

GCSC

11/17/2021

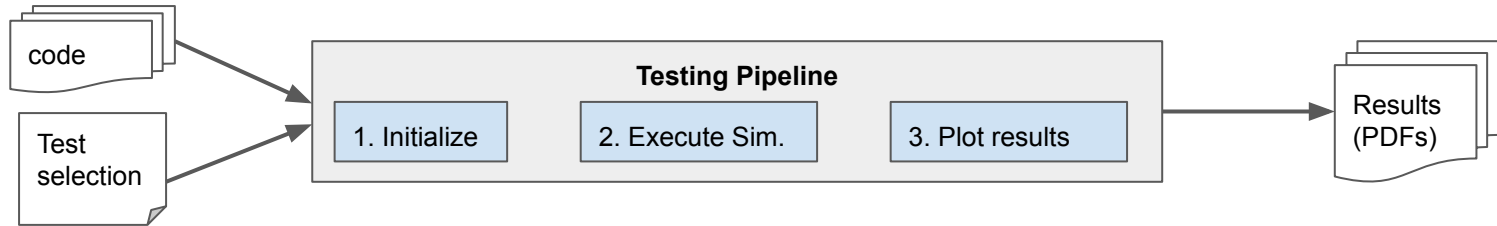
Tooling for the cloud

Objective

Streamline testing GEOS-Chem pull-requests and structural updates

Solution

Use automated pipelines on AWS to initiate, execute, and plot the results of test simulations.



What we need

1. Premade environments to execute the tests in (i.e., container images)
2. A means of synchronizing GEOS-Chem input data (i.e., keeping the data repository up-to-date)
3. The actual “pipeline infrastructure”

Pre-built test environments: EC2 Image Builder

What we need

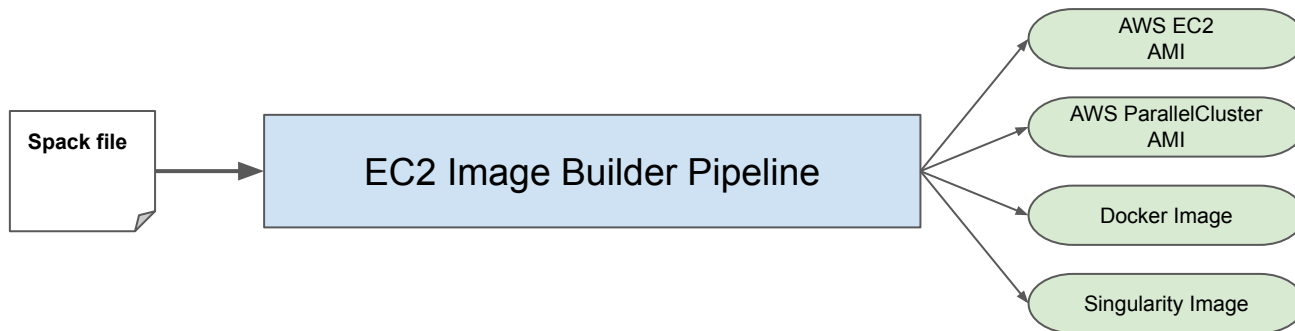
A set of container images for testing GEOS-Chem with different compilers and MPI.

Challenge

Maintaining more than a few images is a burden.

The Solution

We can use EC2 Image Builder to build images from a spack configuration file.



Synchronizing In

What we need

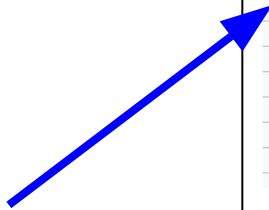
A way to automatically get new data

Solution

Use the bashdatacatalog to facilitate

Example **catalog**: EmissionsInputs-13.2.1.csv

	A	B	C	D
1	Path to collection	Collection URL	Enabled	Notes
2	HEMCO/ACET/v2014-07	http://geoschemdata.wustl.edu/ExtData/HEMCO/ACET/v2014-07		1
3	HEMCO/AEIC/v2015-01	http://geoschemdata.wustl.edu/ExtData/HEMCO/AEIC/v2015-01		1
4	HEMCO/AFCID/v2018-04	http://geoschemdata.wustl.edu/ExtData/HEMCO/AFCID/v2018-04		1
5	HEMCO/ALD2/v2017-03	http://geoschemdata.wustl.edu/ExtData/HEMCO/ALD2/v2017-03		1
6	HEMCO/AnnualScalar/v2014-07	http://geoschemdata.wustl.edu/ExtData/HEMCO/AnnualScalar/v2014-07		1
7	HEMCO/APEI/v2016-11	http://geoschemdata.wustl.edu/ExtData/HEMCO/APEI/v2016-11		1
8	HEMCO/BIOFUEL/v2019-08	http://geoschemdata.wustl.edu/ExtData/HEMCO/BIOFUEL/v2019-08		1
9	HEMCO/BROMINE/v2015-02	http://geoschemdata.wustl.edu/ExtData/HEMCO/BROMINE/v2015-02		1
10	HEMCO/C2H6_2010/v2019-06	http://geoschemdata.wustl.edu/ExtData/HEMCO/C2H6_2010/v2019-06		1
11	HEMCO/CEDS/v2021-06	http://geoschemdata.wustl.edu/ExtData/HEMCO/CEDS/v2021-06		1
12	HEMCO/CH3I/v2014-07	http://geoschemdata.wustl.edu/ExtData/HEMCO/CH3I/v2014-07		1
13	HEMCO/CMIP6/v2020-03	http://geoschemdata.wustl.edu/ExtData/HEMCO/CMIP6/v2020-03		1
14	HEMCO/DICE_Africa/v2016-10	http://geoschemdata.wustl.edu/ExtData/HEMCO/DICE_Africa/v2016-10		1
15	HEMCO/DMS/v2015-07	http://geoschemdata.wustl.edu/ExtData/HEMCO/DMS/v2015-07		1
16	HEMCO/DUST_DEAD/v2019-06	http://geoschemdata.wustl.edu/ExtData/HEMCO/DUST_DEAD/v2019-06		1
17	HEMCO/EDGARv42/v2015-02	http://geoschemdata.wustl.edu/ExtData/HEMCO/EDGARv42/v2015-02		1
18	HEMCO/EDGARv43/v2016-11	http://geoschemdata.wustl.edu/ExtData/HEMCO/EDGARv43/v2016-11		1
19	HEMCO/GEIA/v2014-07	http://geoschemdata.wustl.edu/ExtData/HEMCO/GEIA/v2014-07		1
20	HEMCO/GFED4/v2015-10	http://geoschemdata.wustl.edu/ExtData/HEMCO/GFED4/v2015-10		1
21	HEMCO/GFED4/v2020-02	http://geoschemdata.wustl.edu/ExtData/HEMCO/GFED4/v2020-02		1



Automating Simulations on AWS

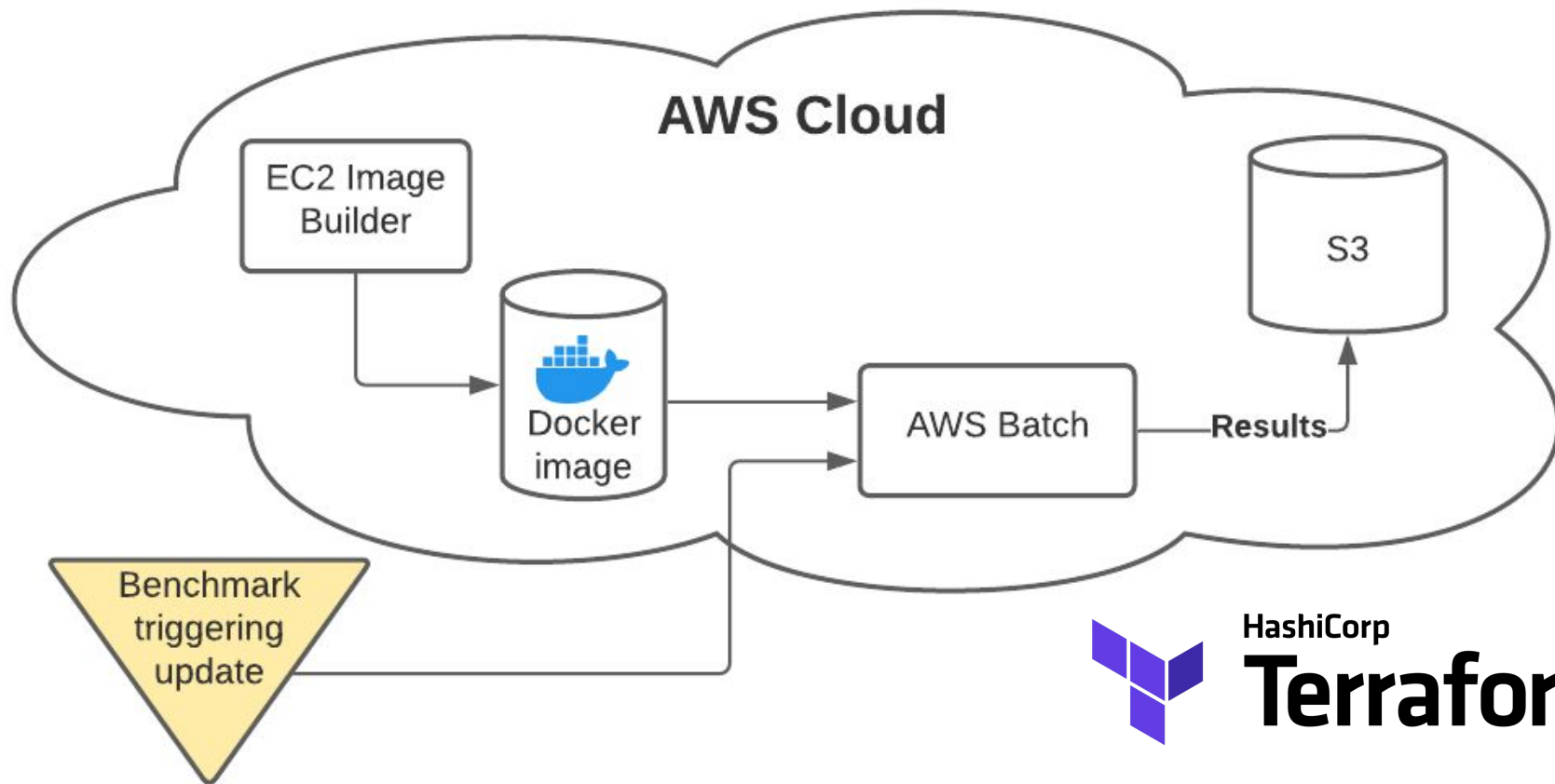
Goal:

- Creation of automated pipelines to trigger benchmarks runs for GCClassic and GCHP

Plan:

- Creation of runtime docker environment with EC2 image builder
- Management and deployment of EC2 instances with AWS Batch
- Plots automatically pushed to S3
- Management of AWS infrastructure with Terraform (IaC)

Pipeline Infrastructure: Proposed Architecture



GEOS-Chem Website

Updates that were applied:

- People and Projects map is back!
- Removed the Harvard-specific branding
- Greater prominence of citation and new developments information
- Adjusted mobile display of meetings menu
- Small cosmetic updates (font and coloring)

Thanks for your feedback and feel free to send more thoughts!