too slow...

**Resolved**

**06/03/2015**

Blocks Arvados - Story #3408: [Keep] Implement Production Data Manager

**Resolved**

**07/29/2014**

#### Associated revisions

**Revision e8f91c87 - 05/01/2015 08:49 PM - Brett Smith**

Merge branch '5834-api-max-response-size-wip'

Closes #5834, #5871.

#### History

**#1 - 04/28/2015 03:35 PM - Ward Vandewege**

Consistently, after reading 25,149 collections with datamanager, when I query for the next set, the SDK returns: #<NoMemoryError: failed to allocate memory>

On the machine that runs datamanager, I'm logging the memory usage, and it never gets above 2.4 Gigs.

Datamanager was built with go1.4 linux/amd64

## #2 - 04/28/2015 08:07 PM - Tom Clegg

- Target version changed from Bug Triage to 2015-05-20 sprint

## #3 - 04/28/2015 08:51 PM - Ward Vandewege

- Description updated

Server

## #4 - 04/28/2015 09:10 PM - Tom Clegg

Manifests are allowed to be over 64 MB which means a page of 100 can easily be over 6.4 GB.

API server doesn't *guarantee* the client will get any particular page size, so this might be the way to resolve it:

- In `Arvados::V1::CollectionsController#find_objects_for_index`, before getting the collection records, do a "select length(manifest\_text)" on the requested set.
- Walk the results of the "length()" query, adding up the lengths. If the cumulative length of the first N collections exceeds some configured limit (say `Rails.configuration.max_response_size`), set `@limit` to `min(1,N-1)`.
- Call `super` to do the usual query.

`max_response_size` should be documented in `application.default.yml` with a note that it's not a hard limit: actual response sizes will sometimes be bigger because we only bother checking `manifest_text` (and only in `collections#index`) and because we return 1 result even if that result is much bigger than `max_response_size`.

Default can be 128 MiB (just like `max_request_size`).

## #5 - 04/29/2015 07:28 PM - Tom Clegg

- Subject changed from *[API] memory issues on the API server when data manager runs* to *[API] Be careful about memory usage when responding to collections#index queries*

- Story points set to 0.5

## #6 - 04/29/2015 07:40 PM - Brett Smith

- Assigned To set to Brett Smith

## #7 - 04/29/2015 09:27 PM - Misha Zatsman

That sounds like a great solution Tom. The data manager doesn't really care how many results it gets, as long as it gets enough to move the modified date forward from the last set.

## #8 - 04/30/2015 06:09 PM - Brett Smith

- Status changed from New to In Progress

## #9 - 04/30/2015 07:24 PM - Brett Smith

(03:06:25 PM) Me: tomclegg: Do you want any compensation for the fact that the `manifest_text` in the database will be smaller than the `manifest_text` that we send over the wire, when signatures get added?

(03:07:16 PM) tomclegg: brett: hm, good point, that can inflate ~2x when the files are big...

(03:09:21 PM) tomclegg: brett: how does "no, but it would be worthwhile to explain this in `application.default.yml`, and does this mean we should rename the config since it's now likely to be so far away from response size?" sound?

(03:10:11 PM) Me: tomclegg: I can work with that. Should we just give up all pretense of this being a generic configuration knob, and have the name reflect that it's specific to `/collections`?

(03:10:44 PM) Me: `collection_max_manifest_read` or something.

(03:13:53 PM) tomclegg: brett: I suppose the reason we're doing this is API memory use, not response size, so should this be a "try not to use much more than X memory" config?

(03:15:17 PM) Me: tomclegg: That seems equally hazy though. I don't have a good intuitive sense of what kind of scaling factor there is for database record size, to ActiveRecord object size, to serialized JSON size.

(03:15:40 PM) Me: And it seems like it could vary a lot based on factors we're unlikely to track, like Rails upgrades, Oj optimizations, and so on.

(03:15:44 PM) tomclegg: it seems to me when we find a similar problem, e.g., `logs#index`, we'll probably want to use the same limit there

(03:15:58 PM) Me: I agree with that.

(03:16:38 PM) Me: `max_database_index_read`?

(03:16:56 PM) Me: To more closely say "We'll track large column(s) from the database and only read this much."

(03:17:16 PM) Me: I feel relatively comfortable saying this is the strategy we're going to keep using.

(03:17:36 PM) tomclegg: that sgtm.

(03:18:01 PM) Me: Because selecting length(whatever) from the database seems to do the best job of letting you look before you leap.

(03:18:57 PM) tomclegg: it also makes it more obvious that it's an internal server thing and clients shouldn't expect to see how that number relates to their response sizes, admins shouldn't expect to see that number in "top", etc.

(03:19:04 PM) Me: Yup.

## #10 - 05/01/2015 08:18 PM - Peter Amstutz

[f87f91c](#)

per the spec:

"max\_response\_size should be documented in application.default.yml with a note that it's not a hard limit: actual response sizes will sometimes be bigger because we only bother checking manifest\_text (and only in collections#index) and because we return 1 result even if that result is much bigger than max\_response\_size."

Please improve the comment in application.default.yml to clarify the soft limit behavior and then merge. Thanks.

**#11 - 05/01/2015 08:55 PM - Brett Smith**

- *Status changed from In Progress to Resolved*

- *% Done changed from 0 to 100*

Applied in changeset arvados|commit:e8f91c87ef238ddf6bd96914a196d2d90825e8ee.