

Arvados - Story #6096

[OPS] Implement a process to regularly deploy a Docker image for running jobs to Arvados clusters

05/20/2015 07:28 PM - Brett Smith

Status:	Resolved	Start date:	05/22/2015
Priority:	Normal	Due date:	
Assigned To:	Ward Vandewege	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	2015-07-22 sprint		
Description			
<p>We're now in a place where we install all components via debs.</p> <p>There's a script in arvados-dev called</p> <pre>jenkins/run-deploy.sh</pre> <p>That script should be updated to make sure the default jobs image that matches the deployment is installed.</p> <p>This is the line that needs to be added:</p> <pre>arv keep docker --project-uuid=<public docker image project> --name="Docker image for compute nodes \$TAG" TAG</pre> <p>where TAG is the git_hash (that will be PINNED in hiera!)</p> <p>So we need to create a public project for the arvados docker images on each installation.</p> <p>Another todo: making that project needs to be a part of the install instructions. I think we should add that here:</p> <p>http://doc.arvados.org/install/create-standard-objects.html</p>			
Subtasks:			
Task # 6237: create debian package for Workbench			Resolved
Task # 6586: figure out arv-keepdocker permissions issue on 4xphq			Resolved
Task # 6528: split off arvados/jobs docker image from docker image			Resolved
Task # 6569: clean up and test arvados/jobs image			Resolved
Task # 6119: Review 6096-deploy-jobs-image			Resolved
Task # 6236: create debian package for API server			Resolved
Task # 6241: review 6096-package-rails-apps			Resolved
Task # 6573: review 6569-smarter-jobs-image			Resolved
Task # 6136: review 6135-docker-git-tag			Resolved
Task # 6135: 'docker' jenkins jobs has to tag with the git revision on push			Resolved
Related issues:			
Blocks Arvados - Feature #6348: [Deployment] [Documentation] Minimize system-...	New	06/18/2015	
Precedes (1 day) Arvados - Bug #5990: [SDKs] arv-run defaults to using arvado...	Resolved	05/25/2015	05/25/2015

Associated revisions

Revision 6c26aa2c - 06/03/2015 08:27 PM - Ward Vandewege

Build API server and Workbench debian packages.

Refs #6096

Revision 6c26aa2c - 06/03/2015 08:27 PM - Ward Vandewege

Build API server and Workbench debian packages.

Refs #6096

Revision 7761d1a6 - 06/03/2015 08:40 PM - Ward Vandewege

Fix license names in API and workbench packages.

refs #6096

Revision 7761d1a6 - 06/03/2015 08:40 PM - Ward Vandewege

Fix license names in API and workbench packages.

refs #6096

Revision 626ea7db - 06/03/2015 08:45 PM - Ward Vandewege

Default BUILD_BUNDLE_PACKAGES to false.

refs #6096

Revision 626ea7db - 06/03/2015 08:45 PM - Ward Vandewege

Default BUILD_BUNDLE_PACKAGES to false.

refs #6096

Revision 0dd51747 - 06/03/2015 09:10 PM - Ward Vandewege

Simplified run-deploy.sh now that we have moved away from Capistrano.

refs #6096

Revision 0dd51747 - 06/03/2015 09:10 PM - Ward Vandewege

Simplified run-deploy.sh now that we have moved away from Capistrano.

refs #6096

Revision 05ee30be - 06/04/2015 02:16 PM - Ward Vandewege

Name change for the arvados-api-server and arvados-workbench packages. The slim versions are now just

```
arvados-api-server
arvados-workbench
```

and the bigger version of the package that includes vendor/bundle is called

```
arvados-api-server-with-bundle
arvados-workbench-with-bundle
```

refs #6096

Revision 05ee30be - 06/04/2015 02:16 PM - Ward Vandewege

Name change for the arvados-api-server and arvados-workbench packages. The slim versions are now just

```
arvados-api-server
arvados-workbench
```

and the bigger version of the package that includes vendor/bundle is called

```
arvados-api-server-with-bundle
arvados-workbench-with-bundle
```

refs #6096

Revision 301f4e3a - 06/04/2015 02:40 PM - Ward Vandewege

Name change for the arvados-api-server and arvados-workbench packages - part 2.

refs #6096

Revision 301f4e3a - 06/04/2015 02:40 PM - Ward Vandewege

Name change for the arvados-api-server and arvados-workbench packages - part 2.

refs #6096

Revision 381b79bf - 06/04/2015 04:40 PM - Ward Vandewege

Be sure to include the pre-generated assets (tmp/cache/assets) in the workbench packages.

refs #6096

Revision 381b79bf - 06/04/2015 04:40 PM - Ward Vandewege

Be sure to include the pre-generated assets (tmp/cache/assets) in the workbench packages.

refs #6096

Revision 855e16ba - 06/05/2015 08:29 PM - Ward Vandewege

Merge branch '6096-package-rails-apps'

refs #6096

Revision 855e16ba - 06/05/2015 08:29 PM - Ward Vandewege

Merge branch '6096-package-rails-apps'

refs #6096

Revision b9bcac3e - 06/09/2015 08:35 PM - Ward Vandewege

Make the run-deploy.sh script much more general. It can now be used to deploy any of our clusters after an upgrade of the versions in hiera.

refs #6096

Revision b9bcac3e - 06/09/2015 08:35 PM - Ward Vandewege

Make the run-deploy.sh script much more general. It can now be used to deploy any of our clusters after an upgrade of the versions in hiera.

refs #6096

Revision 54cfcb7c - 06/10/2015 02:32 PM - Ward Vandewege

Add instructions to the 'Create standard objects' page to create a cluster-wide readable project for standard Arvados Docker images.

refs #6096

Revision f288c6cc - 06/11/2015 01:13 AM - Ward Vandewege

Store the full git hash of the checkout in the arvados-src package, in a file called git-commit.version

refs #6096

Revision f288c6cc - 06/11/2015 01:13 AM - Ward Vandewege

Store the full git hash of the checkout in the arvados-src package, in a file called git-commit.version

refs #6096

Revision af519091 - 06/11/2015 01:30 AM - Ward Vandewege

Make run-deploy.rb install the correct version of the arvados/jobs docker image on deploy.

refs #6096

Revision 0a139ea3 - 07/07/2015 09:12 PM - Ward Vandewege

Make run-deploy.rb install the correct version of the arvados/jobs docker image on deploy.

refs #6096

Revision 0a139ea3 - 07/07/2015 09:12 PM - Ward Vandewege

Make run-deploy.rb install the correct version of the arvados/jobs docker image on deploy.

refs #6096

Revision 6086e2a5 - 07/08/2015 02:05 AM - Ward Vandewege

Small fixes for run-deploy.sh

refs #6096

Revision 6086e2a5 - 07/08/2015 02:05 AM - Ward Vandewege

Small fixes for run-deploy.sh

refs #6096

Revision 510a6b6a - 07/08/2015 01:59 PM - Ward Vandewege

Fix up git log statements:

- remove duplication between -n1 and --max-count=1
- make sure every git log statement has --first-parent

refs #6096

Revision 510a6b6a - 07/08/2015 01:59 PM - Ward Vandewege

Fix up git log statements:

- remove duplication between -n1 and --max-count=1
- make sure every git log statement has --first-parent

refs #6096

History

#1 - 05/20/2015 07:35 PM - Nico César

- *Assigned To set to Nico César*

mental note: this is a arv-keep docker thing

#2 - 05/21/2015 07:20 PM - Nico César

- *Status changed from New to In Progress*

As far as I understand this is as simple as:

```
arv keep docker arvados/jobs
```

My question is where should this be executed and why. Anyone could help me?

#3 - 05/21/2015 07:34 PM - Brett Smith

Nico Cesar wrote:

As far as I understand this is as simple as:

```
arv keep docker arvados/jobs
```

My question is where should this be executed and why. Anyone could help me?

You've got the right basic idea. A few details to consider from here:

- We may want to be able to specify a particular version of the arvados/jobs image to deploy. `arv keep docker` can accept an image hash as input... but whether or not that's doable will depend on where you're getting the arvados/jobs image from. If you're planning on pulling from the public Docker registry, you'll need to check that you can either fetch by image hash, or specify unique version tags during the build+push process that can be retrieved later.
- We'll need to save the image this way to different clusters.
- You'll need to figure out a way to make sure all Arvados users on the cluster can see the image. This is easy enough to do if you put it in a public project (`arv keep docker` accepts a `--project-uuid` option), but you'll need to know the UUID of that project on each cluster.

So you have a sort of balancing act here. The closer you run to Jenkins (where arvados/jobs gets run), the easier it is to be sure you're working with the right image, but the more work you'll have to do to orchestrate client uploads. If you orchestrate this saving to happen somewhere on each cluster, you have less configuration to worry about in each client, but you'll have to be more careful to ensure each cluster gets the right image.

In general, running on each cluster seems to fit within our existing infrastructure better... except I'm not really sure where you would run the process. But does this at least illuminate the trade-offs for you to chew on more, at least?

#4 - 05/21/2015 08:19 PM - Nico César

Ok .. I understand that we can do a new Jenkins job that triggers this update.

I see from "docker" Job:

```
***** upload arvados/jobs image *****
```

```
The push refers to a repository [arvados/jobs] (len: 1)
c1614e5dfdfe: Buffering to Disk
c1614e5dfdfe: Image successfully pushed
c1614e5dfdfe: Image already exists
12a071038bfb: Buffering to Disk
12a071038bfb: Image successfully pushed
b98aeeb7d234: Buffering to Disk
b98aeeb7d234: Image successfully pushed
5793765dfc54: Buffering to Disk
5793765dfc54: Image successfully pushed
abad26f56450: Buffering to Disk
abad26f56450: Image successfully pushed
c78304a261ed: Buffering to Disk
c78304a261ed: Image successfully pushed
3325980672f3: Image already exists
8ce15197d12a: Image already exists
59bc1380e0a6: Image already exists
79ace1046749: Image already exists
df9ac9bc06e6: Image already exists
a74ae6b4dab6: Image already exists
1b430bab60ed: Image already exists
e3551d68778e: Image already exists
Digest: sha256:473ac041771ecf2b0e22d0ef42f650764cc42d82d515746d9d3573da9fa9a7d1
```

I also see that run-deploy.sh has

```
ssh -p2222 root@$IDENTIFIER.arvadosapi.com -C "/...
```

I think I could do a combination of both inside run-docker-test.sh with a flag (something like --update-clusters qr1hi,4xphq,9tee4) and use the sha256 that it's returned by the docker push

Opinions on this?

#5 - 05/22/2015 12:57 PM - Nico César

mental note:

from [#3847](#)

```
arv keep docker --project-uuid=qr1hi-j7d0g-5931q8oed0gynt3 --name="Docker image for compute nodes $TAG" TAG
```

where TAG is the git_hash (that will be PINNED in hiera!)

Brett, how do I obtain the project-uuid's to apply?

#6 - 05/22/2015 01:30 PM - Nico César

see [#6135](#) . docker images will have tags

#7 - 06/01/2015 09:15 PM - Nico César

docker images now HAVE tags.

see: <https://registry.hub.docker.com/u/arvados/jobs/tags/manage/>

you can already see:

```
178d3f36265e0e9e9cc0bb6ac8c7c47a9c701687
1ec1d552c77e18e2912e400ae395ca00f4e51c3c
7a53d874994a5a9af273cee1329d9635b7e03edb
9413eb733015601af699f2027d9a7a5bad3f3dea
```

#8 - 06/04/2015 07:46 PM - Nico César

I reviewed 381b79bf5cfefe790bdcc24dd33296b8518e4c19 6096-package-rails-apps

looks good to me. Let's merge

#9 - 06/05/2015 08:44 PM - Ward Vandewege

Cool, merged 6096-package-rails-apps into arvados-dev.

#10 - 06/05/2015 08:59 PM - Ward Vandewege

- Description updated

#11 - 06/09/2015 05:39 PM - Ward Vandewege

- Description updated

#12 - 06/11/2015 01:43 PM - Radhika Chippada

- Target version changed from 2015-06-10 sprint to 2015-07-08 sprint

#13 - 06/15/2015 06:38 PM - Nico César

I'm checking 2b93735fc87e447301afafe6556d8571afef2bcf

I see that there are 3 commits on 6096-deploy-jobs-image:

2b93735fc87e447301afafe6556d8571afef2bcf
6f1a22656665643dbba71e59099171d69554b2ad
c06542419737cddd5adda84d4b14e0b88912d0f7

that are related to rpm packages from FPM ... is that needed for this ticket? Anyways, I went through the code,

- It assumes that "ssh shell.\$IDENTIFIER" remote execution and does /usr/local/rvm/bin/rvm-exec.... which user in the other end will be executing this? what's your ssh config? the same with ssh \$IDENTIFIER cat /usr/local/arvados/src/git-commit.version

```
◦ Host *.qrlhi  
ProxyCommand ssh turnout@switchyard.qrlhi.arvadosapi.com $SSH_PROXY_FLAGS %h
```

#14 - 07/07/2015 06:51 PM - Ward Vandewege

Nico Cesar wrote:

I'm checking 2b93735fc87e447301afafe6556d8571afef2bcf

I see that there are 3 commits on 6096-deploy-jobs-image:

2b93735fc87e447301afafe6556d8571afef2bcf
6f1a22656665643dbba71e59099171d69554b2ad
c06542419737cddd5adda84d4b14e0b88912d0f7

that are related to rpm packages from FPM ... is that needed for this ticket?

I should have made a separate branch for these, sorry. Only tangentially related.

Anyways, I went through the code,

- It assumes that "ssh shell.\$IDENTIFIER" remote execution and does /usr/local/rvm/bin/rvm-exec.... which user in the other end will be executing this? what's your ssh config? the same with ssh \$IDENTIFIER cat /usr/local/arvados/src/git-commit.version
 - [...]

All this code runs as you - the user who runs the deploy script.

#15 - 07/08/2015 07:12 PM - Brett Smith

- Target version changed from 2015-07-08 sprint to 2015-07-22 sprint

#16 - 07/08/2015 08:13 PM - Ward Vandewege

- Assigned To changed from Nico César to Ward Vandewege

- Story points changed from 2.0 to 0.5

#17 - 07/08/2015 08:19 PM - Ward Vandewege

- Subject changed from *[OPS] Implement a process to regularly deploy the a Docker image for running jobs to Arvados clusters* to *[OPS] Implement a process to regularly deploy a Docker image for running jobs to Arvados clusters*

#18 - 07/13/2015 03:41 PM - Ward Vandewege

- Status changed from *In Progress* to *Resolved*