

## Arvados - Feature #14325

### [crunch-dispatch-cloud] Dispatch containers to cloud VMs directly, without slurm or nodemanager

10/10/2018 03:44 PM - Tom Clegg

<b>Status:</b>	Resolved	<b>Start date:</b>	01/28/2019
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assigned To:</b>	Tom Clegg	<b>% Done:</b>	100%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	2019-02-13 Sprint		
<b>Description</b>			
This issue covers the smallest version of <a href="#">Dispatching containers to cloud VMs</a> that can be deployed on a dev cluster.			
Background -- already done in <a href="#">#14360</a> :			
<ul style="list-style-type: none"><li>• Bring up nodes and run containers on them</li><li>• Structured logs for diagnostics+statistics: cloud API errors, node lifecycle, container lifecycle</li><li>• HTTP status report with current set of containers (queued/running) and VMs (busy/idle) -- see <a href="#">Dispatching containers to cloud VMs</a> "Operator view"</li><li>• Shutdown idle nodes automatically</li><li>• Handle cloud API quota errors</li><li>• Package-building changes are in place, but commented out</li></ul>			
Requirements covered here:			
<ul style="list-style-type: none"><li>• Ops mechanism for draining a node (e.g., curl command using a management token) -- see <a href="#">Dispatching containers to cloud VMs</a> "Management API"</li><li>• Resource consumption metrics (number of instances, number of containers running, total hourly price of all existing VMs) -- see <a href="#">Dispatching containers to cloud VMs</a> "Metrics"</li><li>• Drain (not kill) instances that exist at startup, fail boot probe, but are already running containers -- see <a href="#">Dispatching containers to cloud VMs</a> "Special cases / synchronizing state"</li><li>• Configurable port number for connecting to VM SSH servers</li><li>• Pass API host and dispatcher's token to crunch-run command via ARVADOS_API_* environment variables</li><li>• Test SSH host key verification (dispatcher's token is not sent to a remote host unless the host's SSH key passes the VerifyHostKey() method provided by the cloud driver)</li><li>• Test container.Queue using real railsAPI/controller</li><li>• Test resuming state after restart (some instances are booting, some idle, some running containers, some draining, some on admin-hold)</li><li>• Cancel container after some number of start/requeue cycles (i.e., crunch-run --detach succeeded, but child exited without moving container past Locked state)</li><li>• Cancel container with no suitable instance type</li><li>• Enable package build</li><li>• Handle cloud API ratelimit errors (obey holdoff time returned by driver... incl. test)</li><li>• Update management API response format (lowercase keys)</li><li>• Confirm all probe failures are logged once instance is booted (see <a href="#">#14360#note-38</a>, fixed in <a href="#">7a047d8b6</a>)</li></ul>			
Requirements covered elsewhere:			
<ul style="list-style-type: none"><li>• One cloud vendor driver (Azure = <a href="#">#14324</a>)</li><li>• Production-readiness (<a href="#">#14807</a>)</li></ul>			
Refs			
<ul style="list-style-type: none"><li>• <a href="#">Dispatching containers to cloud VMs</a></li><li>• <a href="#">#13964</a> spike</li></ul>			
<b>Subtasks:</b>			
Task # 14664: Review 14325-dispatch-cloud			<b>Resolved</b>
<b>Related issues:</b>			
Related to Arvados - Feature #14324: [crunch-dispatch-cloud] Azure driver		<b>Resolved</b>	<b>01/09/2019</b>
Related to Arvados - Bug #13964: crunch-dispatch-cloud spike		<b>Resolved</b>	
Related to Arvados - Story #13908: [Epic] Replace SLURM for cloud job schedul...		<b>Resolved</b>	

Related to Arvados - Story #14360: [crunch-dispatch-cloud] Merge incomplete i...	<b>Resolved</b>	<b>10/26/2018</b>
Precedes Arvados - Story #14796: [crunch-dispatch-cloud] Document installatio...	<b>Resolved</b>	<b>01/29/2019</b>
Precedes Arvados - Story #14807: [arvados-dispatch-cloud] Features/fixes need...	<b>Resolved</b>	<b>01/29/2019</b>

## Associated revisions

---

### Revision 800139c8 - 02/06/2019 09:29 PM - Tom Clegg

Merge branch '14325-dispatch-cloud'

closes #14325

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <[tclegg@veritasgenetics.com](mailto:tclegg@veritasgenetics.com)>

### Revision 4f2aab84 - 02/12/2019 06:10 AM - Tom Clegg

Merge branch '14325-dispatch-cloud'

refs #14325

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <[tclegg@veritasgenetics.com](mailto:tclegg@veritasgenetics.com)>

### Revision b56e2857 - 02/12/2019 04:25 PM - Tom Clegg

14325: Fix PrivateKey config type.

JSON decoder expects []byte fields to be base64-encoded, which we don't want here.

refs #14325

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <[tclegg@veritasgenetics.com](mailto:tclegg@veritasgenetics.com)>

### Revision a4396e18 - 02/12/2019 07:17 PM - Tom Clegg

14325: Start up immediately if there are no stale locks.

...instead of waiting for the pool to send a notification to trigger the first loop iteration.

refs #14325

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <[tclegg@veritasgenetics.com](mailto:tclegg@veritasgenetics.com)>

### Revision d06417e5 - 02/13/2019 07:00 AM - Tom Clegg

14325: Start up immediately if there are no stale locks.

...instead of waiting for the pool to send a notification to trigger the first loop iteration.

refs #14325

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <[tclegg@veritasgenetics.com](mailto:tclegg@veritasgenetics.com)>

### Revision e6cb52f3 - 02/15/2019 05:03 AM - Tom Clegg

14325: Start up immediately if there are no stale locks.

...instead of waiting for the pool to send a notification to trigger the first loop iteration.

refs #14325

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <[tclegg@veritasgenetics.com](mailto:tclegg@veritasgenetics.com)>

## History

---

### #1 - 10/10/2018 03:44 PM - Tom Clegg

- Related to Feature #14324: [crunch-dispatch-cloud] Azure driver added

### #2 - 10/10/2018 03:44 PM - Tom Clegg

- Related to Bug #13964: crunch-dispatch-cloud spike added

**#3 - 10/10/2018 03:44 PM - Tom Clegg**

- Related to Story #13908: [Epic] Replace SLURM for cloud job scheduling/dispatching added

**#4 - 10/10/2018 04:59 PM - Tom Clegg**

- Description updated

**#5 - 10/10/2018 05:08 PM - Tom Clegg**

- Description updated

**#6 - 10/10/2018 05:31 PM - Tom Clegg**

- Description updated

**#7 - 10/16/2018 03:07 PM - Tom Clegg**

- Description updated

**#8 - 10/17/2018 05:59 AM - Tom Clegg**

- Related to Story #14360: [crunch-dispatch-cloud] Merge incomplete implementation added

**#9 - 11/14/2018 09:29 PM - Tom Clegg**

- Description updated

**#10 - 11/15/2018 07:31 PM - Tom Clegg**

- Description updated

**#11 - 11/15/2018 08:20 PM - Tom Clegg**

- Description updated

**#12 - 11/19/2018 03:05 PM - Tom Clegg**

- Description updated

**#13 - 12/07/2018 10:02 PM - Tom Clegg**

- Description updated

**#14 - 12/12/2018 04:46 PM - Tom Morris**

- Target version set to To Be Groomed

**#15 - 12/12/2018 09:23 PM - Tom Clegg**

- Description updated

**#16 - 12/17/2018 02:53 PM - Tom Clegg**

- Description updated

**#17 - 12/18/2018 08:53 PM - Tom Clegg**

- Description updated

- Target version deleted (To Be Groomed)

**#18 - 12/19/2018 03:43 PM - Tom Clegg**

- Description updated

**#19 - 12/19/2018 03:51 PM - Tom Clegg**

- Description updated

**#20 - 12/19/2018 03:55 PM - Tom Clegg**

- Description updated

**#21 - 12/19/2018 05:56 PM - Tom Clegg**

- Description updated

**#22 - 12/19/2018 06:07 PM - Tom Clegg**

- Description updated

**#23 - 12/19/2018 06:15 PM - Tom Clegg**

- Target version set to Arvados Future Sprints

- Story points set to 4.0

**#24 - 12/20/2018 05:55 PM - Peter Amstutz**

- Description updated

**#25 - 12/20/2018 06:08 PM - Peter Amstutz**

Management APIs should return {"items": [...]} not {"Items": [...]} for consistency with the Arvados API.

**#26 - 12/20/2018 06:09 PM - Peter Amstutz**

- Description updated

**#27 - 12/20/2018 06:39 PM - Tom Clegg**

- Description updated

**#28 - 12/20/2018 06:42 PM - Tom Clegg**

- Description updated

**#29 - 12/21/2018 08:38 PM - Tom Clegg**

- Description updated

**#30 - 12/21/2018 08:39 PM - Tom Clegg**

- Status changed from New to In Progress

- Assigned To set to Tom Clegg

**#31 - 01/02/2019 04:21 PM - Tom Morris**

- Target version changed from Arvados Future Sprints to 2019-01-16 Sprint

**#32 - 01/16/2019 04:14 PM - Tom Clegg**

- Target version changed from 2019-01-16 Sprint to 2019-01-30 Sprint

**#33 - 01/25/2019 09:55 PM - Tom Clegg**

- Ops mechanism for draining a node (e.g., curl command using a management token) -- see [Dispatching containers to cloud VMs](#) "Management API"

Added "hold" and "drain". (Wiki also mentions a "kill" API -- not included here.)

- Resource consumption metrics (number of instances, number of containers running, total hourly price of all existing VMs) -- see [Dispatching containers to cloud VMs](#) "Metrics"

Added total hourly price. The others were already in place.

- Drain (not kill) instances that exist at startup, fail boot probe, but are already running containers -- see [Dispatching containers to cloud VMs](#) "Special cases / synchronizing state"

Added what the wiki says, which is a little different:

"...instances are left alive at least until the containers finish. After that, the usual rules apply: if boot probe succeeds before boot timeout, start scheduling containers; otherwise, shut down."

This is a bit more consistent since it's more consistent with the "inherited node is *not* running a container and fails boot probe" case: we allow the boot

timeout to run out before killing it, rather than expecting its boot probe to succeed before the existing container finishes.

- Configurable port number for connecting to VM SSH servers

CloudVMs→SSHPort can be given as a name ("ssh") or number ("22").

- Pass API host and dispatcher's token to crunch-run command via ARVADOS\_API\_\* environment variables

Added.

- Test SSH host key verification (dispatcher's token is not sent to a remote host unless the host's SSH key passes the VerifyHostKey() method provided by the cloud driver)

Added.

- Test container.Queue using real railsAPI/controller

Added. Revealed & fixed SDK bug, see [f696f142e](#).

- Test resuming state after restart (some instances are booting, some idle, some running containers, some draining, some on admin-hold)

Added restart/resume test to confirm "hold" and instance-type labels are saved/loaded effectively.

Added a slew of worker tests to confirm proper state changes in probeAndUpdate.

- Cancel container after some number of start/requeue cycles (i.e., crunch-run --detach succeeded, but child exited without moving container past Locked state)

Didn't do this. (We've already implemented it on the API side.)

- Cancel container with no suitable instance type

Added.

- Enable package build

Uncommented.

- Handle cloud API ratelimit errors (obey holdoff time returned by driver... incl. test)

Added.

- Update management API response format (lowercase keys)

Updated.

- Confirm all probe failures are logged once instance is booted (see [#14360#note-38](#), fixed in [7a047d8b6](#))

Confirmed.

14325-dispatch-cloud @ [b105602902e38f18a48505e2091ffea77b2c7c89https://ci.curoverse.com/view/Developer/job/developer-run-tests/1040/](#)

**#34 - 01/28/2019 02:46 PM - Tom Clegg**

Now at [a27b2bf3e](#) with some test cleanup (move LamelInstanceSet's one remaining useful feature to StubDriver and retire LamelInstanceSet).

**#35 - 01/30/2019 04:12 PM - Tom Clegg**

- Target version changed from 2019-01-30 Sprint to 2019-02-13 Sprint

**#36 - 01/30/2019 04:23 PM - Tom Clegg**

- Story points changed from 4.0 to 1.0

### #37 - 01/30/2019 04:48 PM - Tom Clegg

- Precedes Story #14796: [crunch-dispatch-cloud] Document installation / migration from c-d-slurm + node manager added

### #38 - 01/31/2019 10:35 PM - Peter Amstutz

worker.shutdownIfIdle():

```
if !(wkr.state == StateIdle || (wkr.state == StateBooting && wkr.idleBehavior == IdleBehaviorDrain)) {
    return false
}
```

The double-negative logic (do nothing if these things are NOT true...) makes this expression hard to read. Please add comments clarifying the intention that we want to shut down when certain things are true.

```
if wkr.idleBehavior != IdleBehaviorDrain && age < wkr.wp.timeoutIdle {
    return false
}
```

Same comment about confusing expression. If I'm understanding the intended behavior, it would be clearer to write `wkr.idleBehavior == IdleBehaviorRun && age < wkr.wp.timeoutIdle` because the `IdleBehaviorHold` case has already been eliminated, and `IdleBehaviorDrain` ignores the timeout (but having `IdleBehaviorDrain` and `timeoutIdle` appear on the same line implies they are related).

Queue.Update():

```
if _, keep := cq.dontupdate[uuid]; keep {
    continue
}
```

...

```
if _, keep := cq.dontupdate[uuid]; keep {
    continue
} else if _, keep = next[uuid]; keep {
    continue
} else {
    delete(cq.current, uuid)
}
```

Comment from last time that "keeplocal" was confusing and was renamed to "dontupdate" but there's still a few local variables called "keep" and I don't know how to read it. Should those also be called "dontupdate"? Maybe add some comments?

In `Queue.addEnt()` there's an embedded assumption that if the current dispatcher can't find an instance type for a container, nobody can, so it should always cancel the container (even if it has to lock it first). I think that's fine (heterogeneous dispatchers has complexity we don't want to get into yet, if ever) but should probably be mentioned in a comment.

worker.probeAndUpdate():

```
for _, uuid := range ctrUUIDs {
    running[uuid] = struct{}{}
    if _, ok := wkr.running[uuid]; !ok {
        changed = true
    }
}
```

Another place that would benefit from some more comments expressing the intent / context of the code. I think what this is doing is determining if there a container UUID was found on the node which isn't present in `wkr.running`. The next block looks like it checks the opposite case where a container is known to `wkr.running` but not present on the instance.

```
if wkr.state == StateUnknown || wkr.state == StateBooting {
    wkr.state = StateIdle
}
```

It is implied by getting to this point in the code that `probeBooted()` and `probeRunning()` both passed successfully, could use a comment making that assumption explicit.

... to be continued ....

### #39 - 02/01/2019 03:36 PM - Peter Amstutz

Pool.Unallocated():

```
if !(wkr.state == StateIdle || wkr.state == StateBooting || wkr.state == StateUnknown) || wkr.idleBehavior != IdleBehaviorRun || len(wkr.running) > 0 {
```

```
        continue
    }
```

This line is way too long.

Similar comment to earlier about hard-to-follow double-negative logic. Here `!(wkr.state StateIdle || wkr.state StateBooting || wkr.state == StateUnknown)` is much clearer written as `(wkr.state != StateIdle && wkr.state != StateBooting && wkr.state != StateUnknown)`

#### #40 - 02/01/2019 04:18 PM - Peter Amstutz

Cancel container after some number of start/requeue cycles (i.e., `crunch-run --detach` succeeded, but child exited without moving container past Locked state)

Didn't do this. (We've already implemented it on the API side.)

We've agreed to do so, but haven't actually done it yet ([#11561](#))

```
# git.curoverse.com/arvados.git/lib/dispatchcloud/container
./queue_test.go:38:17: undefined: test
./queue_test.go:95:17: undefined: test
FAIL    git.curoverse.com/arvados.git/lib/dispatchcloud/container [build failed]
```

```
import (
    "github.com/julienschmidt/httprouter"
)
```

What's the goal of introducing yet another routing framework here? We already use both `http.ServeMux` and `gorilla/mux`.

```
# Layouter fails if we add these
```

Maybe use `graphviz` instead? (Requires slightly different notation).

#### #41 - 02/01/2019 09:47 PM - Tom Clegg

Indeed, those are some confusing conditional expressions, thanks. Clarified and added comments.

Peter Amstutz wrote:

Didn't do this. (We've already implemented it on the API side.)

We've agreed to do so, but haven't actually done it yet ([#11561](#))

Ah, indeed.

What's the goal of introducing yet another routing framework here? We already use both `http.ServeMux` and `gorilla/mux`.

Cheap, easy to use, does what we need (filter on methods + extract path params), didn't think of a reason not to use it. (It happens to be much more efficient with time and memory than `gorilla`, not that that's a big concern here.)

Maybe use `graphviz` instead? (Requires slightly different notation).

Maybe. I didn't find this ascii art exercise particularly rewarding. If I were to sink more time into different ways of doing this, I'd probably just give up and make a drawing in Google Drive.

`lib/dispatchcloud/container` tests are fixed.

14325-dispatch-cloud @ [71fd4da18b22100682ae7e2079aadfd66360d310https://ci.curoverse.com/view/Developer/job/developer-run-tests/1051/](https://ci.curoverse.com/view/Developer/job/developer-run-tests/1051/)

#### #42 - 02/04/2019 10:17 PM - Tom Clegg

- Precedes Story #14807: *[arvados-dispatch-cloud] Features/fixes needed before first production deploy added*

#### #43 - 02/04/2019 10:18 PM - Tom Clegg

- Description updated

**#44 - 02/05/2019 10:27 PM - Tom Clegg**

Added one more fix that wasn't mentioned here: Log stderr from last boot-probe when giving up on boot.

14325-dispatch-cloud @ [ee53a267ded17bc50eaf4dfebba5ff4a3273753](https://ci.curoverse.com/view/Developer/job/developer-run-tests/1053/)

**#45 - 02/06/2019 06:00 PM - Peter Amstutz**

Tom Clegg wrote:

Added one more fix that wasn't mentioned here: Log stderr from last boot-probe when giving up on boot.

14325-dispatch-cloud @ [ee53a267ded17bc50eaf4dfebba5ff4a3273753](https://ci.curoverse.com/view/Developer/job/developer-run-tests/1053/)

This LGTM, thanks.

**#46 - 02/06/2019 10:38 PM - Tom Clegg**

- *Status changed from In Progress to Resolved*

Applied in changeset [arvados|800139c8dee7d9a563a8a2dca9e45e283c55c22c](https://arvados.org/changeset/arvados|800139c8dee7d9a563a8a2dca9e45e283c55c22c).

**#47 - 03/01/2019 07:33 PM - Tom Morris**

- *Release set to 15*