

Arvados - Story #14374

Multi-site object search

10/17/2018 04:34 PM - Tom Morris

Status: New	Start date:
Priority: Normal	Due date:
Assigned To:	% Done: 0%
Category:	Estimated time: 0.00 hour
Target version: Arvados Future Sprints	
Description Implemented in controller. <ol style="list-style-type: none">1. Distribute query to each cluster in the federation.2. Collect responses, collate by order and truncate to response item limit3. Construct a "page token" which consists of the original filters plus additional information needed to get remaining unseen items on each cluster4. Respond with object list and page token5. Subsequent requests should only use the page token Query parameters: <ul style="list-style-type: none">• enable_multisite_search<ul style="list-style-type: none">◦ true enables multisite search◦ true by default?◦ provide cluster_id= to search only an individual cluster?• return_partial_results<ul style="list-style-type: none">◦ when true, if a query to a remote cluster fails, return a response with that cluster stripped out and an additional "warning" field mentioning there was a problem◦ when false, if any query to a remote cluster fails, the whole request fails• multisite_search_timeout<ul style="list-style-type: none">◦ how long to wait for a result from one of the remote clusters Query considerations: <ul style="list-style-type: none">• If "select" is specified and "uuid" is missing, must add in uuid for remote queries (needed for paging) and strip it out in the response• Only order on scalar columns. Disambiguate rows with the same value by ordering by uuid.• Cannot have "offset"• Must be "count=none"• Must be "distinct=false"	
Related issues: Related to Arvados - Bug #14710: [Workbench] Child containers run on federate... New	

History

#1 - 10/24/2018 02:37 PM - Tom Morris

- Subject changed from Multi-site workflow search to Multi-site object search

- Description updated

#2 - 10/24/2018 03:37 PM - Peter Amstutz

[THIS IS NO LONGER RELEVANT]

The main challenges are

1. return responses as we get them (so one slow site doesn't slow down the whole response)
2. stateless paging

Items in the response can only be globally ordered by waiting for and merging responses from every site. We can restrict/redefine ordering behavior, such as: objects on the same cluster will be returned sequentially and ordered internally, the client is expected to re-sort.

If we write the response json object incrementally (as remote responses come back) we could use streaming (event driven) parsing in the browser with <http://oboejs.com/>

Note: writing incrementally probably means using chunked Transfer-Encoding. See also <https://github.com/jonnyreeves/chunked-request>

On paging.

Given the ordering restriction described above, the client could use a recipe like:

For each sequence of items on the same cluster, look for the last one (as defined by a cut-over in the list from one uuid prefix to another, or the end of the list). Use each "final" item to construct a query that fetches the next page of items.

Note: we might not know that we have the "last" item if the following response is delayed, so we are delayed in seeing the cut-over. Maybe "items" should be a list of lists, so we see the termination of the list instead? Or maybe "items" should be a map of cluster → items.

Maybe a federated response special form:

```
{
  "kind": "arvados#federatedList",
  "parts": {
    "qrlhi": {
      "items": [...]
    },
    "4xphq": {
      "items": [...]
    }
  }
}
```

With corresponding federated query?

```
filters: [
  ["any", "@@", "my query string"],
  {
    "parts": {
      "qrlhi": [{"uuid", ">", "123"}],
      "4xphq": [{"uuid", ">", "456"}]
    }
  }
]
```

#3 - 10/24/2018 05:57 PM - Peter Amstutz

- Description updated

#4 - 10/24/2018 05:59 PM - Peter Amstutz

- Description updated

#5 - 10/24/2018 06:00 PM - Peter Amstutz

Strike note-2 in favor of current description.

#6 - 10/24/2018 06:40 PM - Peter Amstutz

- Description updated

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

- Story points set to 3.0

#7 - 02/20/2019 03:55 PM - Tom Morris

- Related to Bug #14710: [Workbench] Child containers run on federated clusters do not show up added