

Arvados - Feature #14324

[crunch-dispatch-cloud] Azure driver

10/10/2018 03:30 PM - Tom Clegg

Status: Resolved	Start date: 01/09/2019
Priority: Normal	Due date:
Assigned To: Peter Amstutz	% Done: 100%
Category:	Estimated time: 0.00 hour
Target version: 2019-01-16 Sprint	
Description Implement crunch-dispatch-cloud's cloud.Driver interface using Azure client library . See <ul style="list-style-type: none">source:lib/cloud/interfaces.go@13964-dispatch-cloudsource:lib/dispatchcloud/test/stub_driver.go@13964-dispatch-cloud	
Subtasks: Task # 14661: Review 14324-cdc-azure Resolved	
Related issues: Related to Arvados - Feature #14325: [crunch-dispatch-cloud] Dispatch contain... Resolved 01/28/2019 Related to Arvados - Story #13908: [Epic] Replace SLURM for cloud job schedul... Resolved	

Associated revisions

Revision e62dc0ff - 01/15/2019 06:40 PM - Peter Amstutz

Merge branch '14324-cdc-azure' refs #14324

Arvados-DCO-1.1-Signed-off-by: Peter Amstutz <pamstutz@veritasgenetics.com>

History

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

- Related to Feature #14325: [crunch-dispatch-cloud] Dispatch containers to cloud VMs directly, without slurm or nodemanager added

#2 - 10/10/2018 04:22 PM - Tom Clegg

- Tracker changed from Bug to Feature

#3 - 10/10/2018 04:24 PM - Peter Amstutz

previous work:

13964-cdc-azure @ [2e609dce27fc6b9f39857cf7b0a9904719d2d526](#)

#4 - 10/10/2018 04:26 PM - Tom Clegg

- Description updated

#5 - 10/10/2018 04:48 PM - Tom Morris

- Target version set to Arvados Future Sprints

- Story points set to 3.0

Time boxed with 2 pts of tests.

#6 - 12/19/2018 03:26 PM - Tom Clegg

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

#7 - 01/02/2019 04:36 PM - Tom Morris

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

#8 - 01/02/2019 04:39 PM - Tom Morris

- Assigned To set to Peter Amstutz

#9 - 01/09/2019 09:41 PM - Peter Amstutz

14324-cdc-azure @ [1b2a4c98249e3a46242f7d10c00a70feeb2a843](#)

<https://ci.curoverse.com/view/Developer/job/developer-run-tests/1025>

#10 - 01/10/2019 02:18 PM - Peter Amstutz

A few things I realized I still need to do:

- Testing lib/dispatchcloud doesn't run tests in sub-directories, so I overlooked a test that's failing to compile
- Document how to run individual test cases against real cloud
- ManageBlobs should run periodically in the background
- Logging probably needs to be cleaned up
- Determine if the underlying SDK handles refreshing expired tokens automatically or if our code needs to do it (needed to do this in libcloud)

#11 - 01/10/2019 06:29 PM - Peter Amstutz

Peter Amstutz wrote:

A few things I realized I still need to do:

- Testing lib/dispatchcloud doesn't run tests in sub-directories, so I overlooked a test that's failing to compile

Fixed

- Document how to run individual test cases against real cloud

Done

- ManageBlobs should run periodically in the background

Done

- Determine if the underlying SDK handles refreshing expired tokens automatically or if our code needs to do it (needed to do this in libcloud)

It appears this is handled by the SDK.

- Logging probably needs to be cleaned up

Getting rid of spurious development logging, but need to decide what to do about background tasks that can experience errors

#12 - 01/11/2019 03:40 PM - Peter Amstutz

14324-cdc-azure @ [443a0b96316ed46600dc5035193adae6ac4d1f74](#)

<https://ci.curoverse.com/view/Developer/job/developer-run-tests/1027/>

Sorted out above issues, ready for review comments.

#13 - 01/14/2019 03:50 PM - Peter Amstutz

How about 14324-cdc-azure @ [f398e9d75d32fbfa6804041353997bf74195e489](#)

<https://ci.curoverse.com/view/Developer/job/developer-run-tests/1031/>

#14 - 01/14/2019 04:42 PM - Lucas Di Pentima

Reviewing [f398e9d75d32fbfa6804041353997bf74195e489](#)

- Didn't run the driver tests against the real cloud, as I don't have an account on Azure. Should I ask ops for some credentials? OTOH, running them without credentials returns [no tests to run], I think it's expected to run anyways, right?
- On file cloud/azure.go:
 - Line 162: Are there other error statuses that would require throttling?
 - Line 169: Should that conditional need to be reversed?
 - Lines 173-177: Should that be an else block on line 171?
 - Line 258: Shouldn't ManageBlobs() be called one more time before returning?

- Line 268: Are these 4 iterations a way to add 4 attempts to NIC & blob cleanup? Maybe this value should be configurable?

#15 - 01/14/2019 07:23 PM - Peter Amstutz

Lucas Di Pentima wrote:

Reviewing [f398e9d75d32fbfa6804041353997bf74195e489](https://github.com/14324-cdc-azure/commit/f398e9d75d32fbfa6804041353997bf74195e489)

- Didn't run the driver tests against the real cloud, as I don't have an account on Azure. Should I ask ops for some credentials? OTOH, running them without credentials returns [no tests to run], I think it's expected to run anyways, right?

There was a missing file that called check.TestingT(t). Fixed.

- On file cloud/azure.go:
 - Line 162: Are there other error statuses that would require throttling?

I don't know. Node manager goes into a backoff-and-retry loop with error codes ≥ 500 , but in crunch-dispatch-cloud the scheduler handles retries and doesn't do any backoff. We could react to 500 errors as "wait and try again in 20 seconds".

- Line 169: Should that conditional need to be reversed?
- Lines 173-177: Should that be an else block on line 171?

Yes, good catch.

- Line 258: Shouldn't ManageBlobs() be called one more time before returning?

Not sure. Depends on whether it is more desirable that when context is Done (because Stop() was called) it should terminate right away, or that it maximizes the opportunity to garbage collect blobs.

- Line 268: Are these 4 iterations a way to add 4 attempts to NIC & blob cleanup? Maybe this value should be configurable?

It is parallelizing to 4 threads for the case where it is deleting a batch of NICs or blobs and individual operations have latency. If an individual operation fails, it will get retried on next call to ManageNICs or ManageBlobs.

14324-cdc-azure @ [6756ee790a6c685b96bc1ec80b67e7d6c94e9846](https://github.com/14324-cdc-azure/commit/6756ee790a6c685b96bc1ec80b67e7d6c94e9846)

#16 - 01/14/2019 09:25 PM - Peter Amstutz

- Status changed from New to In Progress

#17 - 01/15/2019 02:44 PM - Lucas Di Pentima

I think there's a file missing:

```
-----  
FAIL: azure_test.go:137: AzureInstanceSetSuite.TestCreate
```

```
azure_test.go:144:  
    c.Assert(err, check.IsNil)  
... value *os.PathError = &os.PathError{Op:"open", Path:"azconfig_sshkey.pub", Err:0x2} ("open azconfig_sshkey  
.pub: no such file or directory")
```

```
time="2019-01-15T09:36:29-05:00" level=warning msg="Couldn't get account keys" error="storage.AccountsClient#L  
istKeys: Invalid input: autorest/validation: validation failed: parameter=resourceGroupName constraint=MinLeng  
th value=\\\" details: value length must be greater than or equal to 1"
```

```
OOPS: 8 passed, 1 FAILED  
--- FAIL: Test (0.00s)
```

#18 - 01/15/2019 06:32 PM - Peter Amstutz

Lucas Di Pentima wrote:

I think there's a file missing:

[...]

Ah yea, fixed. 14324-cdc-azure @ [a27348943c6e35eee983d6a333eb6977394c3f13](https://github.com/14324-cdc-azure/commit/a27348943c6e35eee983d6a333eb6977394c3f13)

#19 - 01/15/2019 06:38 PM - Lucas Di Pentima

This LGTM, thanks!

#20 - 01/15/2019 06:41 PM - Peter Amstutz

- *Status changed from In Progress to Resolved*

#21 - 03/01/2019 06:30 PM - Tom Morris

- *Release set to 15*