NextGen In A Box: NextGen Framework National Water Model Community Release

Unlock the potential for collaborative research

Arpita Patel – The University of Alabama

CUAHSI Summer Institute ‘23

THE UNIVERSITY OF ALABAMA | Alabama Water Institute

AGENDA

- Introduction – CIROH, Team
- NOAA-OWP's **ngen**
- **NextGen In A Box @ CIROH**
- **Technology** Stack Learning Session
- Hands-on **demo of NextGen In A Box (local machine)**
- **Conclusion**
- **Q&A**

ABOUT CIROH

- The Cooperative Institute for Research to Operations in Hydrology (CIROH) is a **consortium of 28 institutions** to advance the National Oceanic and Atmospheric Administration's (**NOAA**) science and services capabilities to provide actionable water resources intelligence.
- CIROH's research aims to **improve water prediction** and supports **four broad themes**:

  (1) **water** resources **prediction capabilities**;

  (2) community **water** resources **modeling**;

  (3) **hydroinformatics**; and

  (4) application of **social**, economic and **behavioral science to water resources prediction.**

- CIROH outcomes will inform hydrological process understanding, operational forecasting techniques and workflows, community engagement in water modeling, open-source software development, translation of forecasts to actionable products, and use of predictions in decision making.

OUR TEAM

CIROH TEAM

The University of Alabama (UA) Team:

- **Dr. Steve Burian -** Executive Director of CIROH, Director of Science, AWI
- **James Halgren -** Assistant Director of Science
- **Dr. Puri Bangalore -** Professor in the Department of Computer Science
- **Jeff Carver -** Professor in the Department of Computer Science
- **Arpita Patel -** DevOps Manager & Enterprise Architect
- **Benjamin Lee -** DevOps Engineer
- **Shahab Alam** - Postdoc
- **Sepehr Karimi** – Software Research Engineer, Data Scientist
- **Hari Jajula –** Student
- **Rohan Sunkarapalli –** Student

Lynker Team:

- **Zach Wills -** Software Engineer
- **Nels Frazier -** Lead Technical Architect
- **Jordan Laser -** Software Engineer
- **Mike Johnson –** Hydrofabric Lead

RESEARCH TO OPERATIONS PATHWAY

NEXTGEN

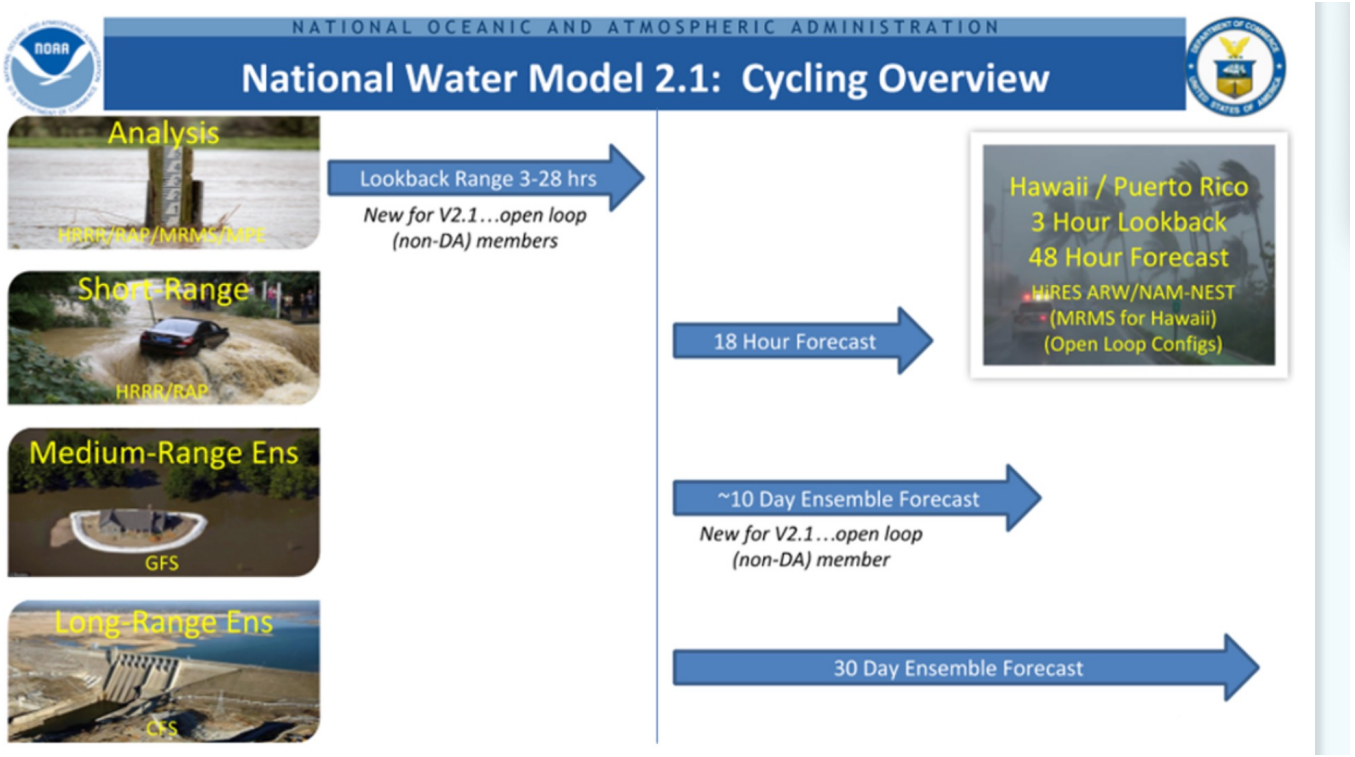
- NEXTGEN- HYDRO FABRIC
- NEXTGEN- CATCHMENTS, NEXUSES
- NEXTGEN– REALIZATION

WHAT IS THE NATIONAL WATER MODEL?

- The National Water Model (NWM) is a **hydrologic model developed by** the National Oceanic and Atmospheric Administration (**NOAA**) to simulate and forecast water conditions across the United States.
- This water prediction model creates forecast guidance for over 3.4 million miles of rivers and streams across the United States and its territories.
- The NWM supercomputer is fed nonstop data covering everything from current stream-flow to the snowpack in mountain ranges above.
- The job of the supercomputer is to take all of that data and produce stream flow forecasts for every stream and river in the continental United States and its territories. There are three different flavors of forecasts: short, medium, and long range.
- Just like how your local weather station can tell you what the weather will be like tomorrow, in 3 days, or a week from now; the NWM can tell you what the stream flow will be like in 18 hours, 10 days, or even up to 30 days from now.
- NOAA Central Library Seminar Recording from OWP (Next-Generation Water Resources Modeling): https://youtu.be/DLIi3PruYxo

NWM 2.1

- The NWM currently runs in four configurations:
  - **Analysis and assimilation provides** a snapshot of current hydrologic conditions
  - **Short-Range** produces hourly forecasts of streamflow and hydrologic states out to 15 hours
  - **Medium-Rang**e produces 3-hourly forecasts out to 10 days
  - **Long-Range** generates 30-day ensemble forecasts.



Source : https://water.noaa.gov/about/nwm

SO...WHAT IS NEXTGEN?

NextGen is a framework designed for building and integrating models rather than a model itself.

- Model agnostic!

The framework's focus is on a data-centric process that abstracts the addition of processes and data behind a standard.

- Data-centric standards

NextGen design allows for greater flexibility and standardization in model creation and integration.

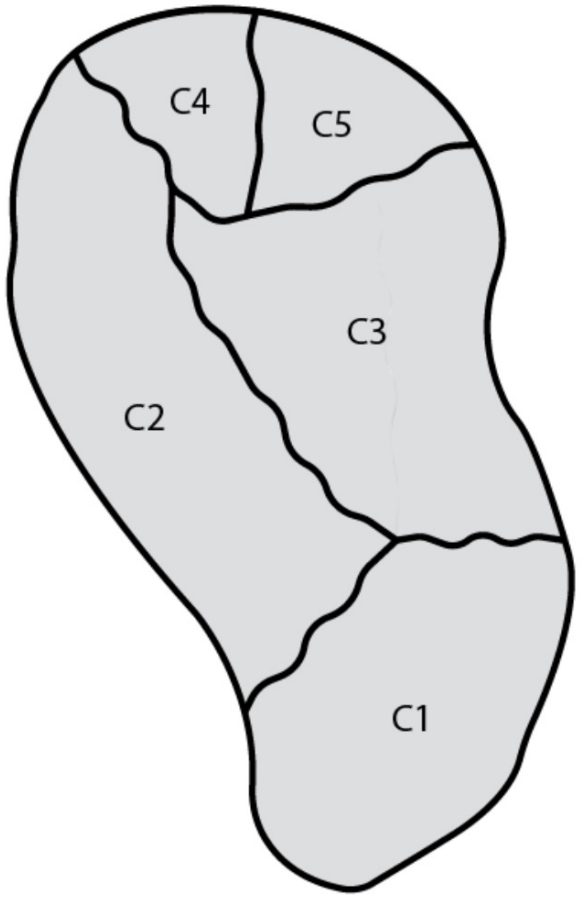
Github URL: https://github.com/NOAA-OWP/ngen

NEXTGEN CATCHMENT

- A catchment is a geographical area in which water drains into a common watercourse, such as a river or lake.

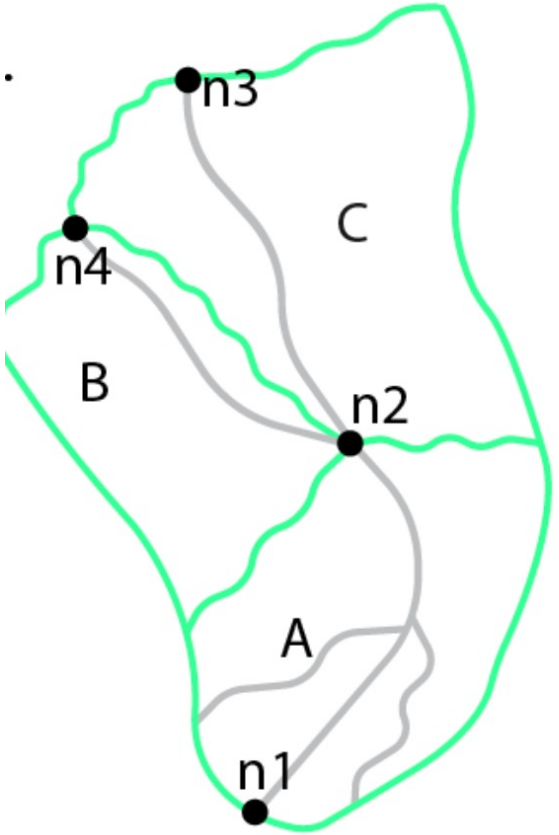
- The NextGen Hydrofabric divides the landscape into catchments, each of which is configured and parameterized then simulated independently by the model.

- More info: https://github.com/NOAA-OWP/ngen/blob/master/data/README.md

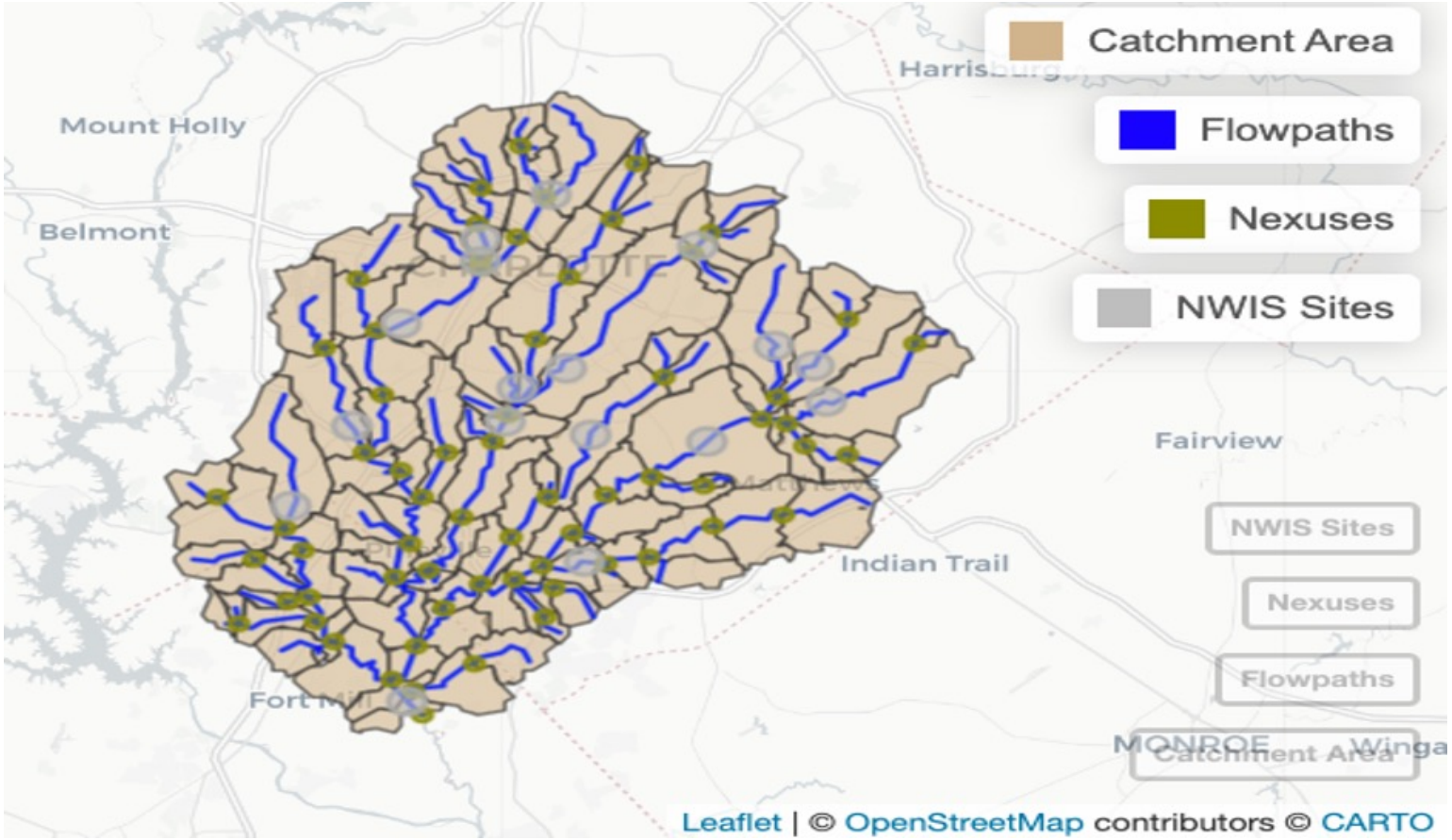


NEXTGEN NEXUSES

- The nexus is the connection between the surface and subsurface water systems within a catchment.
- It includes the movement of water through soil and rock, as well as the interactions between groundwater and surface water.
- The NextGen framework simulates these processes to estimate the amount of water at a nexus.
- It is a location where one or more nested models or groups of models are integrated.
- Nexuses are also planned be placed at hydrologically significant locations, (like at a known gauge) to output data for easier comparison.



CATCHMENT AND NEXUSES



NEXTGEN HYDROFABRIC

FILE REPRESENTATION OF THE PHYSICAL LAND AND FLOW PATHS OF WATER IN THE CONUS

https://noaa-owp.github.io/hydrofabric/

The hydrofabric is a representation of the physical land and flow paths of water in the contiguous United States (CONUS).

Its files are the first argument to run ngen model.

URL: https://noaa-owp.github.io/hydrofabric/

S3 bucket: https://nextgen-hydrofabric.s3.amazonaws.com/index.html

REALIZATION FILES (CONFIG FILE)

```
{
    "global": {
        "formulations": [
            {
                "name": "bmi_multi",
                "params": {
                    "name": "bmi_multi",
                    "model_type_name": "NoahOWP_CFE",
                    "main_output_variable": "Q_OUT",
                    "init_config": "",
                    "allow_exceed_end_time": false,
                    "fixed_time_step": false,
                    "uses_forcing_file": false,
                    "modules": [
                        {
                            "name": "bmi_c",
                            "params": {
                                "name": "bmi_c",
                                "model_type_name": "CFE",
                                "main_output_variable": "Q_OUT",
                                "init_config": "/ngen/AWI_001/config/awi_config.ini",
                                "allow_exceed_end_time": true,
                                "fixed_time_step": false,
                                "uses_forcing_file": false,
                                "variables_names_map": {
                                    "atmosphere_water__liquid_equivalent_precipitation_rate": "precip_rate",
                                    "water_potential_evaporation_flux": "EVAPOTRANS",
                                    "ice_fraction_schaake": "sloth_ice_fraction_schaake",
                                    "ice_fraction_xinan": "sloth_ice_fraction_xinan",
                                    "soil_moisture_profile": "sloth_smp"
                                },
                                "model_params": {
                                    "b": 8.660529385231255,
                                    "satdk": 0.00011760880965802808,
                                    "maxsmc": 0.543673362985325,
                                    "refkdt": 3.661344050458613 4,
                                    "slope": 0.8154788969461678,
                                    "max_gw_storage": 0.040219944149233 59,
                                    "expon": 7.308820146231674,
                                    "Cgw": 0.0004609207383395736,
                                    "Klf": 0.1681695665829872,
                                    "Kn": 0.4017865685354076
                                },
```

- A realization is a set of inputs to the NWM that are used to simulate the behavior of a catchment. This includes meteorological data, such as precipitation and temperature, as well as information about the land surface, such as soil moisture and vegetation cover. The NWM uses realizations to forecast the water conditions in a catchment over time.

Realizations provide a description of:

the model(s) that run in catchments

Model parameters

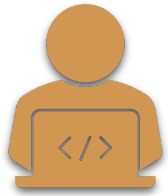
Associated forcings

Model libraries

Realization files offer a powerful interface:

For automating model generation and calibration using machine-readable file formats

which provide ability to support multi-model implementations, routing changes, and parameter variations.



Realization makes the process more efficient and streamlined.

The JSON format allows for systematic generation and calibration

Also allows for customizable interfaces and metadata collection

```
2022-08-24 21:00:00,0.0,366.85886363636365,515.3281818181817,99616.97272727272,294.55954545454534,-1.4730909090909092,-2.742090909090909,0.0
2022-08-24 22:00:00,0.0,369.80590909090915,346.70995454545454,99594.38181818184,294.9004545454545,-2.3777272727272734,-2.3822727272727273,0.0
2022-08-24 23:00:00,5.465763e-07,371.1136363636363,174.92749999999998,99608.5090909091,295.19363636363636,-1.9541818181818185,-2.111181818181819,5.465763e-07
2022-08-25 00:00:00,6.7041565e-06,370.8944545454545,27.839636363636362,99634.64090909091,295.3190909090909,-0.7035,-2.5654090909090908,6.7041565e-06
2022-08-25 01:00:00,8.889083e-05,373.98027272727273,0.0,99668.03636363636,295.04681818181814,-0.05204545454545454,-2.0010909090909093,8.889083e-05
2022-08-25 02:00:00,6.747035e-05,404.54200000000003,0.0,99688.96363636364,295.08454545454555,-0.968909090909091,-0.9149090909090912,6.747035e-05
2022-08-25 03:00:00,2.7146058e-07,387.6152727272727,0.0,99682.16363636364,294.980909090909,-1.3147272727272727,1.1156818181818182,2.7146058e-07
2022-08-25 04:00:00,2.0956728e-07,405.59327272727273,0.0,99679.1,294.8190909090909,-1.3546363636363634,1.233409090909091,2.0956728e-07
2022-08-25 05:00:00,4.5842498e-08,377.20518181818176,0.0,99692.15454545454,294.6122727272727,-1.1915,1.1515454545454544,4.5842498e-08
2022-08-25 06:00:00,4.2404034e-07,364.4026363636363,0.0,99692.88636363637,294.4031818181818,-1.0278636363636362,0.20149999999999998,4.2404034e-07
2022-08-25 07:00:00,0.0,362.955,0.0,99655.15909090909,294.1154545454545,-1.1871818181818183,-0.2055454545454545,0.0
2022-08-25 08:00:00,4.348889e-07,361.8862272727273,0.0,99638.67272727274,293.834090909091,-1.486181818181818,1.2407727272727274,4.348889e-07
2022-08-25 09:00:00,2.1378277e-07,363.37718181818184,0.0,99628.05909090908,293.4740909090908,-1.6439545454545454,1.550909090909091,2.1378277e-07
2022-08-25 10:00:00,1.0461959e-09,372.4043181818181,0.0,99611.38181818184,293.11227272727274,-1.4667272727272727,2.35,1.0461959e-09
2022-08-25 11:00:00,1.06891385e-07,373.8485,8.853818181818182,99626.06363636366,292.8454545454545,-1.2619090909090909,2.775363636363636,1.06891385e-07
2022-08-25 12:00:00,1.06891385e-07,377.8324090909091,89.28663636363636,99651.32727272726,292.62409090909097,-1.1993636363636364,2.8759545454545457,1.06891385e-07
2022-08-25 13:00:00,1.9111683e-06,377.84145454545455,224.03386363636366,99667.52272727272,292.4804545454545,-1.3387272727272725,2.5170909090909093,1.9111683e-06
2022-08-25 14:00:00,2.8642868e-05,368.8635,448.12922727272723,99674.15454545456,292.33909090909094,-0.992545454545454,2.269545454545454,2.8642868e-05
2022-08-25 15:00:00,1.7014056e-05,371.45863636363634,579.4265454545455,99670.86818181818,292.4509090909091,-1.3245,1.785500000000003,1.7014056e-05
2022-08-25 16:00:00,5.9881e-06,390.7098181818182,489.1470909090909,99656.4,292.5909090909091,-0.2307272727272727,1.7757272727272722,5.9881e-06
2022-08-25 17:00:00,6.974531e-06,389.9126363636364,569.3464545454545,99598.58636363636,292.8409090909091,0.061454545454545456,1.4741363636363638,6.974531e-06
2022-08-25 18:00:00,7.915695e-05,387.3512727272727,575.4722727272728,99570.09545454546,293.02318181818185,1.437090909090909,1.698363636363637,7.915695e-05
2022-08-25 19:00:00,5.2005063e-05,400.9825909090909,329.6197272727273,99490.45909090912,293.45409090909095,0.234,-0.8528181818181817,5.2005063e-05
2022-08-25 20:00:00,0.00046328641,405.85604545454555,119.39481818181182,99471.23636363637,294.23545454545456,1.8572727272727276,-2.158181818181818,0.00046328641
2022-08-25 21:00:00,2.3808841e-05,403.6048181818183,43.77959090909092,99448.34545454546,294.26045454545454,0.1495,-0.3169090909090909,2.3808841e-05
```

NEXTGEN FORCING

Per Catchment:

cat03w_cat-106885.csv

Link to forcing files in S3 bucket:

S3-bucket-forcing-files

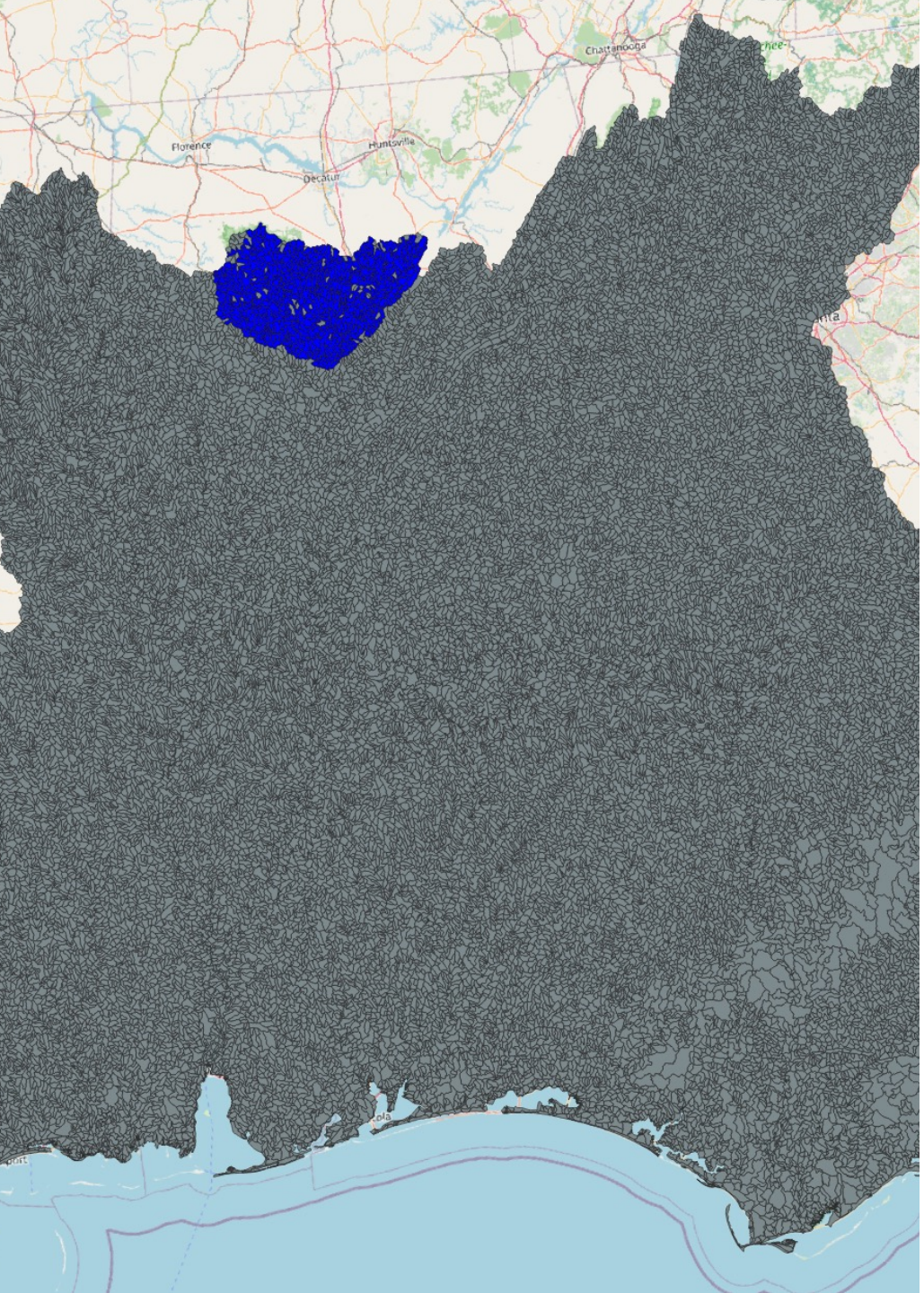
ADDITIONAL NWM INFORMATION

- Existing National Water Model basics course on HydroLearn:

https://edx.hydrolearn.org/courses/course-v1:BYU+NWM101+2021/about

NEXTGEN IN A BOX

- USE CASE
- OVERVIEW OF NEXTGEN
- OBJECTIVES AND BENEFITS OF NEXTGEN
- KEY FEATURES



USE CASE

- We are going to do a case study NWM run for Sipsey Fork, Black Warrior river
  - We don't want to run all of CONUS
  - We want to run NextGen locally
  - We want to have control over inputs / config.
  - How can we do it? Answer: NextGen-In-A-Box
- Finally, - Generate outputs of the NextGen
  - NextGen currently outputs CSV at each catchment & nexus

OVERVIEW OF NEXTGEN IN A BOX COMMUNITY WATER MODEL FOR NWM

- Introducing the Nextgen In A Box community model for NWM to Academic community.

- Encouraging Academic researchers and developers to consider adopting this model in their own projects

OBJECTIVES AND BENEFITS OF THE COMMUNITY MODEL

To foster collaboration among academic researchers and developers in the Github repo community

This model provides a platform for individuals to share ideas, resources, and knowledge to advance research and development in the NWM.This model provides a platform for individuals to share ideas, resources, and knowledge to advance research and development in the NWM.

To enable rapid and efficient development of NWM repo-related tools and technologies

The model leverages the collective skills, expertise, and resources of the community to accelerate the development and deployment of NWM repo-related tools and technologies.

To promote open and transparent research practices

The model encourages open sharing of data, methods, and results to facilitate reproducibility and collaboration in the NWM repo field.

To facilitate adoption of NWM repo-related tools and technologies by the wider academic community

The model promotes the dissemination and adoption of NWM repo-related tools and technologies to enhance research efficiency and impact.

KEY FEATURES OF THE NEXTGEN IN A BOX COMMUNITY MODEL FOR NWM

Open and collaborative development process for all community members

Ensures that all members have a voice and can contribute to the development of the model

Flexible and modular structure that allows for customization

Enables users to adapt the NextGen workflow to fit their specific research and development needs

Reusable components that can be shared across multiple projects

Saves time and effort by avoiding the duplication of work and promoting knowledge sharing

Dedicated community support and resources

Provides documentation and guidance to users throughout the implementation and development process

TECHNICAL STACK: NEXTGEN IN A BOX

- OPENSOURCE TECHNOLOGY

“NEXTGEN IN A BOX”- TECHNOLOGY STACK

OPENSOURCE TECHNOLOGIES

Git – Distributed version control system that allows developers to track changes to their code over time and collaborate with other team members on code development.

Github – web-based central hub for Git repositories.

Github Actions - platform for automating software workflows, such as building, testing, and deploying code, directly within the GitHub repository

Docker – platform that allows developers to create, deploy and run applications in containers.

Terraform – Infrastructure as code tool that allows developers to provision and manage cloud infrastructure using declarative code.

Cloud Infrastructure, Storage – Collection of hardware and software components that are used to provide cloud computing service, such as VM, storage and networking.

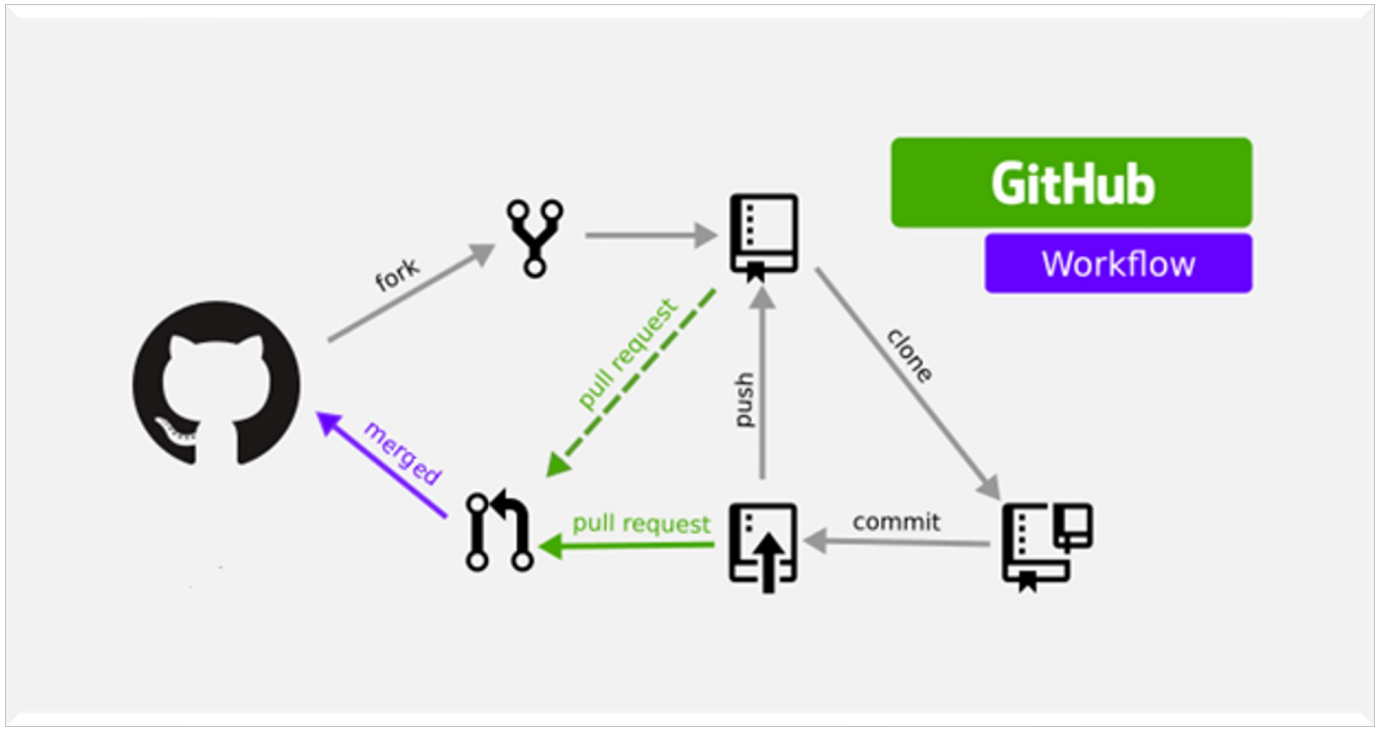
GIT BASICS!

- Git is a **version control system** for tracking the changes to the files and for **collaborating** work on those files among multiple people.
- Some basic **GIT commands** are:

🔷 **git init** -> Create a new git repo in the directory

🔷 **git clone**-> Clone the repository

🔷 **git branch** -> Create a new local branch
🔷 **git checkout** -> Switch branches
🔷 **git add** -> Add a new file to your staging area
🔷 **git commit** -> Adds staged changes to your local repository
🔷 **git pull** -> pull code from your remote repo to your local directory
🔷 **git push** -> Push local repository changes to your remote repo
🔷 **git status** -> Show which files are being tracked (and untracked)
🔷 **git diff** -> See the actual difference in code between your Working Directory and your Staging Area



GITHUB

- GitHub is a web-based **platform** for hosting and sharing Git repositories
- GitHub makes it easier for teams to **collaborate** on code development, even when working remotely
- GitHub provides **version control** and change tracking, so developers can easily see what changes have been made to the code over time
- GitHub provides a platform for sharing and discovering **open source** projects
- **Key features:**
  - Code hosting and version control
  - Collaboration tools, such as pull requests and code reviews
  - **Github Issues & Github projects** - Issue tracking and project management
  - Continuous integration and deployment **(CI/CD) with GitHub Actions**

NEXTGEN IN A BOX : GITHUB REPO

Repository location:

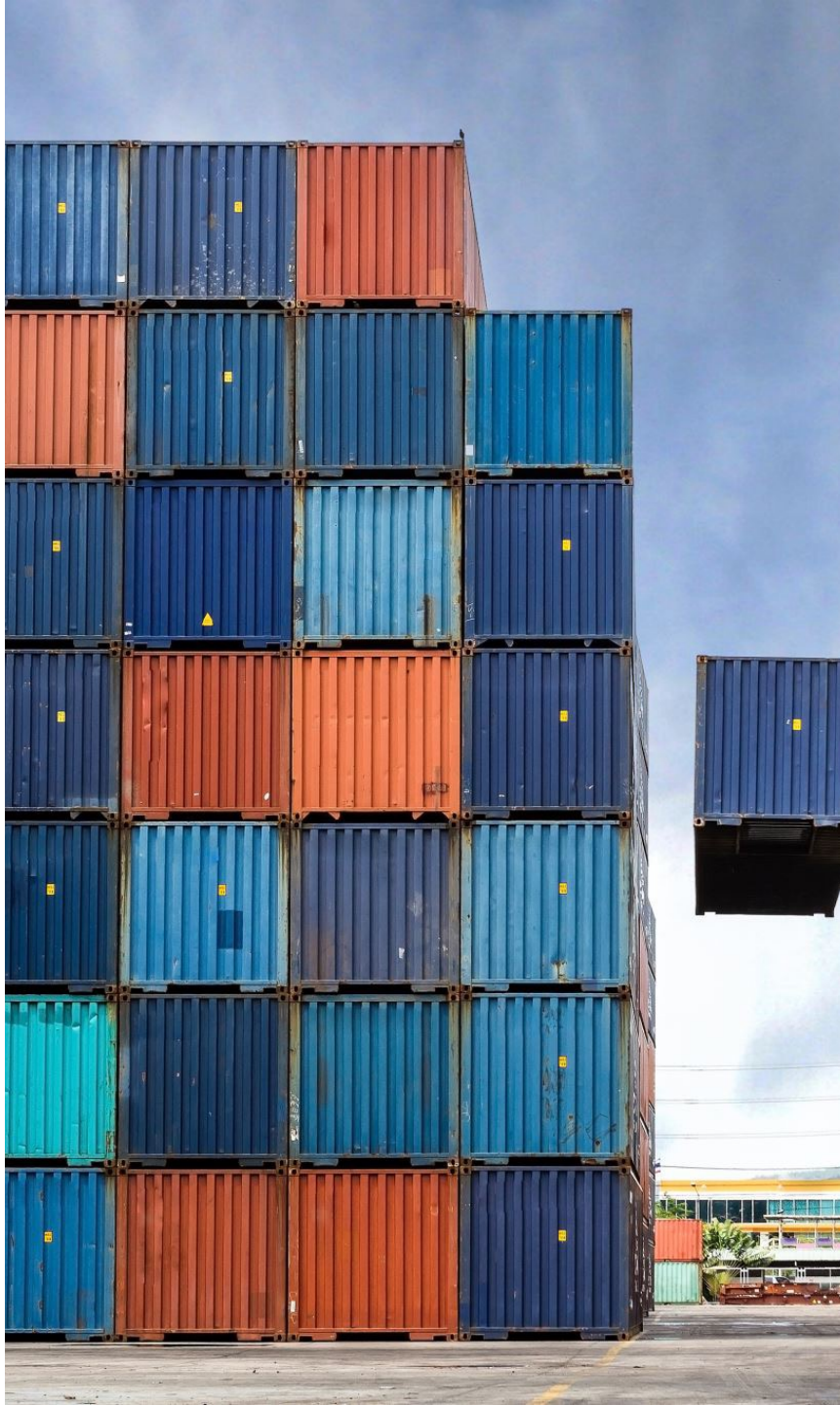
https://github.com/CIROH-UA/CloudInfra

Contains:

Dockerfiles
Terraform scripts
Github actions yaml files
Shell Scripts

DOCKER

- Standard unit of software
- **Packages code and dependencies**
- Can be shared as Docker **Images**
- **Multiple** containers can be run simultaneously
- **Portable**
- **Lightweight** - Uses the host operating system
- **Secure** - Strong default isolation features
- Sometimes used in addition to VMs



WHY USE DOCKER?

- Develop **faster**
- Develop applications that work on **any OS**
- Easy to **share** applications among teams
- Easy to **scale** across multiple servers
- Large applications can be broken into **multiple** containers - one for each microservice
- Great solution for **Cloud Computing**
- Big community and **library** of Docker Images

DOCKERFILE COMMANDS

- **FROM**: Specifies the base image for the Docker image you're building.
- **RUN**: Runs a command in the Docker image.
- **COPY**: Copies files from your local machine to the Docker image.
- **WORKDIR**: Sets the working directory for the Docker image.
- **EXPOSE**: Specifies which ports should be exposed by the Docker image.
- **CMD**: Specifies the command that should be run when the Docker container is started.

DOCKER COMMANDS:

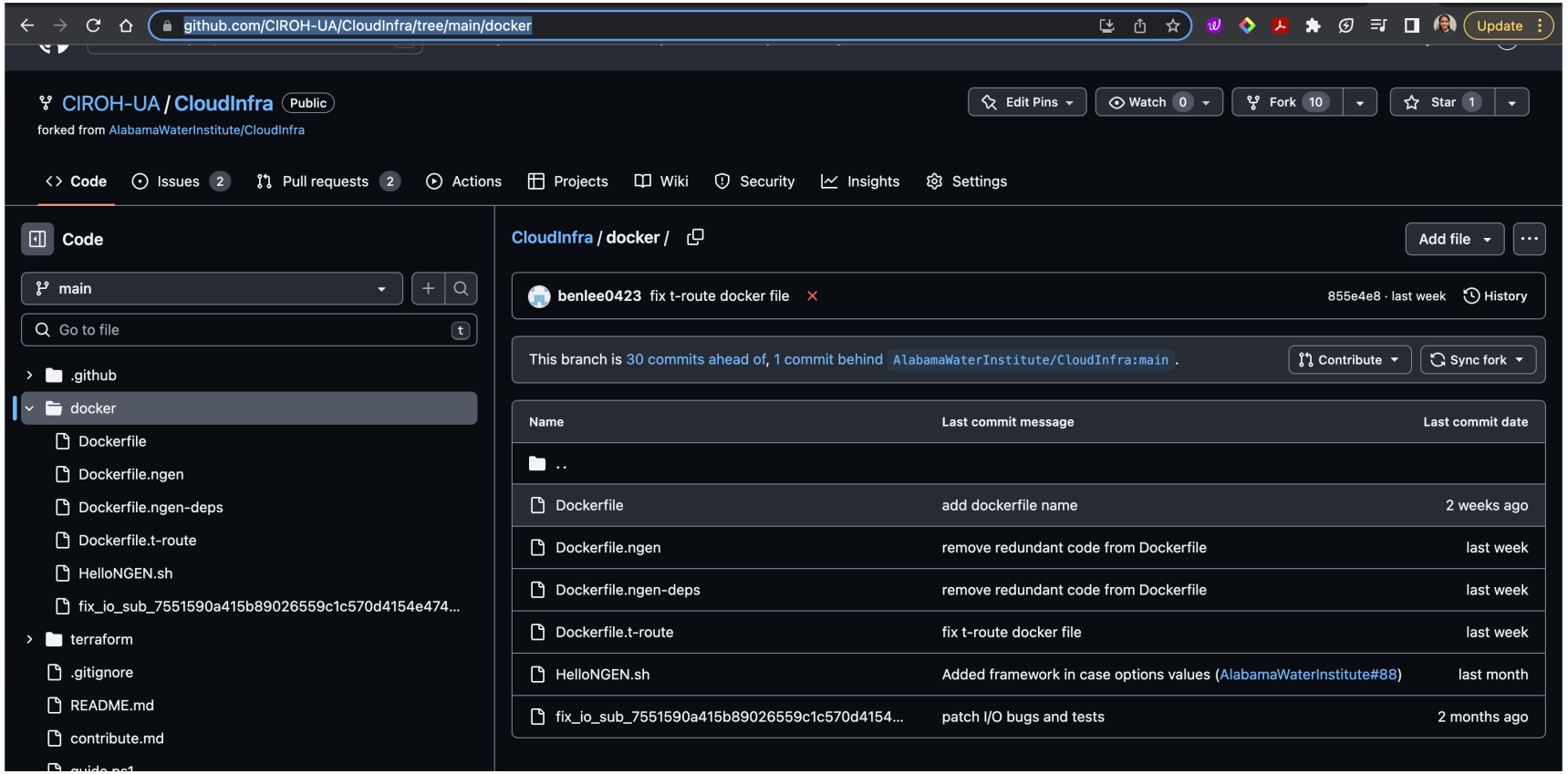
- **docker build:** Builds a Docker image from a Dockerfile.
- **docker run:** Runs a Docker container from a Docker image.
- **docker ps:** Lists the running Docker containers.
- **docker stop**: Stops a running Docker container.
- **docker rm:** Deletes a stopped Docker container.
- **docker images:** Lists the Docker images on your machine.
- **docker rmi:** Deletes a Docker image.

GETTING STARTED WITH DOCKER

- To get started with Docker, you need to install the **Docker engine** on your machine
- Docker provides a range of c**ommand-line tools and APIs** for working with containers and images
- Docker also provides a range of services, such as **Docker Hub** and **Docker Store,** for finding and sharing container images

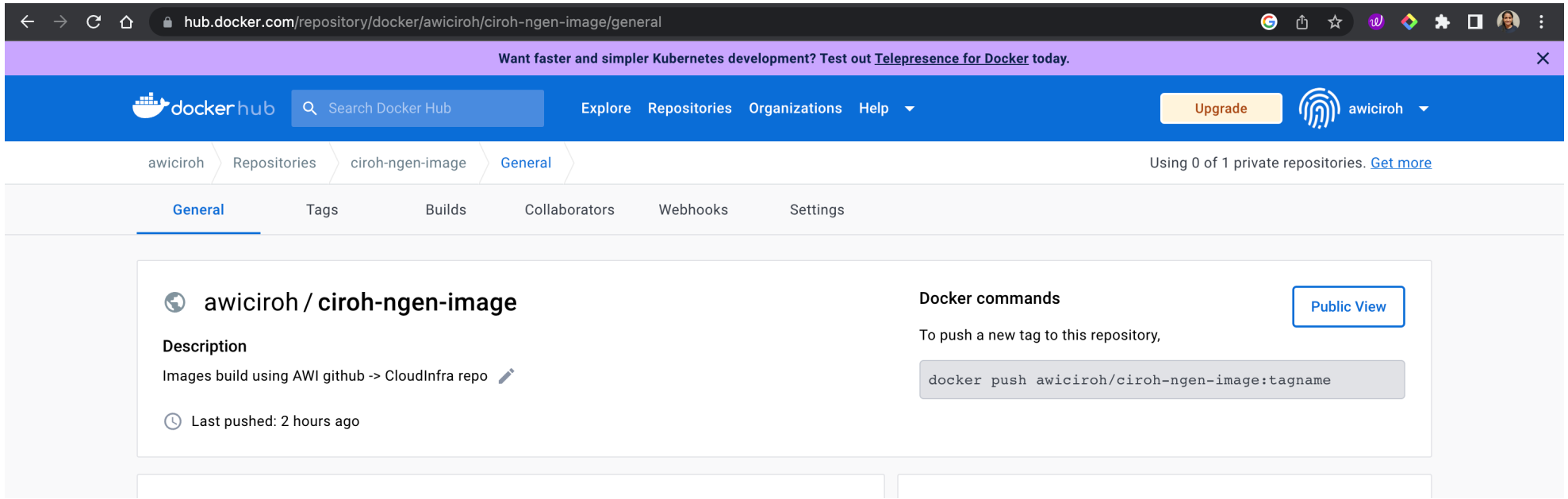
NGEN DOCKERFILE

- NextGen Dockerfile is available at
- https://github.com/CIROH-UA/CloudInfra/tree/main/docker



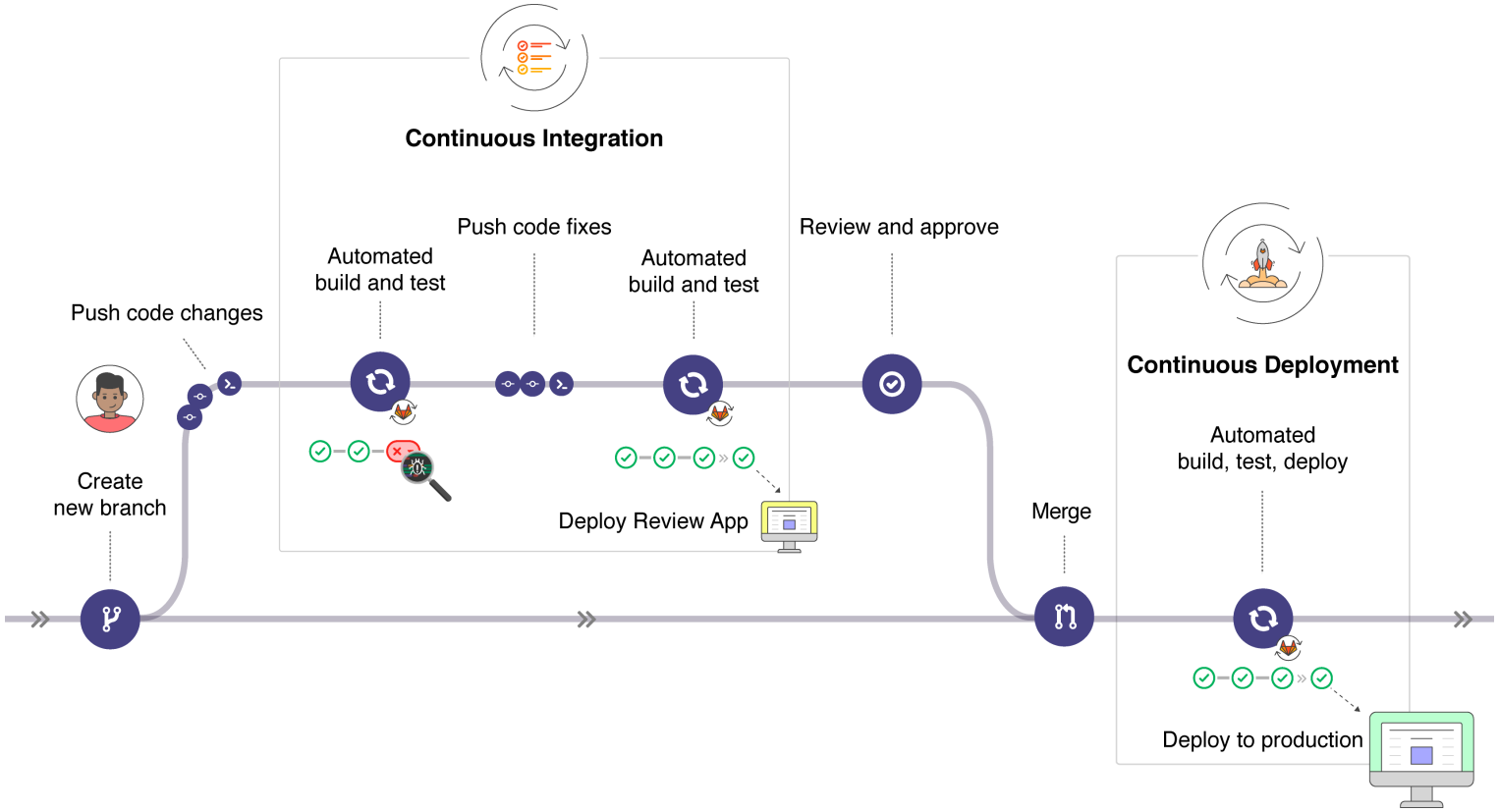
NEXTGEN IN A BOX: DOCKER HUB

- **Dockerhub account** : awiciroh
- **Image name:** ciroh-ngen-image
- **Image tags:** latest_x86, latest_arm

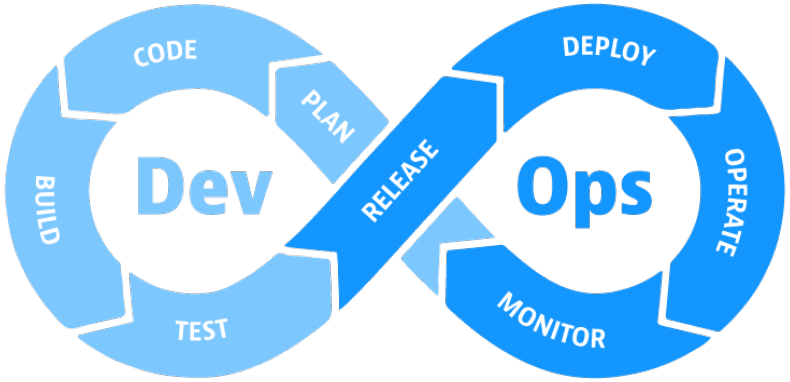


WHAT IS CI/CD?

- Continuous Integration & Continuous Deployment is an **automated process** for continuously integrating code changes in a small interval, testing them and releasing them out to the community.



WHAT IS DEVOPS?

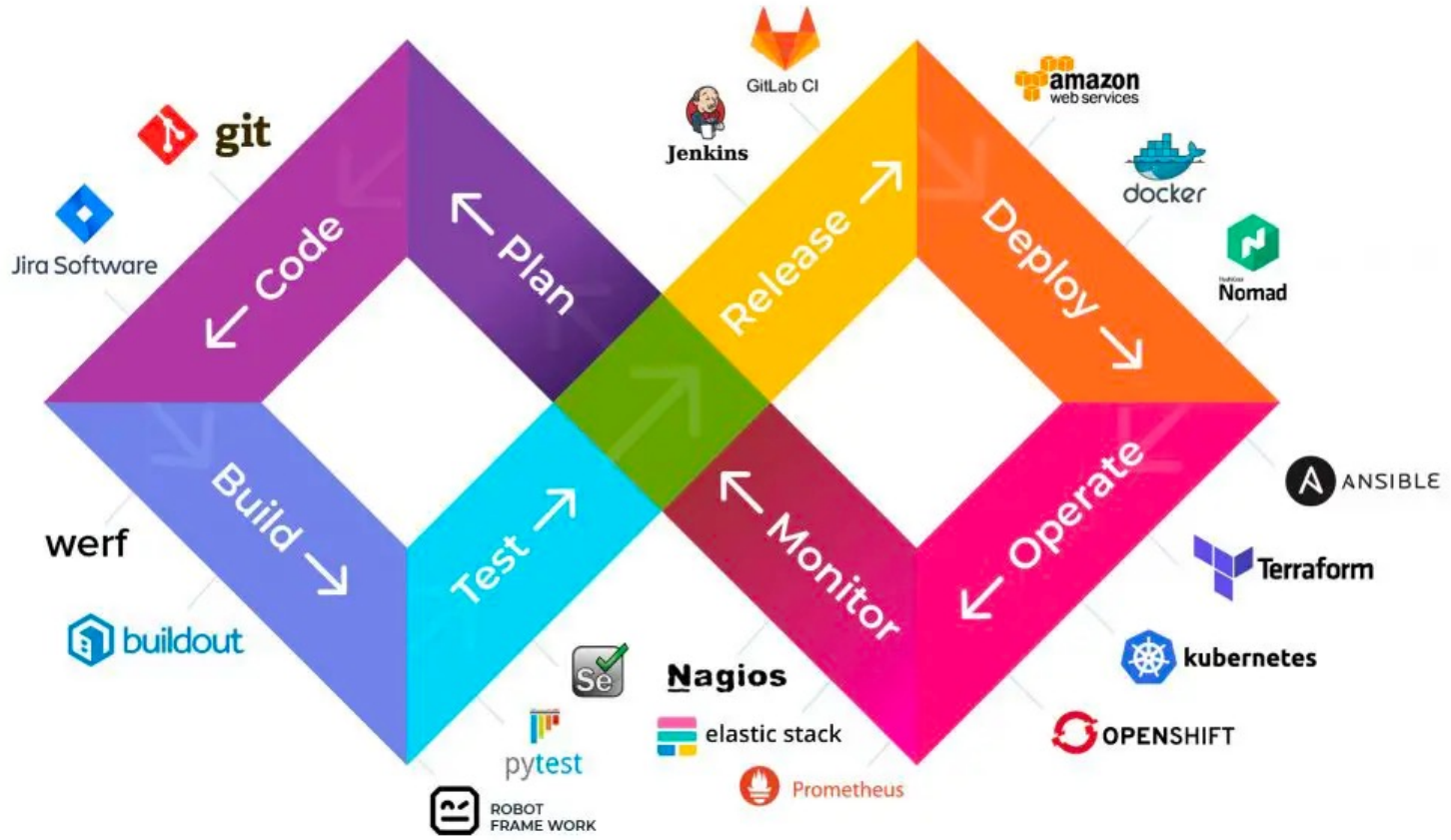


DevOps is a combination of software development (dev) and operations (ops).
It is defined as a software engineering methodology which aims to integrate the work of development teams and operations teams by facilitating a culture of collaboration and shared responsibility.

- DevOps is the union of **people**, **process**, and **products** to enable continuous delivery of value to your end users.
- Let's see life before DevOps and after DevOps using pit stops example: (DevOps analogy)
- Some key benefits of using DevOps -
  - Better tasks & backlog tracking, **Reduce overall cost**
  - Continuous Integration, Delivery & Deployments **(CI/CD)**
  - Release **faster** and work smarter
  - Culture of **Collaboration**
  - **Scalability**
  - **Security** and monitoring (**DevSecOps**)

DEVOPS CONT.

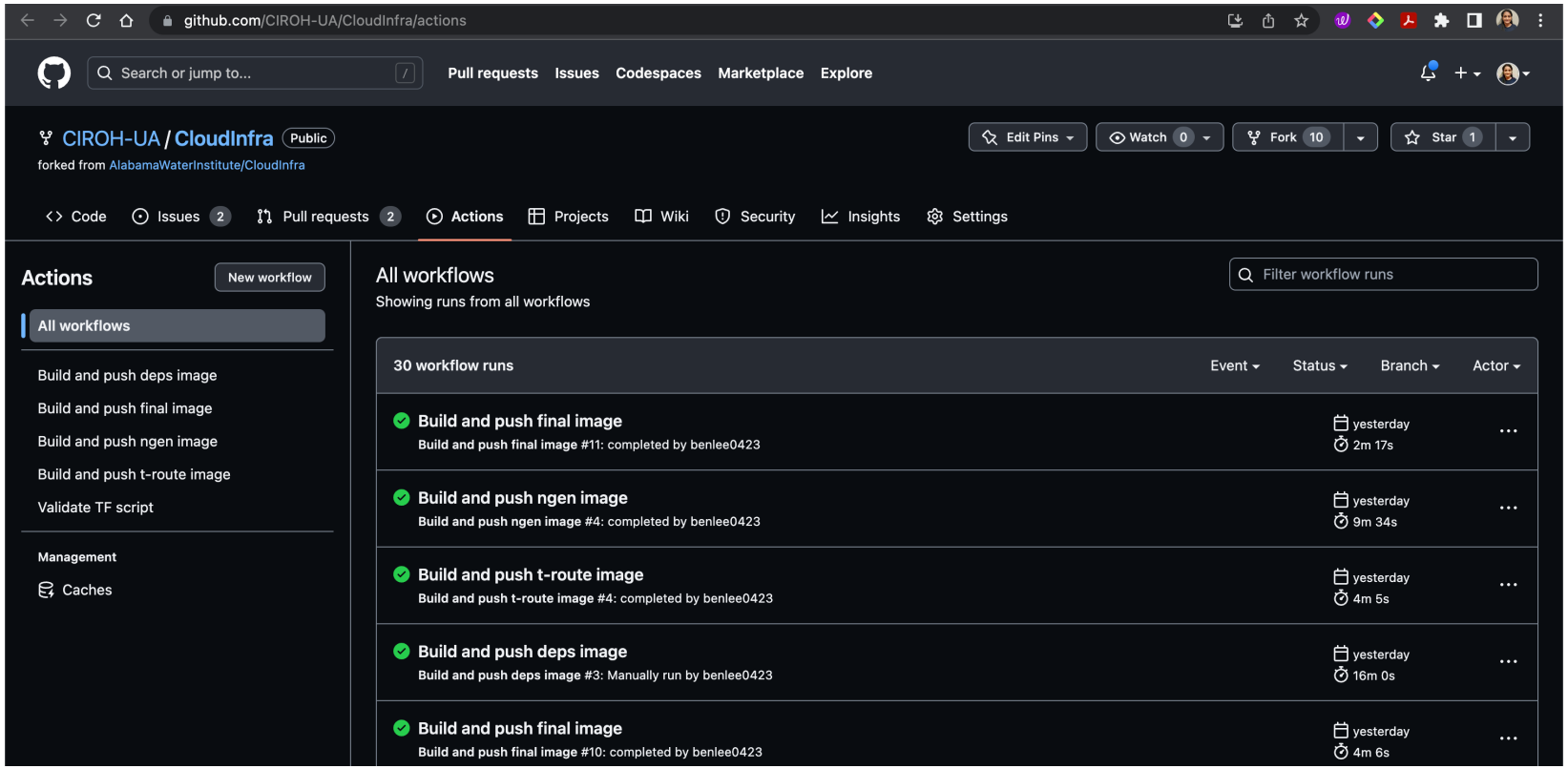
- Azure DevOps
- Git, Github, Github Issues
- Github Actions
- Terraform / K8s / ARM templates
- Docker
- AWS
- Prometheus / Grafana



NEXTGEN IN A BOX : CI WORKFLOW

- Using **Github Actions** to run workflows.
- Automated flow to builds and push the docker image to Dockerhub once Pull Request is merged.
  - On Pull Request, triggers action workflow and validates changes.
  - On merge to main branch, triggers action, builds docker image and pushes to Dockerhub.

CI WORKFLOWS:



HANDS-ON DEMO

- RUNNING ON LOCAL MACHINE USING DOCKER

PREREQUISITES

- Make sure Prerequisites are installed on your local machine:
  - git
  - Docker (installed and started)

  Make sure below is working in your instance:

  $ docker ps

  - Make sure input data is downloaded.

NEXTGEN IN A BOX: README.MD FILE

Open README file:

https://github.com/CIROH-UA/CloudInfra#readme

3 Steps:

1. Install docker
2. Download sample data
3. Run NextGen In A Box – guide.sh

STEP 1 : INSTALL DOCKER

On Windows:
Install Docker Desktop on Windows : https://docs.docker.com/desktop/install/windows-install/#install-docker-desktop-on-windows

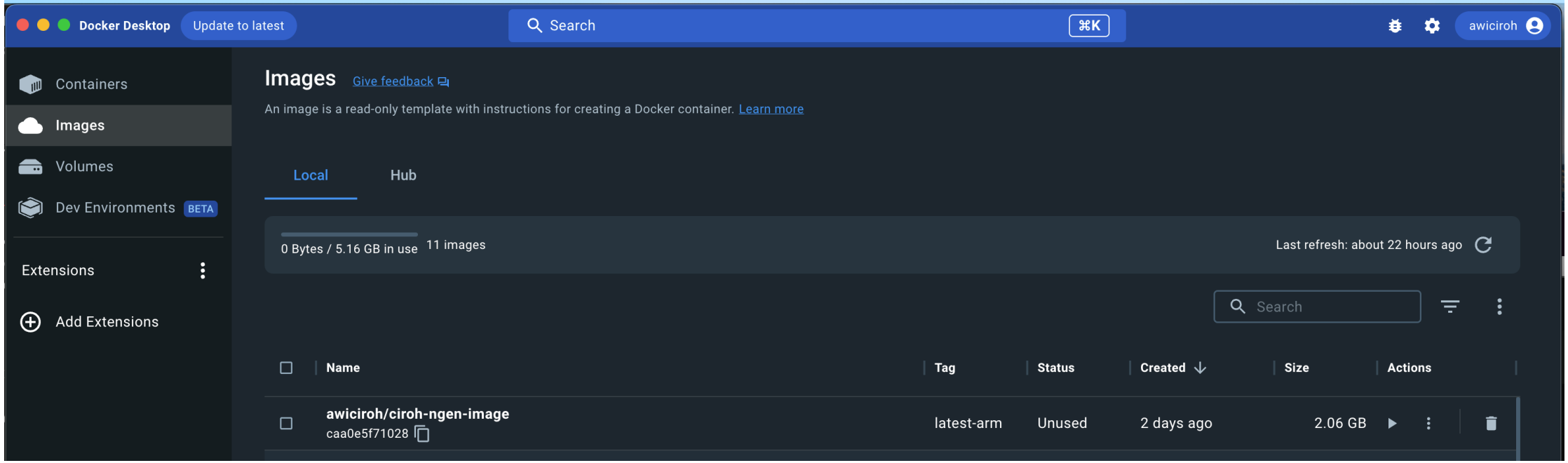
- Once docker is installed, start **Docker Desktop.**
- Open **powershell** -> right click and `Run as an Administrator`
- Type `docker ps -a` to make sure docker is working.

On Mac:
Install docker on Mac : https://docs.docker.com/desktop/install/mac-install

- Once docker is installed, start **Docker Desktop**.
  Open **terminal** app
  Type 'docker ps –a' to make sure docker is working.

On Linux:
Install docker on Linux : https://docs.docker.com/desktop/install/linux-install

- Follow similar steps as Mac for starting Docker and verifying the installation



STEP 2 : DOWNLOAD DATA

Use Case: AWI_001 - Sipsey Fork, Black Warrior river **- 03W VPU (catchment-113060) Subset data:**

Index file for sample data on S3 bucket: index.html

Steps :

$ mkdir NextGen

$ cd NextGen

$ mkdir ngen-data

$ cd ngen-data

$ wget --no-parent https://ciroh-ua-ngen-data.s3.us-east-2.amazonaws.com/AWI-001/AWI_03W_113060_001.tar.gz

$ tar -xf AWI_03W_113060_001.tar.gz

$ cd AWI_03W_113060_001

$ pwd

```
apatel54@UA-W2RP43G:~/NextGen/ngen-data|⇒  wget --no-parent https://ciroh-ua-ngen-data.s3.us-east-2.amazonaws.com/AWI-001/AWI_03W_
113060_001.tar.gz
--2023-05-12 08:45:00--  https://ciroh-ua-ngen-data.s3.us-east-2.amazonaws.com/AWI-001/AWI_03W_113060_001.tar.gz
Resolving ciroh-ua-ngen-data.s3.us-east-2.amazonaws.com (ciroh-ua-ngen-data.s3.us-east-2.amazonaws.com)... 52.219.176.106
Connecting to ciroh-ua-ngen-data.s3.us-east-2.amazonaws.com (ciroh-ua-ngen-data.s3.us-east-2.amazonaws.com)|52.219.176.106|:443...
 connected.
HTTP request sent, awaiting response... 200 OK
Length: 329467276 (314M) [application/x-tar]
Saving to: 'AWI_03W_113060_001.tar.gz'

AWI_03W_113060_001.tar.gz       100%[=================================================>] 314.20M  23.0MB/s    in 18s

2023-05-12 08:45:19 (17.7 MB/s) - 'AWI_03W_113060_001.tar.gz' saved [329467276/329467276]

apatel54@UA-W2RP43G:~/NextGen/ngen-data|⇒
```

```
apatel54@UA-W2RP43G:~/NextGen/ngen-data|⇒  ls -lart
total 655952
drwxr-xr-x   5 apatel54  staff        160 May 10 23:24 AWI_03W_113060_001
-rw-r--r--   1 apatel54  staff  329467276 May 11 10:51 AWI_03W_113060_001.tar.gz
drwxr-xr-x  10 apatel54  staff        320 May 12 08:44 ..
drwxr-xr-x   4 apatel54  staff        128 May 12 09:23 .
```

Please contact below for more info on how to prepare the dataset:

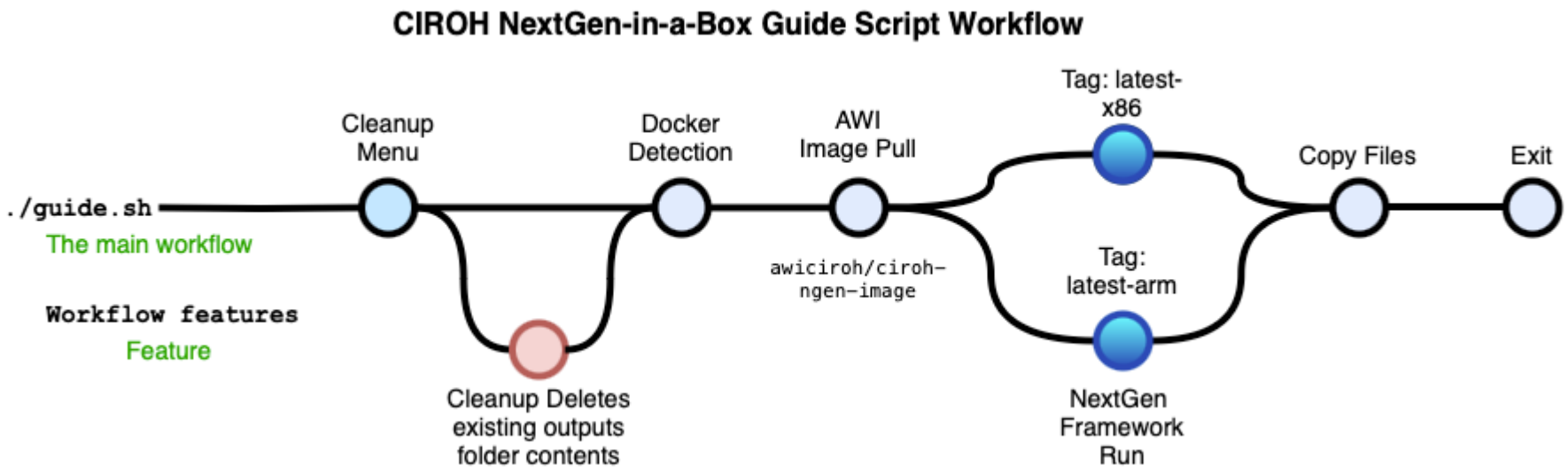
Preparing Data for NextGen-Based Simulations

- Irene Garousi-Nejad, CUAHSI
- James Halgren, University of Alabama

STEP 3: NEXTGEN IN A BOX: GUIDE.SH WORKFLOW

Clone repo:
$ cd NextGen
$ git clone https://github.com/CIROH-UA/CloudInfra.git
$ cd CloudInfra

Latest guide.sh location: https://github.com/CIROH-UA/CloudInfra/blob/main/guide.sh



STEP 3: RUN GUIDE.SH

~/NextGen/CloudInfra => ./guide.sh

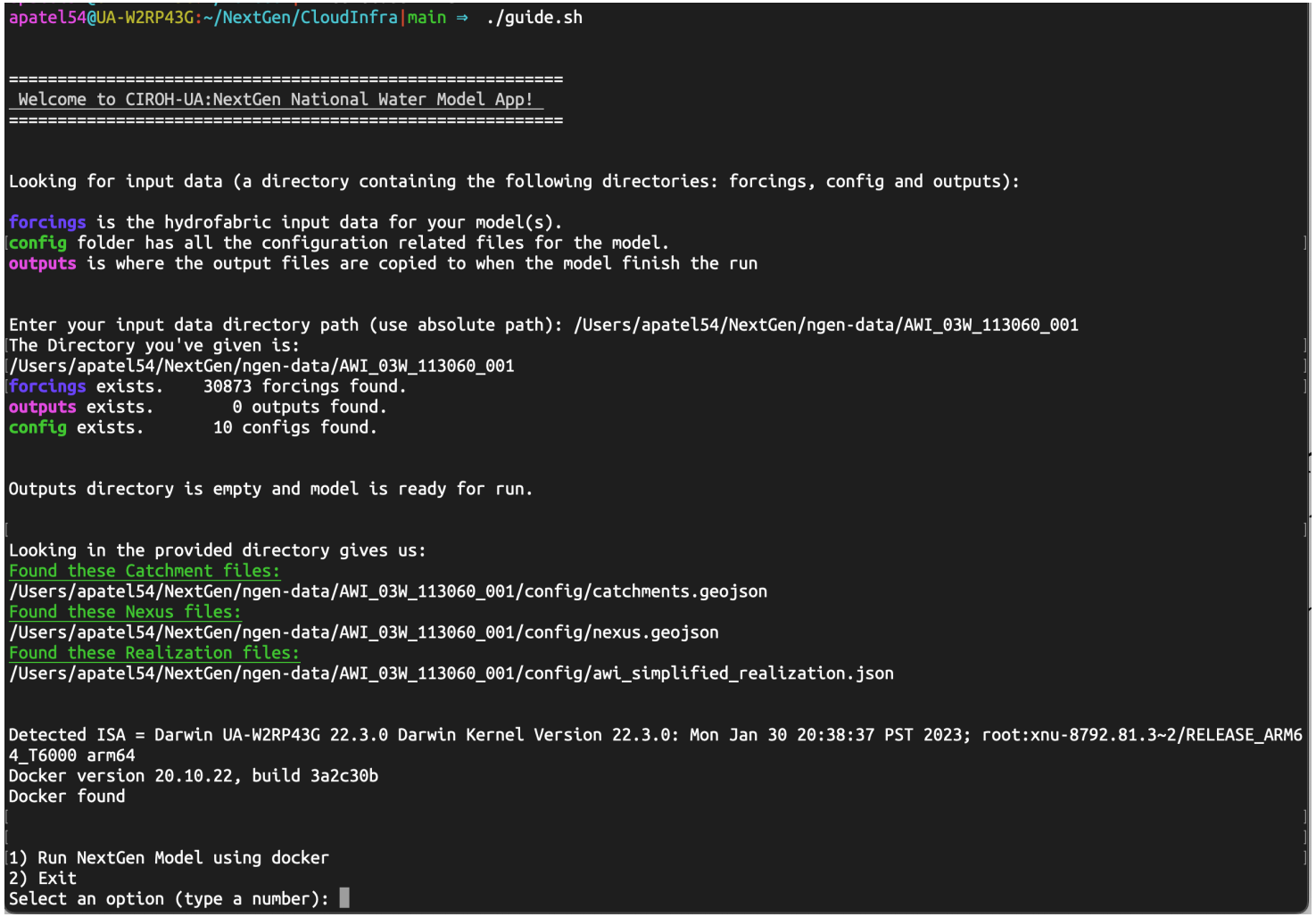
....

....

....

Script will execute and prompt the user for entering the absolute path of input data.

Enter absolute input data path. (copied in notepad)



~/NextGen/CloudInfra => ./guide.sh

....

....

Script will display catchment, nexus and realization files under input data.

Script will detect the OS architecture (arm vs x86) and make sure Docker is running.

Script will prompt user to enter "1" if ready to run the Model!

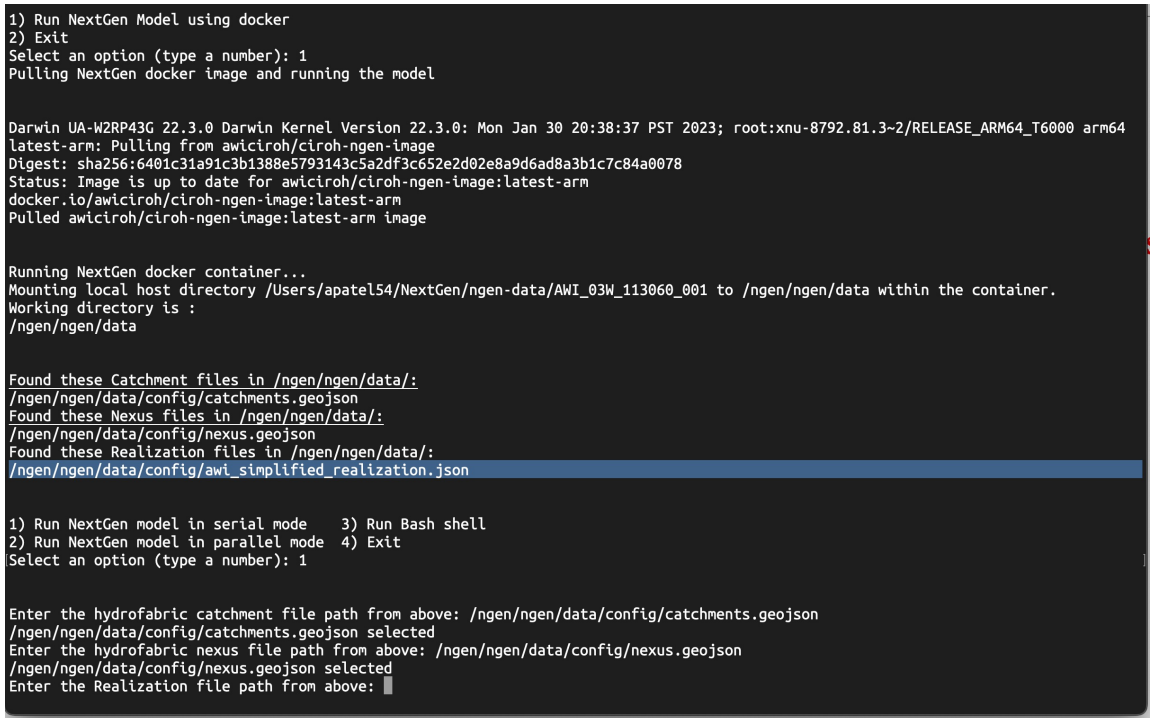
```
Detected ISA = Darwin UA-W2RP43G 22.3.0 Darwin Kernel Version 22.3.0: Mon Jan 30 20:38:37 PST 2023; root:xnu-8792.81.3~2/RELEASE_ARM6
4_T6000 arm64
Docker version 20.10.22, build 3a2c30b
Docker found


1) Run NextGen Model using docker
2) Exit
Select an option (type a number): 1
Pulling NextGen docker image and running the model


Darwin UA-W2RP43G 22.3.0 Darwin Kernel Version 22.3.0: Mon Jan 30 20:38:37 PST 2023; root:xnu-8792.81.3~2/RELEASE_ARM64_T6000 arm64
latest-arm: Pulling from awiciroh/ciroh-ngen-image
Digest: sha256:6401c31a91c3b1388e5793143c5a2df3c652e2d02e8a9d6ad8a3b1c7c84a0078
Status: Image is up to date for awiciroh/ciroh-ngen-image:latest-arm
docker.io/awiciroh/ciroh-ngen-image:latest-arm
Pulled awiciroh/ciroh-ngen-image:latest-arm image


Running NextGen docker container...
Mounting local host directory /Users/apatel54/NextGen/ngen-data/AWI_03W_113060_001 to /ngen/ngen/data within the container.
Working directory is :
/ngen/ngen/data


Found these Catchment files in /ngen/ngen/data/:
/ngen/ngen/data/config/catchments.geojson
Found these Nexus files in /ngen/ngen/data/:
/ngen/ngen/data/config/nexus.geojson
Found these Realization files in /ngen/ngen/data/:
/ngen/ngen/data/config/awi_simplified_realization.json


1) Run NextGen model in serial mode    3) Run Bash shell
2) Run NextGen model in parallel mode  4) Exit
Select an option (type a number):
```

~/NextGen/CloudInfra => ./guide.sh

....

- Script will pull docker image based on detected OS.
- Script will mount input data directory inside the container.
- Script will display paths of the files inside the container.
- The user is then prompted to select whether they want to run the model in serial mode (1) or parallel mode (2).

RUN SERIAL MODE

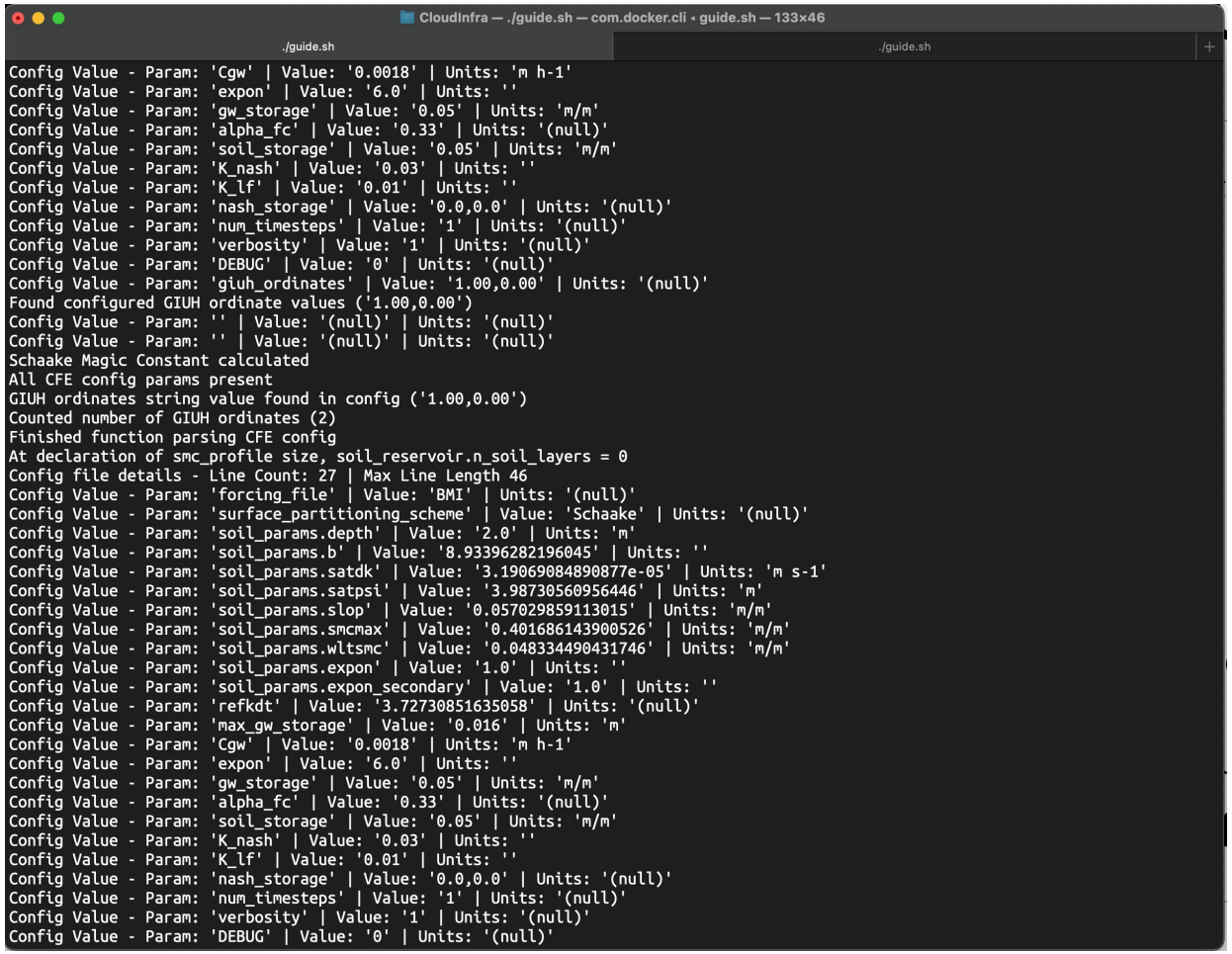


~/NextGen/CloudInfra => ./guide.sh

....

....

- Enter "1" for serial mode run.
- Copy and paste catchment, nexus and realization files path from above list.
- Model will run in serial mode.



~/NextGen/CloudInfra => ./guide.sh

....

....

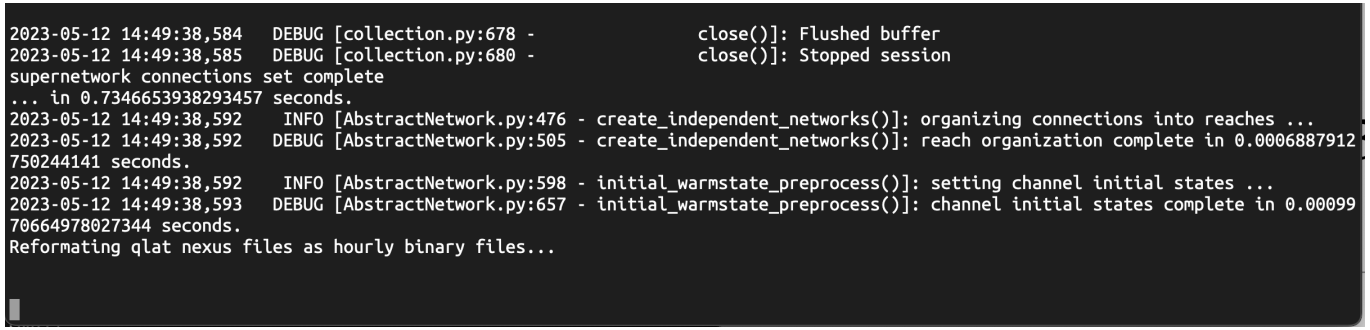
- This output on console indicates script is running.
- This step will take 2-3 mins depending on hardware used.

~/NextGen/CloudInfra => ./guide.sh

....

....

....



- This output on console indicates script is running.
- This step will take 2-3 mins depending on hardware used. "SIP Water!!!" :)

```
2023-05-12 23:01:55,115   DEBUG [output.py:315 - nwm_output_generator()]: writing CSV file took 4.577499628067017 seconds.
2023-05-12 23:01:55,115   DEBUG [output.py:370 - nwm_output_generator()]:          (0, q)    (0, v)    (0, d)  ...  (2879, q)  (2879, v)  (28
113035  0.004222  0.077847  0.010080  ...   0.000120   0.018764   0.001184
113036  0.000000  0.000000  0.000000  ...   0.000747   0.040594   0.003430
113037  0.007119  0.049306  0.010021  ...   0.008429   0.066446   0.015711
113038  0.000000  0.000000  0.000000  ...   0.009002   0.121398   0.012910
113039  0.000000  0.000000  0.000000  ...   0.009086   0.118733   0.014489
...          ...       ...       ...  ...        ...        ...        ...
125120  0.000000  0.000000  0.000000  ...   0.000000   0.000000   0.000000
125121  0.000000  0.000000  0.000000  ...   0.000000   0.000000   0.000000
125122  0.007620  0.094260  0.010146  ...   0.000174   0.024352   0.001323
125123  0.009853  0.086726  0.010178  ...   0.000193   0.028134   0.001865
125124  0.000000  0.000000  0.000000  ...   0.000000   0.000000   0.000000

[539 rows x 8640 columns]
2023-05-12 23:01:55,196   DEBUG [output.py:372 - nwm_output_generator()]: output complete in 4.662291526794434 seconds.
2023-05-12 23:01:55,198   DEBUG [__main__.py:271 -            main_v04()]: process complete in 119.47762989997864 seconds.
Finished routing
creating supernetwork connections set
supernetwork connections set complete
... in 1.8237299919128418 seconds.
Reformating qlat nexus files as hourly binary files...
************ TIMING SUMMARY ************
----------------------------------------
Network graph construction: 1.83 secs, 1.53 %
Forcing array construction: 112.31 secs, 94.0 %
Routing computations: 0.67 secs, 0.56 %
Output writing: 4.66 secs, 3.9 %
----------------------------------------
Total execution time: 119.47 secs
Would you like to continue?
1) Interactive-Shell
2) Copy output data from container to local machine
3) Exit
Select an option (type a number): 2
997 new outputs created.
Any copied files can be found here: /home/ec2-user/NextGen/ngen-data/AWI_03W_113060_001/outputs
Thank you for running NextGen In A Box: National Water Model! Have a nice day!
```

~/NextGen/CloudInfra => ./guide.sh

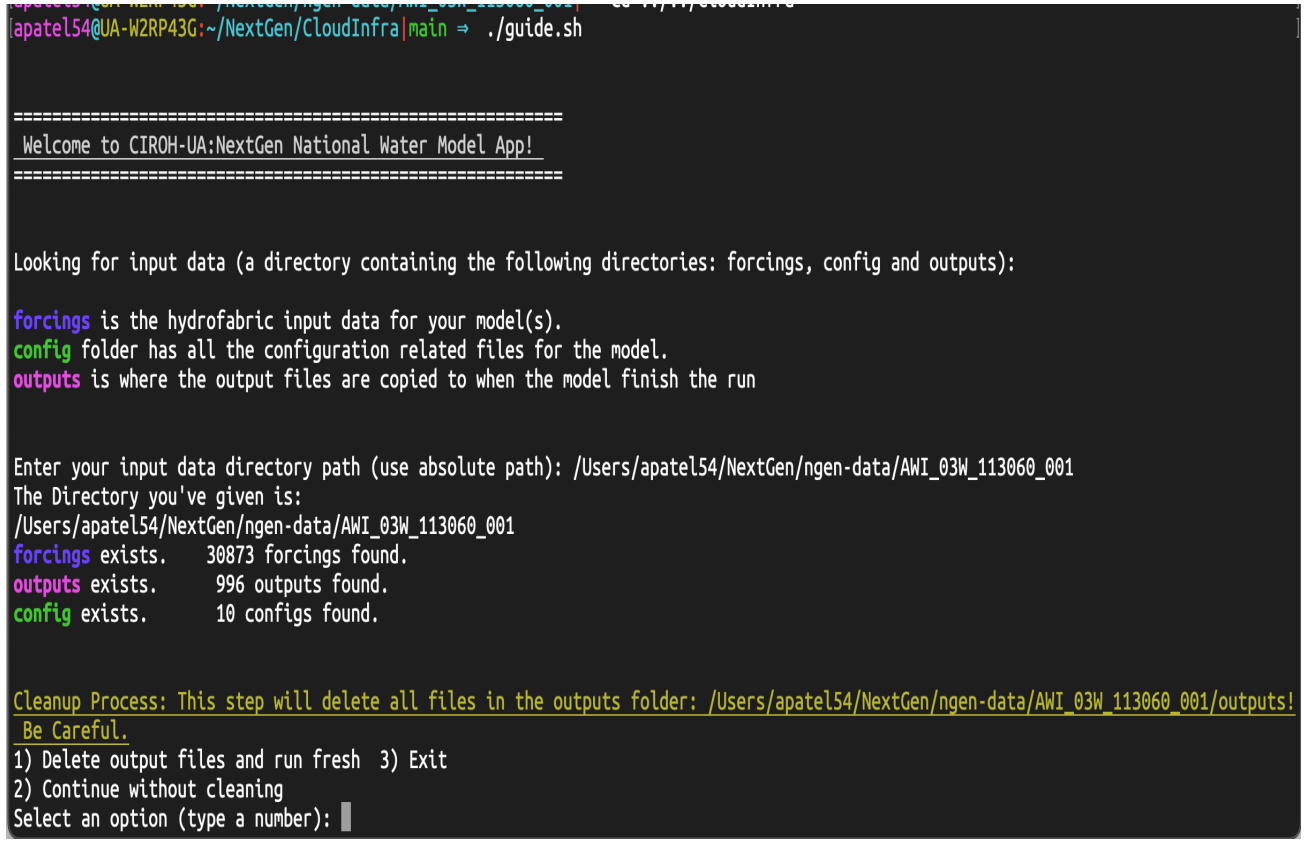
....

....

....

- Timing summary is displayed.
- User is prompted if they want to exec into the container (1) or directly want to copy the output files to the local machine. (2)
- Select "2" to copy the output files.
- Output files are copied to path displayed.

RUN PARALLEL MODE



~/NextGen/CloudInfra => ./guide.sh

....

....

....

- If outputs folder is not empty, script will prompt for cleaning outputs folder before rerunning new flow.
- Enter "1"

```
Cleanup Process: This step will delete all files in the outputs folder: /Users/apatel54/NextGen/ngen-data/AWI_03W_113060_001/outputs!
 Be Careful.
1) Delete output files and run fresh  3) Exit
2) Continue without cleaning
Select an option (type a number): 1
Cleaning Outputs folder for fresh run
Starting Cleanup of Files:      996


Looking in the provided directory gives us:
Found these Catchment files:
/Users/apatel54/NextGen/ngen-data/AWI_03W_113060_001/config/catchments.geojson
Found these Nexus files:
/Users/apatel54/NextGen/ngen-data/AWI_03W_113060_001/config/nexus.geojson
Found these Realization files:
/Users/apatel54/NextGen/ngen-data/AWI_03W_113060_001/config/awi_simplified_realization.json


Detected ISA = Darwin UA-W2RP43G 22.3.0 Darwin Kernel Version 22.3.0: Mon Jan 30 20:38:37 PST 2023; root:xnu-8792.81.3~2/RELEASE_ARM6
4_T6000 arm64
Docker version 20.10.22, build 3a2c30b
Docker found


1) Run NextGen Model using docker
2) Exit
Select an option (type a number): 1
```

~/NextGen/CloudInfra => ./guide.sh

....

....

- Follow same steps as for serial mode.
- Enter "1" for running NextGen Model.

```
Darwin UA-W2RP43G 22.3.0 Darwin Kernel Version 22.3.0: Mon Jan 30 20:38:37 PST 2023; root:xnu-8792.81.3~2/RELEASE_ARM64_T6000 arm64
latest-arm: Pulling from awiciroh/ciroh-ngen-image
Digest: sha256:6401c31a91c3b1388e5793143c5a2df3c652e2d02e8a9d6ad8a3b1c7c84a0078
Status: Image is up to date for awiciroh/ciroh-ngen-image:latest-arm
docker.io/awiciroh/ciroh-ngen-image:latest-arm
Pulled awiciroh/ciroh-ngen-image:latest-arm image


Running NextGen docker container...
Mounting local host directory /Users/apatel54/NextGen/ngen-data/AWI_03W_113060_001 to /ngen/ngen/data within the container.
Working directory is :
/ngen/ngen/data


Found these Catchment files in /ngen/ngen/data/:
/ngen/ngen/data/config/catchments.geojson
Found these Nexus files in /ngen/ngen/data/:
/ngen/ngen/data/config/nexus.geojson
Found these Realization files in /ngen/ngen/data/:
/ngen/ngen/data/config/awi_simplified_realization.json


1) Run NextGen model in serial mode    3) Run Bash shell
2) Run NextGen model in parallel mode  4) Exit
Select an option (type a number): 2


Enter the hydrofabric catchment file path: /ngen/ngen/data/config/catchments.geojson
/ngen/ngen/data/config/catchments.geojson selected
Enter the hydrofabric nexus file path: /ngen/ngen/data/config/nexus.geojson
/ngen/ngen/data/config/nexus.geojson selected
Enter the Realization file path:
```

~/NextGen/CloudInfra => ./guide.sh

....

....

- Follow same steps as for serial mode.
- Enter "2" for running NextGen Model in parallel mode.
- Copy and paste catchment, nexus and realization files path.
- Model will run in parallel mode.

```
Config Value - Param: 'surface_partitioning_scheme' | Value: 'Schaake' | Units: '(null)'
Config Value - Param: 'soil_params.depth' | Value: '2.0' | Units: 'm'
Config Value - Param: 'soil_params.b' | Value: '8.93396282196045' | Units: ''
Config Value - Param: 'soil_params.satdk' | Value: '3.19069084890877e-05' | Units: 'm s-1'
Config Value - Param: 'soil_params.satpsi' | Value: '3.98730560956446' | Units: 'm'
Config Value - Param: 'soil_params.slop' | Value: '0.057029859113015' | Units: 'm/m'
Config Value - Param: 'soil_params.smcmax' | Value: '0.401686143900526' | Units: 'm/m'
Config Value - Param: 'soil_params.wltsmc' | Value: '0.048334490431746' | Units: 'm/m'
Config Value - Param: 'soil_params.expon' | Value: '1.0' | Units: ''
Config Value - Param: 'soil_params.expon_secondary' | Value: '1.0' | Units: ''
Config Value - Param: 'refkdt' | Value: '3.72730851635058' | Units: '(null)'
Config Value - Param: 'max_gw_storage' | Value: '0.016' | Units: 'm'
Config Value - Param: 'Cgw' | Value: '0.0018' | Units: 'm h-1'
Config Value - Param: 'expon' | Value: '6.0' | Units: ''
Config Value - Param: 'gw_storage' | Value: '0.05' | Units: 'm/m'
Config Value - Param: 'alpha_fc' | Value: '0.33' | Units: '(null)'
Config Value - Param: 'soil_storage' | Value: '0.05' | Units: 'm/m'
Config Value - Param: 'K_nash' | Value: '0.03' | Units: ''
Config Value - Param: 'K_lf' | Value: '0.01' | Units: ''
Config Value - Param: 'nash_storage' | Value: '0.0,0.0' | Units: '(null)'
Config Value - Param: 'num_timesteps' | Value: '1' | Units: '(null)'
Config Value - Param: 'verbosity' | Value: '1' | Units: '(null)'
Config Value - Param: 'DEBUG' | Value: '0' | Units: '(null)'
Config Value - Param: 'giuh_ordinates' | Value: '1.00,0.00' | Units: '(null)'
Found configured GIUH ordinate values ('1.00,0.00')
Config Value - Param: '' | Value: '(null)' | Units: '(null)'
Config Value - Param: '' | Value: '(null)' | Units: '(null)'
Schaake Magic Constant calculated
All CFE config params present
GIUH ordinates string value found in config ('1.00,0.00')
Counted number of GIUH ordinates (2)
Finished function parsing CFE config
At declaration of smc_profile size, soil_reservoir.n_soil_layers = 0
```

```
2023-05-12 15:29:07,997   DEBUG [collection.py:678 -                close()]: Flushed buffer
2023-05-12 15:29:07,998   DEBUG [collection.py:680 -                close()]: Stopped session
2023-05-12 15:29:08,006    INFO [AbstractNetwork.py:476 - create_independent_networks()]: organizing connections into reaches ...
2023-05-12 15:29:08,007   DEBUG [AbstractNetwork.py:505 - create_independent_networks()]: reach organization complete in 0.0006945133
209228516 seconds.
2023-05-12 15:29:08,007    INFO [AbstractNetwork.py:598 - initial_warmstate_preprocess()]: setting channel initial states ...
2023-05-12 15:29:08,008   DEBUG [AbstractNetwork.py:657 - initial_warmstate_preprocess()]: channel initial states complete in 0.00103
23524475097656 seconds.
```

~/NextGen/CloudInfra => ./guide.sh

....

....

....

- This output on console indicates script is running.
- This step will few mins depending on hardware used. Please wait...

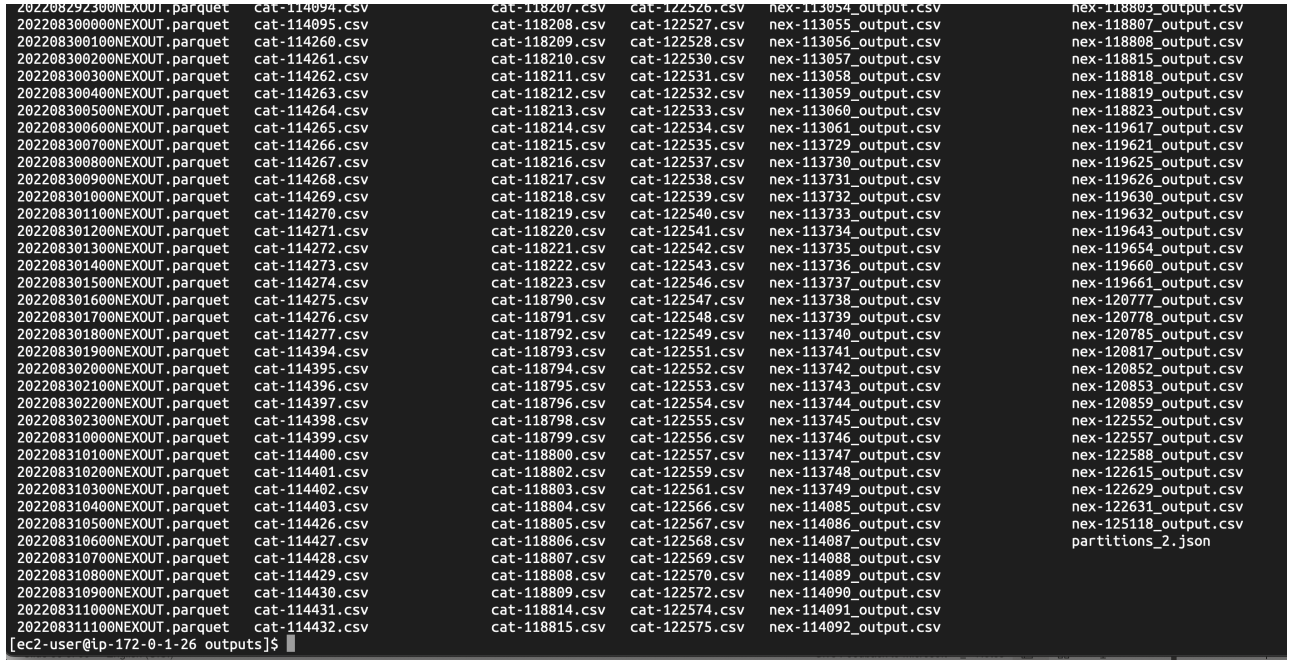
```
2023-05-14 02:15:13,517   DEBUG [output.py:315 - nwm_output_generator()]: writing CSV file took 4.583699703216553 seconds.
2023-05-14 02:15:13,517   DEBUG [output.py:370 - nwm_output_generator()]:          (0, q)    (0, v)    (0, d)  ...  (2879, q)  (2879, v)  (28
113035  0.004222  0.077847  0.010080  ...   0.000120   0.018764   0.001184
113036  0.000000  0.000000  0.000000  ...   0.000747   0.040594   0.003430
113037  0.007119  0.049306  0.010021  ...   0.008429   0.066446   0.015711
113038  0.000000  0.000000  0.000000  ...   0.009002   0.121398   0.012910
113039  0.000000  0.000000  0.000000  ...   0.009086   0.118733   0.014489
...          ...       ...       ...  ...        ...        ...        ...
125120  0.000000  0.000000  0.000000  ...   0.000000   0.000000   0.000000
125121  0.000000  0.000000  0.000000  ...   0.000000   0.000000   0.000000
125122  0.007620  0.094260  0.010146  ...   0.000174   0.024352   0.001323
125123  0.009853  0.086726  0.010178  ...   0.000193   0.028134   0.001865
125124  0.000000  0.000000  0.000000  ...   0.000000   0.000000   0.000000

[539 rows x 8640 columns]
2023-05-14 02:15:13,600   DEBUG [output.py:372 - nwm_output_generator()]: output complete in 4.670687675476074 seconds.
2023-05-14 02:15:13,602   DEBUG [__main__.py:271 -             main_v04()]: process complete in 120.44207549095154 seconds.
Finished routing
creating supernetwork connections set
supernetwork connections set complete
... in 1.7141664028167725 seconds.
Reformating qlat nexus files as hourly binary files...
************ TIMING SUMMARY ************
----------------------------------------
Network graph construction: 1.72 secs, 1.43 %
Forcing array construction: 113.37 secs, 94.13 %
Routing computations: 0.67 secs, 0.56 %
Output writing: 4.67 secs, 3.88 %
----------------------------------------
Total execution time: 120.43 secs
Would you like to continue?
1) Interactive-Shell
2) Copy output data from container to local machine
3) Exit
Select an option (type a number): 2
997 new outputs created.
Any copied files can be found here: /home/ec2-user/NextGen/ngen-data/AWI_03W_113060_001/outputs
Thank you for running NextGen In A Box: National Water Model! Have a nice day!
```

~/NextGen/CloudInfra =>
./guide.sh

....

....

- Timing summary is displayed.
- User is prompted if want to exec into the container (1) or directly want to copy the output files to the local machine. (2)
- Select "2" to copy the output files.
- Output files are copied to path displayed.



```
202208292300NEXOUT.parquet   cat-114094.csv   cat-118207.csv   cat-122526.csv   nex-113054_output.csv   nex-118803_output.csv
202208300000NEXOUT.parquet   cat-114095.csv   cat-118208.csv   cat-122527.csv   nex-113055_output.csv   nex-118807_output.csv
202208300100NEXOUT.parquet   cat-114260.csv   cat-118209.csv   cat-122528.csv   nex-113056_output.csv   nex-118808_output.csv
202208300200NEXOUT.parquet   cat-114261.csv   cat-118210.csv   cat-122530.csv   nex-113057_output.csv   nex-118815_output.csv
202208300300NEXOUT.parquet   cat-114262.csv   cat-118211.csv   cat-122531.csv   nex-113058_output.csv   nex-118818_output.csv
202208300400NEXOUT.parquet   cat-114263.csv   cat-118212.csv   cat-122532.csv   nex-113059_output.csv   nex-118819_output.csv
202208300500NEXOUT.parquet   cat-114264.csv   cat-118213.csv   cat-122533.csv   nex-113060_output.csv   nex-118823_output.csv
202208300600NEXOUT.parquet   cat-114265.csv   cat-118214.csv   cat-122534.csv   nex-113061_output.csv   nex-119617_output.csv
202208300700NEXOUT.parquet   cat-114266.csv   cat-118215.csv   cat-122535.csv   nex-113729_output.csv   nex-119621_output.csv
202208300800NEXOUT.parquet   cat-114267.csv   cat-118216.csv   cat-122537.csv   nex-113730_output.csv   nex-119625_output.csv
202208300900NEXOUT.parquet   cat-114268.csv   cat-118217.csv   cat-122538.csv   nex-113731_output.csv   nex-119626_output.csv
202208301000NEXOUT.parquet   cat-114269.csv   cat-118218.csv   cat-122539.csv   nex-113732_output.csv   nex-119630_output.csv
202208301100NEXOUT.parquet   cat-114270.csv   cat-118219.csv   cat-122540.csv   nex-113733_output.csv   nex-119632_output.csv
202208301200NEXOUT.parquet   cat-114271.csv   cat-118220.csv   cat-122541.csv   nex-113734_output.csv   nex-119643_output.csv
202208301300NEXOUT.parquet   cat-114272.csv   cat-118221.csv   cat-122542.csv   nex-113735_output.csv   nex-119654_output.csv
202208301400NEXOUT.parquet   cat-114273.csv   cat-118222.csv   cat-122543.csv   nex-113736_output.csv   nex-119660_output.csv
202208301500NEXOUT.parquet   cat-114274.csv   cat-118223.csv   cat-122546.csv   nex-113737_output.csv   nex-119661_output.csv
202208301600NEXOUT.parquet   cat-114275.csv   cat-118790.csv   cat-122547.csv   nex-113738_output.csv   nex-120777_output.csv
202208301700NEXOUT.parquet   cat-114276.csv   cat-118791.csv   cat-122548.csv   nex-113739_output.csv   nex-120778_output.csv
202208301800NEXOUT.parquet   cat-114277.csv   cat-118792.csv   cat-122549.csv   nex-113740_output.csv   nex-120785_output.csv
202208301900NEXOUT.parquet   cat-114394.csv   cat-118793.csv   cat-122551.csv   nex-113741_output.csv   nex-120817_output.csv
202208302000NEXOUT.parquet   cat-114395.csv   cat-118794.csv   cat-122552.csv   nex-113742_output.csv   nex-120852_output.csv
202208302100NEXOUT.parquet   cat-114396.csv   cat-118795.csv   cat-122553.csv   nex-113743_output.csv   nex-120853_output.csv
202208302200NEXOUT.parquet   cat-114397.csv   cat-118796.csv   cat-122554.csv   nex-113744_output.csv   nex-120859_output.csv
202208302300NEXOUT.parquet   cat-114398.csv   cat-118798.csv   cat-122555.csv   nex-113745_output.csv   nex-122552_output.csv
202208310000NEXOUT.parquet   cat-114399.csv   cat-118799.csv   cat-122556.csv   nex-113746_output.csv   nex-122557_output.csv
202208310100NEXOUT.parquet   cat-114400.csv   cat-118800.csv   cat-122557.csv   nex-113747_output.csv   nex-122588_output.csv
202208310200NEXOUT.parquet   cat-114401.csv   cat-118802.csv   cat-122559.csv   nex-113748_output.csv   nex-122615_output.csv
202208310300NEXOUT.parquet   cat-114402.csv   cat-118803.csv   cat-122561.csv   nex-113749_output.csv   nex-122629_output.csv
202208310400NEXOUT.parquet   cat-114403.csv   cat-118804.csv   cat-122566.csv   nex-114085_output.csv   nex-122631_output.csv
202208310500NEXOUT.parquet   cat-114426.csv   cat-118805.csv   cat-122567.csv   nex-114086_output.csv   nex-125118_output.csv
202208310600NEXOUT.parquet   cat-114427.csv   cat-118806.csv   cat-122568.csv   nex-114087_output.csv   partitions_2.json
202208310700NEXOUT.parquet   cat-114428.csv   cat-118807.csv   cat-122569.csv   nex-114088_output.csv
202208310800NEXOUT.parquet   cat-114429.csv   cat-118808.csv   cat-122570.csv   nex-114089_output.csv
202208310900NEXOUT.parquet   cat-114430.csv   cat-118809.csv   cat-122572.csv   nex-114090_output.csv
202208311000NEXOUT.parquet   cat-114431.csv   cat-118814.csv   cat-122574.csv   nex-114091_output.csv
202208311100NEXOUT.parquet   cat-114432.csv   cat-118815.csv   cat-122575.csv   nex-114092_output.csv
[ec2-user@ip-172-0-1-26 outputs]$
```

NEXTGEN IN A BOX – OUTPUT FILES

- **Nexus output file** - volumetric flow (m^3/s) from the catchment to the nexus
- e.g: **nex-113731_output.csv**

  0, 2022-08-24 13:00:00, 0.897584

  1, 2022-08-24 14:00:00, 0.814954

  2, 2022-08-24 15:00:00, 0.73377

  3, 2022-08-24 16:00:00, 0.686156

  4, 2022-08-24 17:00:00, 0.633547

- **Catchment output file:**
- **e.g : cat-114432.csv**
  - Time Step,Time,EVAPOTRANS,sloth_ice_fraction_schaake,sloth_ice_fraction_xinan,sloth_smp
  -
  - 0,2022-08-24 13:00:00,0.000000,0.000000,0.000000,0.000000

CONCLUSION

- FUTURE PLANS
- ISSUE TEMPLATE
- CONTRIBUTION GUIDE
- SUPPORT QUESTIONS

FUTURE PLANS FOR NEXTGEN-IN-A-BOX

- Implement "**CIROH Research to Operations Hybrid Cloud – R2OHC"** a flexible computational platform to support collaborative hydrology research surrounding the NextGen framework. – This will be CIROH-supported, community-accessible resource for supporting NextGen framework development in the hybrid cloud.
- **Cloud Migration, Automation, Infrastructure as Code.**
- Collaborate with researchers across CIROH to create **a series of case studies that demonstrate the value of a shared cloud computing environment in hydrology research**
- **GOALS**:
    - Promote **reproducible hydrologic computing experiments** with the NextGen Water Resource Modeling Framework;
    - Provide support for, and reduce the barrier to entry for performing NextGen-related **experiments at** various **scales**; and
    - Accelerate the interconnection and **integration of research products** and hydroinformatics innovations from the various ongoing CIROH experiments.

CONTRIBUTION GUIDE & BUG REPORT

- **Contribution Guide:**

  https://github.com/CIROH-UA/CloudInfra/blob/main/contribute.md

- **Report an issue:**

  https://github.com/CIROH-UA/CloudInfra/issues/new?assignees=&labels=bug&projects=&template=bug_report.md&title=

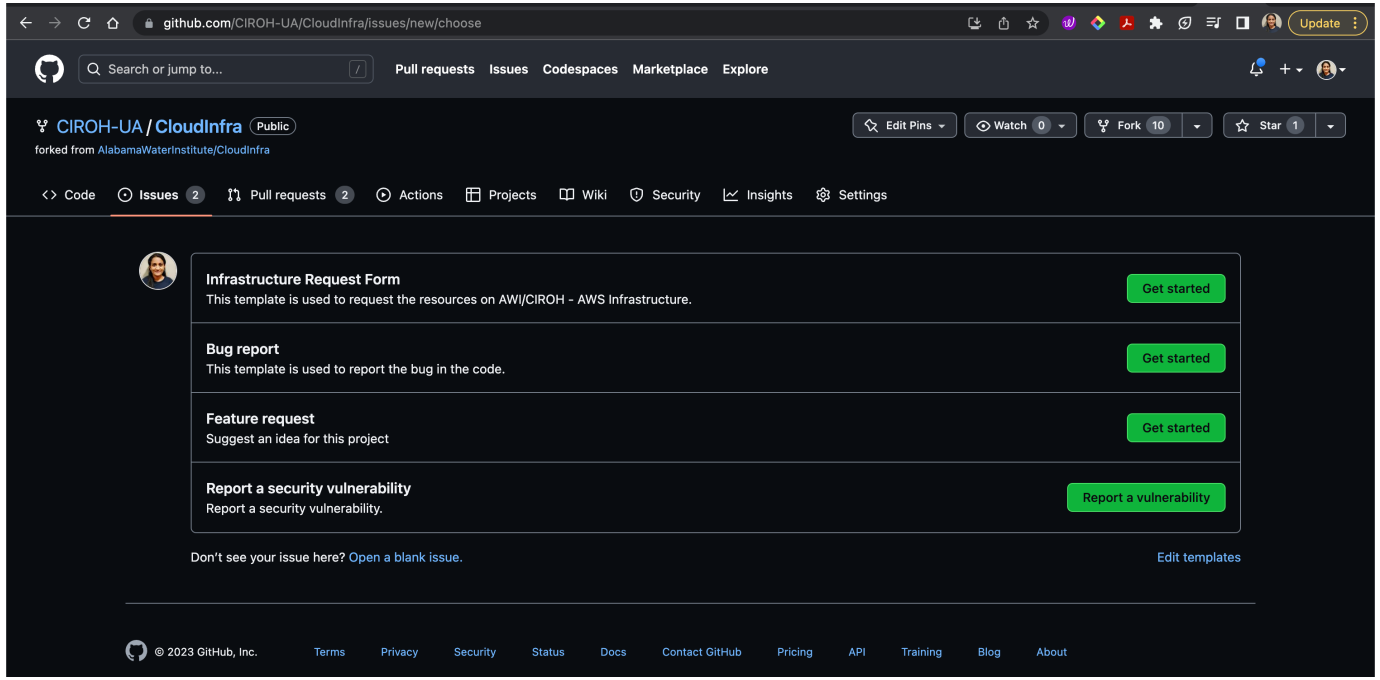
  - Together, let's drive innovation and advance academic research through the power of NextGen In a Box

FEATURE REQUEST

- Github -> Issues -> New Issues

- If you have a feature request, please open a new issue so we can discuss it and prioritize it accordingly.

- https://github.com/CIROH-UA/CloudInfra/issues/new?assignees=&labels=enhancement&projects=&template=feature_request.md&title=

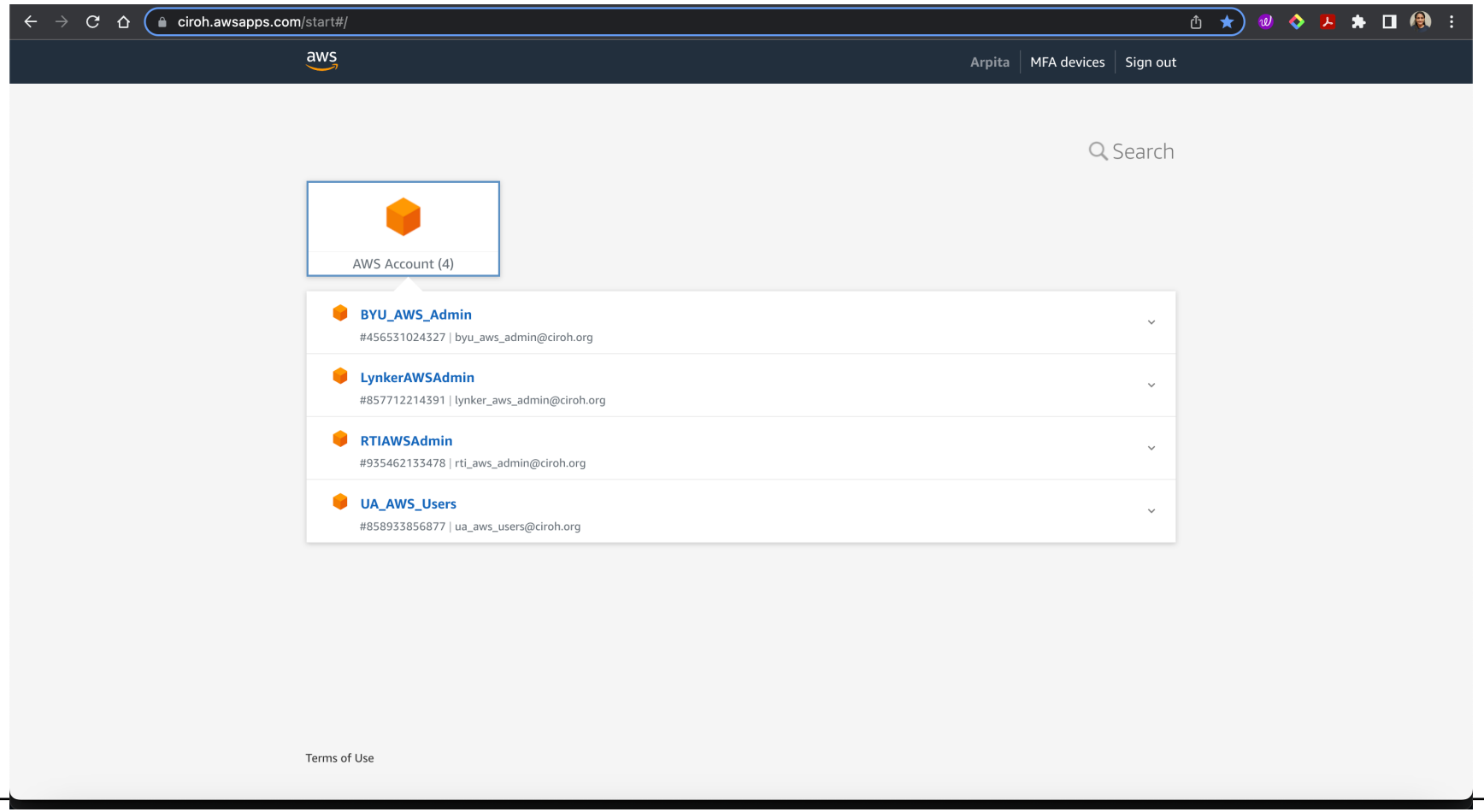


AWS INFRASTRUCTURE REQUEST FORM

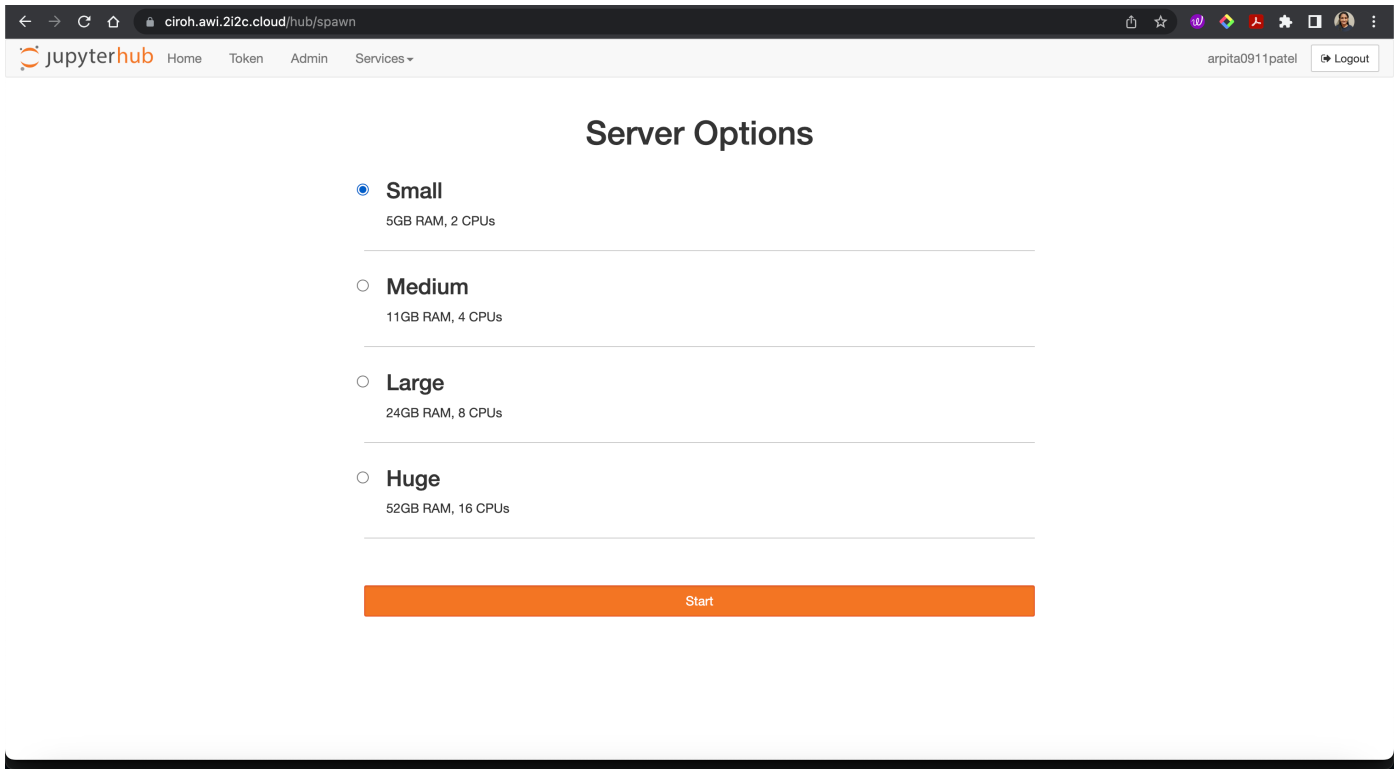
- AWS Infrastructure Request Form

- https://github.com/CIROH-UA/CloudInfra/issues/new?assignees=&labels=infrastructure&projects=&template=aws_infrastructure_request.md&title=

CIROH AWS ORGANIZATION



GOOGLE (VIA 2I2C) - CIROH-UA JUPYTERHUB

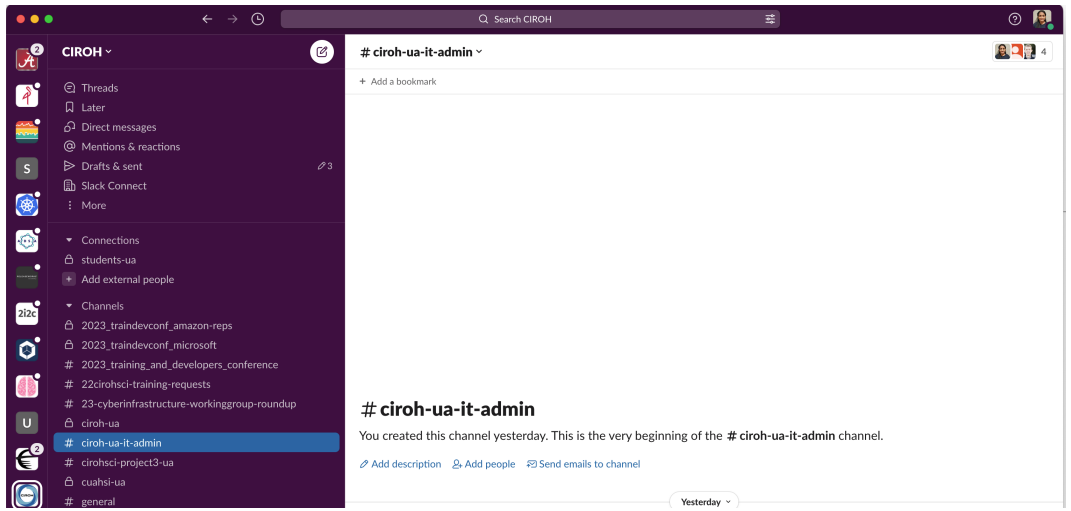


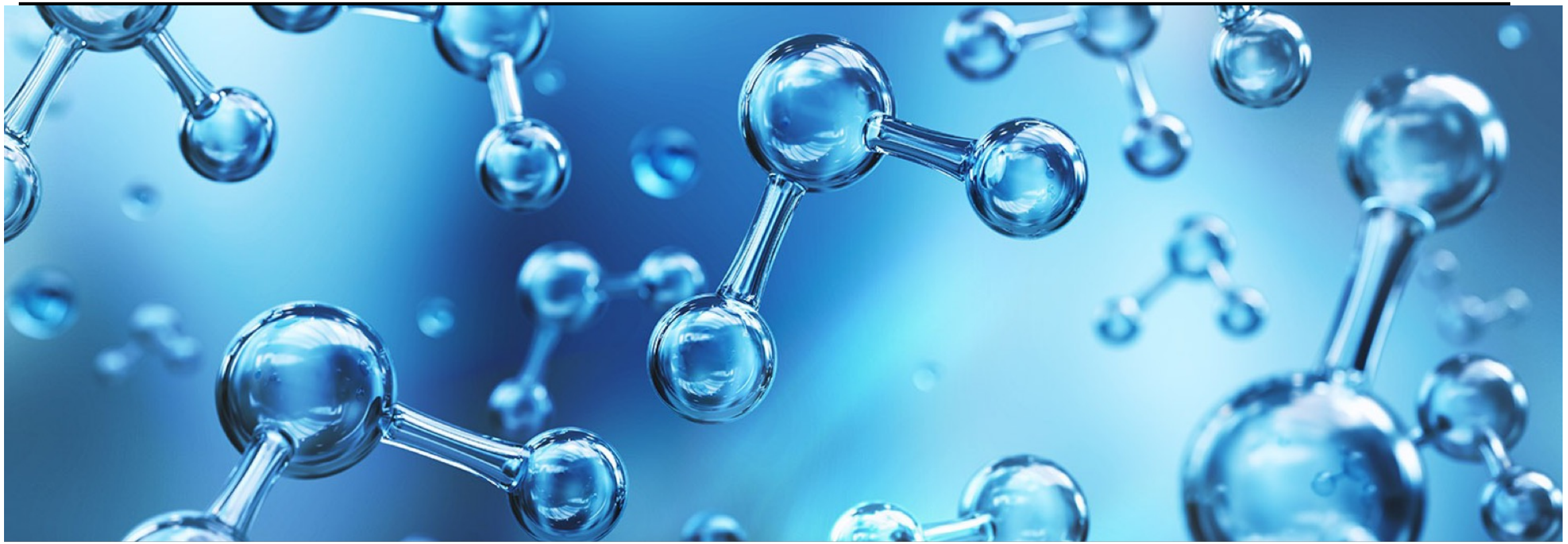
SUPPORT

- For any questions, comments or issues:

Email us at : ciroh-it-admin@ciroh.org

Contact us on CIROH Slack: **#ciroh-ua-it-admin**





Q&A?

Thank you for your time and attention 🙂

The University of Alabama
Alabama Water Institute
248 Kirkbride Ln.
Tuscaloosa, AL 35401

awi.ua.edu

THE UNIVERSITY OF ALABAMA | Alabama Water Institute