Creating ocs pipeline to run tier1 tests 20-30% faster
Gabriella Roman, May-Aug 2020

Introduction

- Tier1 tests are a predefined set of simple tests, compared to tier2 and tier3 tests which involve tests of a higher level of complexity
    - Therefore, this should be one of the easiest places to start in reducing ocs pipeline running time
- After running a series of diagnostics, it was determined that tier1 tests take approximately 210 minutes to run (at start of internship), so the objective was to reduce running time by approximately 40-60 minutes
- By the completion of my internship I had created a new ocs pipeline that ran successfully, and found a way to accurately measure each test's duration

Timeline

May

- Intern orientation & onboarding
- Reviewed team Trello board to see where I could become involved
    - Rook
    - Ceph-fs
    - Cephadm

June

- Went over possible projects with Christina
    - Creating ocs pipeline to reduce tier1 tests running time was assigned to me
    - Was introduced to point of contact & mentor, Zack Cerza
- Much of the month was spent learning about and familiarizing self with related information through GitHub repositories and online documentation
    - Jenkins: server used to automate building, testing, and deploying software
    - Pipeline: a series of data processing elements in which each step's input is dependent on the output of its preview step
    - Groovy language (used to orchestrate Jenkins pipeline)
- Set up virtual environment with help from Zack
- Learned to use some of the ceph-jenkins features
- Created new ocs pipeline by cloning ocs-ci.yml file and making necessary changes
    - Added tier1 testing stage to pipeline
    - Pushed changes to my forked repository (first commit!)
        - Build unfortunately failed
            - Took a look through console output to figure out what might've caused the failure and noticed that several of the stages were being skipped
        - With the help of Zack, the job ran successfully after some more rebuilds
- Learned to use pytest in the virtual environment to determine which tests specifically are failing

- Learned to identify failures in the code by inspecting the log file

July
- Became more familiar with using Git commands, CLI, and Vim
  - Git stash
  - Cherry picking
- Began looking through different tests to try and determine where there is inefficiency
- Learned to utilize build artifacts
  - Jenkins job > build artifacts > logs > report.html
  - Useful in sorting tests with longest running time durations
- Began brainstorming possible solutions to reducing running time
  - Decrease number of loops
  - Use values directly where possible
  - Implementation of algorithms with less time complexity

August

- Continued looking for code in tier1 tests that wasn't optimal
- Had Zack walk me through his process for optimizing a given tier1 test
- Learned to use a profiler, specifically pytest's profiling plugin (runs code and provides detailed breakdown of each test's execution time)
- Worked on redirecting Jenkins job
- Created a new instance of pipeline & it ran successfully!

Obstacles

- Ran into issue with py 3.7 data classes
  - Data classes add generated methods to classes
  - Issue was within setup.py file
    - dataclasses: python_version < "3.7";
  - Solution: version is only needed if dependency is there
- Job build failures
- Cloning of ocs-ci.yml file
  - Zack later realized that this was the wrong approach for creating a second job because it returns a job with a different name

Results

- Tests that were originally slower became faster
  - test_nodes_maintenance.py initially had the longest running time, became nearly twice as fast
- Overall running time of tier1 tests increased significantly due to new code introduced: from 12591.16 seconds to 21479.84 seconds (most recent report.html summary below)

Summary

58 tests ran in 21479.84 seconds.

(Un)check the boxes to filter the results.

☑ 52 passed, ☑ 0 skipped, ☑ 6 failed, ☑ 3 errors, ☑ 0 expected failures, ☑ 0 unexpected passes

Achievements

- Created new pipeline & had it run successfully
- Determined reliable method to measure tier1 tests running time duration

Future Work

- Continue reduction of tier1 tests running time

Finally, a big thank you to my manager, Christina Meno, and mentor, Zack Cerza!