parsl-visualize

Ken Herner

parsl-visualize

- ## What it does
  - Graphical interface for monitoring campaign status
  - Can see job states, dependencies, resource consumption
- ## What it does not do
  - Lacks bps report-like functionality (overview dump to terminal)
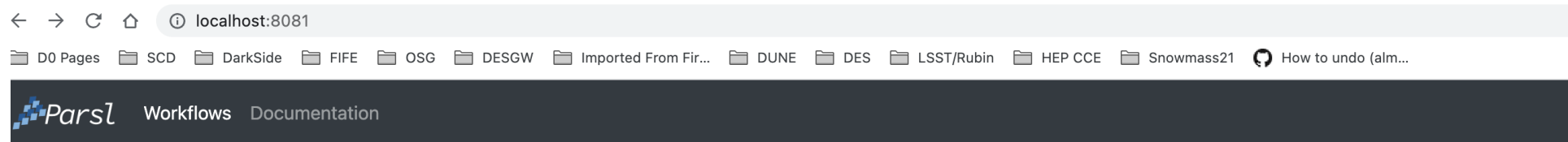 - Does not link directly to logs for viewing in browser (provides path only)

How to use it

- Requires some debugging options to be enabled at submit time
  - You get a monitoring.db file inside the runinfo directory in your submit dir
- Some of the required packages are not in the Rubin stack
  - I've created an extra conda env you're welcome to use
  - Minimal list: python, flask, postgresql, parsl, networkx, numpy, libpq, pandas, libxml2, sqlalchemy + deps
- To set up, do
  - $ conda activate /sdf/group/rubin/u/kherner/testconda
  - $ export PYTHONPATH=/sdf/group/rubin/u/kherner/testconda/lib/python3.10/site-packages:$PYTHONPATH
  - $ parsl-visualize -p portnumber sqlite:////path/to/your/monitoring.db (default port 8080)
- You should see something like

```
(/sdf/group/rubin/u/kherner/testconda) [kherner@sdfrome002 DM-36026]$ parsl-visualize -p 8081 sqlite:////sdf/home/e/elhoward/u/repo-main
-logs/DM-36026/runinfo/monitoring.db &
[1] 3740401
(/sdf/group/rubin/u/kherner/testconda) [kherner@sdfrome002 DM-36026]$  * Serving Flask app 'parsl.monitoring.visualization.app'
 * Debug mode: off
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:8081
Press CTRL+C to quit
```

Setup part 2

- Note what login machine you're on
- Set up ssh tunnel to it on your local machine with your port number, e.g.:
  - `$ ssh -L portnum:localhost:portnum -J username@s3dflogin.slac.stanford.edu sdfrome002.slac.stanford.edu`
- Finally, point your local browser to localhost:portnum (may have multiple runs depending on what you've submitted from that dir)

Looking around

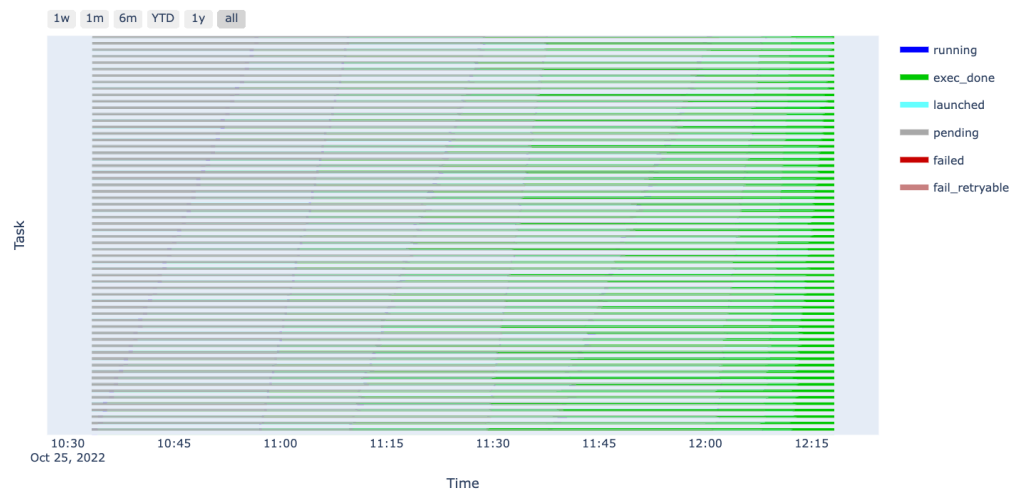## ApPipe-DC2-Test.DM-36026

Workflow Summary

- **Started:** 2022-10-25 10:33:22
- **Completed:** 2022-10-25 12:18:12
- **Workflow duration:** 1:44:49
- **Owner:** elhoward
- **Host:** sdfrome001
- **Run directory:** /sdf/data/rubin/user/elhoward/repo-main-logs/DM-36026/runinfo/001
- **Number of completed tasks:** 2174
- **Number of failed tasks:** 2

App Summary

Name	Count
calibrate	272
characterizeImage	272
detectAndMeasure	272
diaPipe	272
isr	272
retrieveTemplate	272
subtractImages	272
transformDiaSrcCat	272

View workflow DAG -- colored by apps
View workflow DAG -- colored by task states
View workflow resource usage

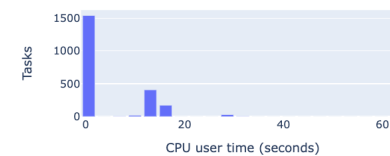Task names not in run order

Legend:
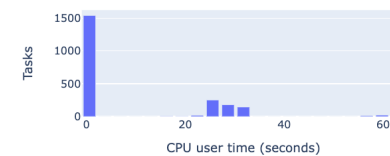- running
- exec_done
- launched
- pending
- failed
- fail_retryable

Resource usage

CPU Usage

CPU Time Distribution(avg) — Tasks vs CPU user time (seconds)

CPU Time Distribution(max) — Tasks vs CPU user time (seconds)

Memory usage

Memory usage — Memory (GB) vs Time (seconds)
- Used memory
- Total memory
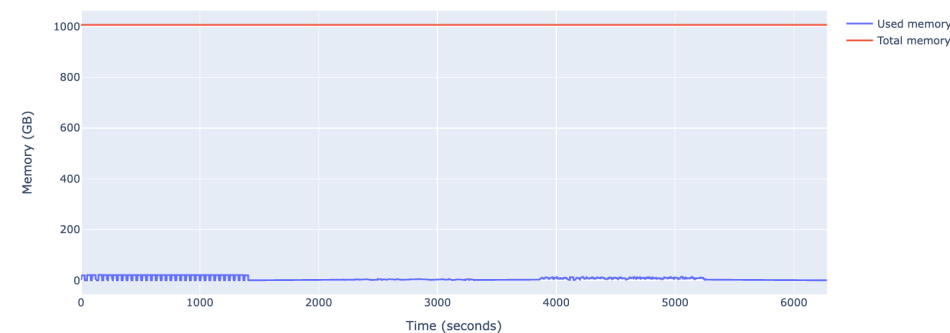
Memory Usage

Memory Distribution(avg) — Tasks vs Memory usage (bytes)

Memory Distribution(max) — Tasks vs Memory usage (bytes)
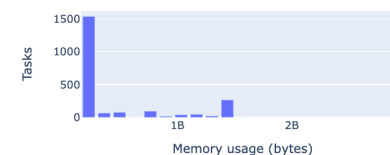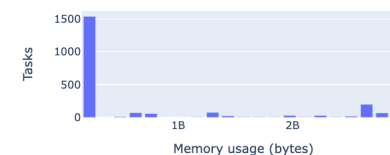
Looking around (2)

Pages grouped by task, with quanta ID and deps

subtractImages

- **Workflow name:** ApPipe-DC2-Test.DM-36026
- **Started:** 2022-10-25 10:33:22
- **Completed:** 2022-10-25 12:18:12
- **Workflow duration:** 1:44:49
- **Owner:** elhoward
- **Host:** sdfrome001
- **Run directory:** /sdf/data/rubin/user/elhoward/repo-main-logs/DM-36026/runinfo/001
- **Number of completed tasks:** 2174
- **Number of failed tasks:** 2

Name	Task ID	Dependencies	Completed
subtractImages	4	3 2	2022-10-25 11:38:17
subtractImages	12	11 10	2022-10-25 11:38:56
subtractImages	20	18 19	2022-10-25 11:38:27
subtractImages	28	26 27	2022-10-25 11:38:25
subtractImages	36	34 35	2022-10-25 11:39:18
subtractImages	44	43 42	2022-10-25 11:38:32
subtractImages	52	50 51	2022-10-25 11:38:27
subtractImages	60	58 59	2022-10-25 11:37:38
subtractImages	68	66 67	2022-10-25 11:39:38
subtractImages	76	75 74	2022-10-25 11:38:00
subtractImages	84	83 82	2022-10-25 11:38:54

Can also look at individual quanta

subtractImages (116)

- **Workflow name:** ApPipe-DC2-Test.DM-36026
- **Started:** 2022-10-25 10:33:22
- **Completed:** 2022-10-25 12:18:12
- **Workflow duration:** 1:44:49
- **Owner:** elhoward
- **task_func_name:** subtractImages
- **task_id:** 116
- **task_depends:** 115 114
- **task_time_invoked:** 2022-10-25 10:33:22
- **task_time_returned:** 2022-10-25 11:39:01
- **task_inputs:** [<AppFuture at 0x7fad1e072d10 state=pending>, <AppFuture at 0x7fad1e072950 state=pending>]
- **task_outputs:** None
- **task_stdin:** None
- **task_stdout:** /sdf/home/e/elhoward/u/repo-main-logs/DM-36026/submit/u/elhoward/DM-36026/DC2-4patch_4431/20221025T173121Z/logs/c7a855a0-de75-4f50-9f9b-f181184e830b_subtractImages_212077_124.stdout
- **task_stderr:** /sdf/home/e/elhoward/u/repo-main-logs/DM-36026/submit/u/elhoward/DM-36026/DC2-4patch_4431/20221025T173121Z/logs/c7a855a0-de75-4f50-9f9b-f181184e830b_subtractImages_212077_124.stderr

Task State

Time	State
2022-10-25 10:33:22	pending
2022-10-25 11:29:29	launched
2022-10-25 11:38:25	running
2022-10-25 11:39:01	exec_done

Memory Usage
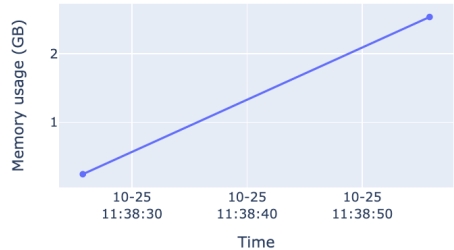
Final comments

- Seems useful as a stopgap solution while other tools (bps report, etc.) are being readied

- Some nice features, though slightly slower and harder to get the overall picture at a glance.

- Advantages of being able to drill down quickly to specific quanta

- All are welcome to use the ill-named conda env for trying it out, but might want to add the packages to the main stack if people are going to do this a lot in the long run
  - Need to re-initialize the Rubin env after doing conda deactivate when leaving the testconda env