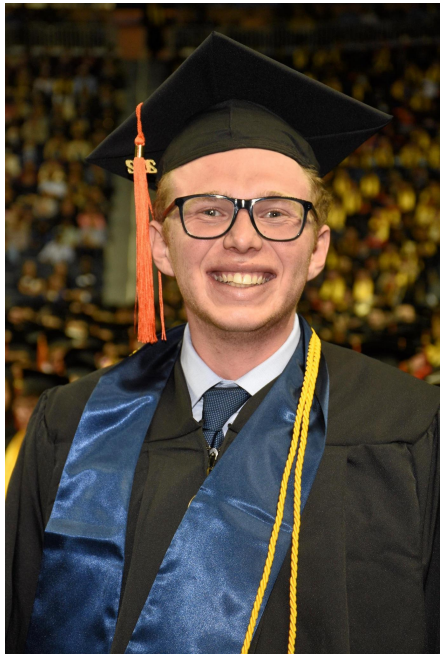
Migrating the UFS Graduate Student Tests to the Cloud

Sam Ephraim, OAR, Silver Spring, MD; and J. H. G. M. Alves, K. V. Kumar, L. Bernardet, D. Heinzeller, and M. Ji

EPIC
Earth Prediction Innovation Center

About Me - Education



University of Michigan
Class of 2022

BSE Climate and Meteorology
BSE Computer Science

University of Miami
Graduate School

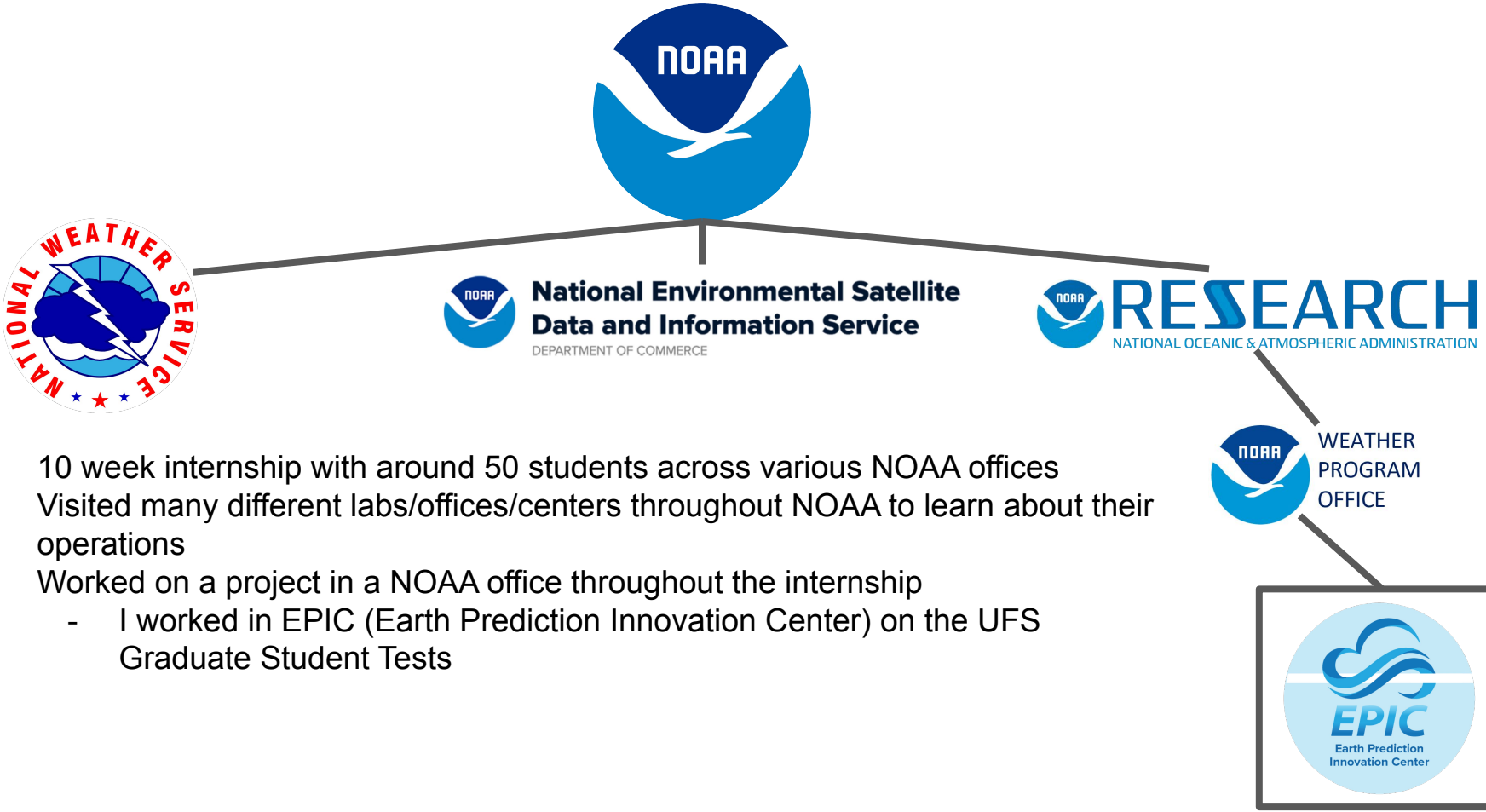
Pursuing MPS in Weather Forecasting

About Me - Hobbies

Lapenta Internship Experience

- 10 week internship with around 50 students across various NOAA offices
- Visited many different labs/offices/centers throughout NOAA to learn about their operations

Lapenta Internship Experience



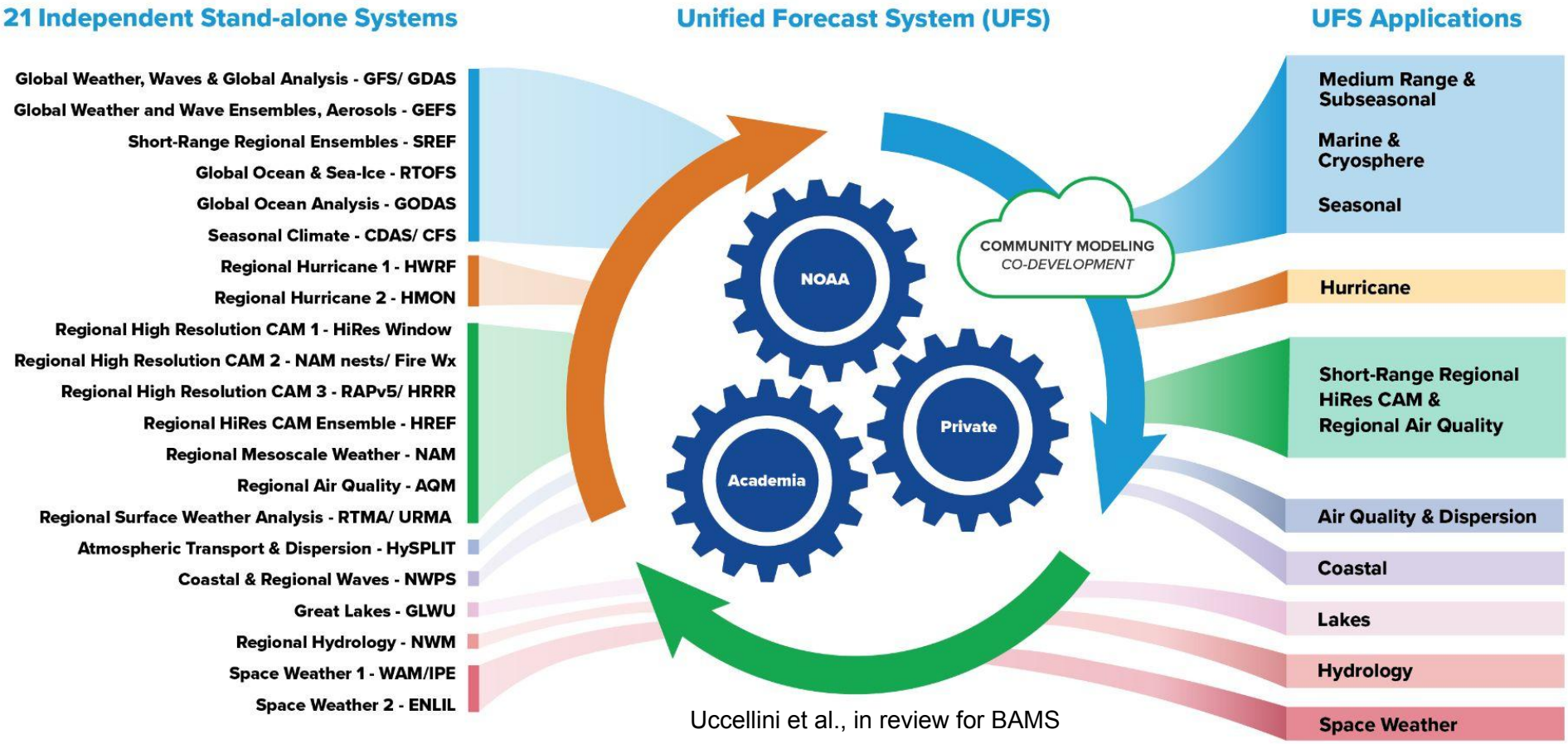
- 10 week internship with around 50 students across various NOAA offices
- Visited many different labs/offices/centers throughout NOAA to learn about their operations
- Worked on a project in a NOAA office throughout the internship
    - I worked in EPIC (Earth Prediction Innovation Center) on the UFS Graduate Student Tests

UFS BACKGROUND

UFS Overview
UFS Applications
Graduate Student Tests

Simplifying NOAA's Operational Forecast Suite

Reducing the 21 Stand-alone Operational Forecast Systems into Eight Applications



Uccellini et al., in review for BAMS

EPIC

Partnering with the community for the benefit of the nation

Vision: *Enable* the most accurate and reliable operational numerical forecast model in the world.

Mission: To be the *catalyst* for community research and modeling system advances that continually inform and accelerate advances in our nation's operational forecast modeling systems.

What EPIC is….

- A virtual community model development environment
- Management of cloud- ready code
- Community access to NOAA observations, data & tools
- Community support & engagement
- Clear research & model transition to operations priorities
- Expected expansion to other additional model components
- EPIC: focus on the Unified Forecast System (UFS)

Community Engagement

UFS
UNIFIED FORECAST SYSTEM

Cloud Use

Community

O2R

R2D

UFS Code Repository

R2O

Research to Operations Screening Funnel

R2O

UFS Code on NCEP WCOSS

Scientists, engineers, graduate students, and collaborators *(NOAA, DOD, NCAR, NASA, Academia, Private Sector)*

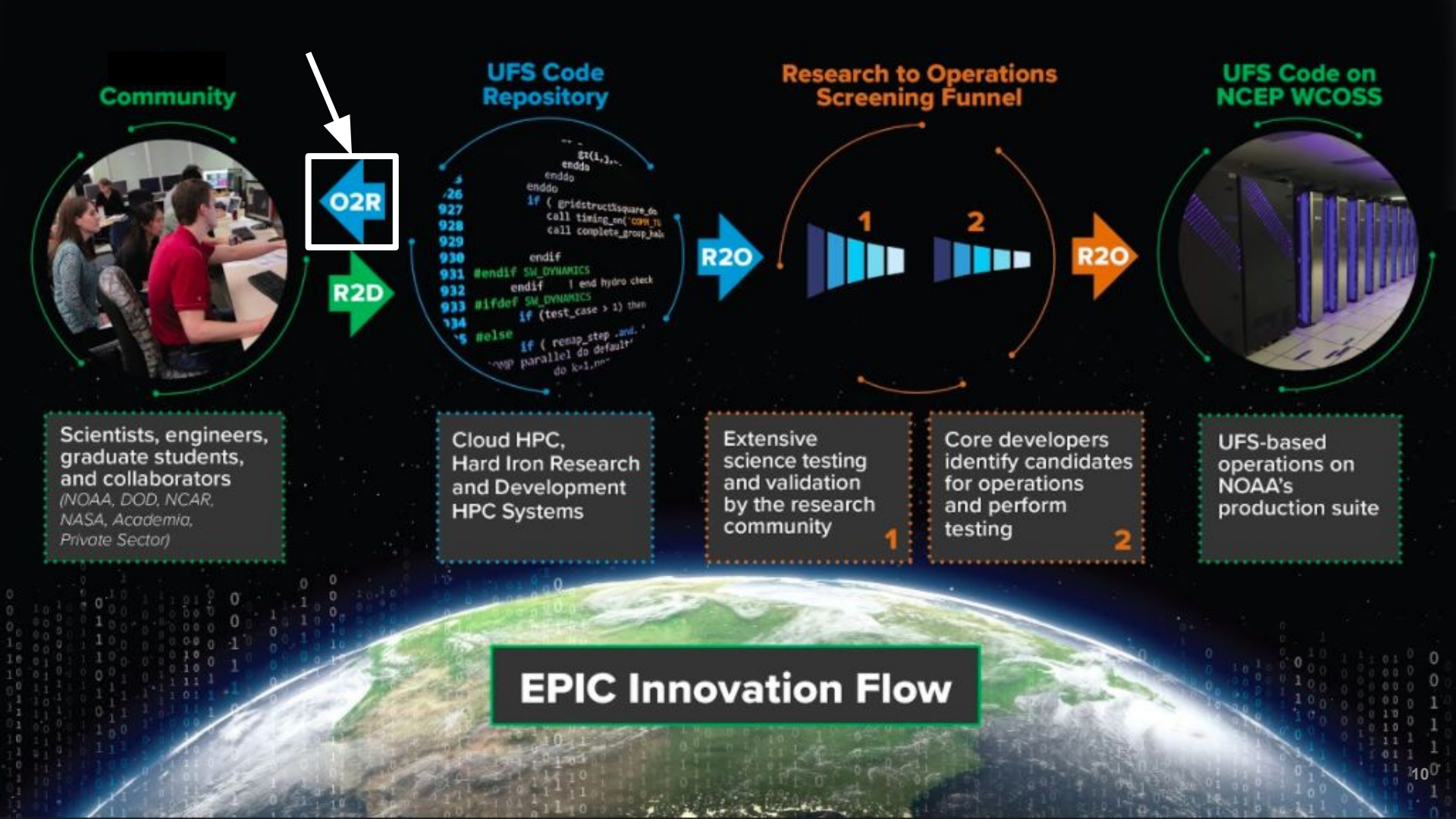
Cloud HPC, Hard Iron Research and Development HPC Systems

Extensive science testing and validation by the research community 1

Core developers identify candidates for operations and perform testing 2

UFS-based operations on NOAA's production suite

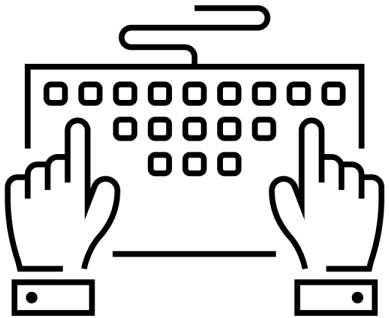
EPIC Innovation Flow

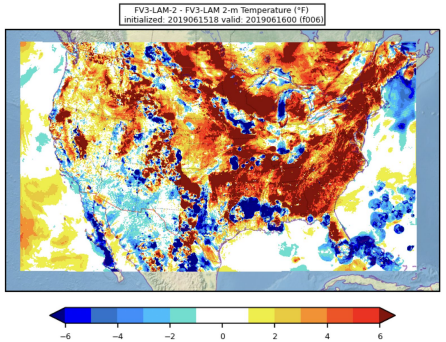

Community
O2R
R2D
UFS Code Repository
R2O
Research to Operations Screening Funnel
1
2
R2O
UFS Code on NCEP WCOSS
Scientists, engineers, graduate students, and collaborators (NOAA, DOD, NCAR, NASA, Academia, Private Sector)
Cloud HPC, Hard Iron Research and Development HPC Systems
Extensive science testing and validation by the research community 1
Core developers identify candidates for operations and perform testing 2
UFS-based operations on NOAA's production suite
EPIC Innovation Flow

UFS Graduate Student Tests

- Usability test: Modules that test how easy it is to run the UFS Short Range Weather (SRW) Application or the UFS Medium Range Weather (MRW) Application
  - Goal is that it's easy enough for a graduate student to complete in < 8 hours
  - Engages young scientists! (you don't have to be a grad student)
  - Get feedback

GitHub





Download Code → Run Code → Modify and Re-run Code → Plot and Compare Outputs

Objectives

Objectives/Deliverables

- Creating documentation to run containerized versions of Graduate Student Tests (GSTs) in the cloud
- Run the GSTs in <8hrs in the cloud
- Coding plotting script in Python to display model output for the Medium Range Weather App

Benefits of Contribution

- Increased accessibility for running the UFS Graduate Student Tests
- Improved visualization of Medium Range Weather App results by making it easy to run and visually appealing
- Sparks interest in young meteorologists that will hopefully contribute to the UFS in the future

MEET THE TEAM

2021 Lapenta Intern

Sam Ephraim

University of Michigan
Climate and Meteorology BSE
Computer Science BSE

Mentors

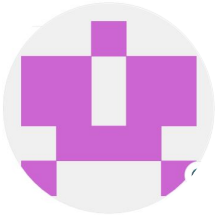
Krishna Kumar

Program Coordinator &
Senior Program Scientist
for EPIC
NOAA/OAR/WPO

Jose-Henrique Alves

Research Physical Scientist
for EPIC
NOAA/OAR/WPO

Tech Support Team



Minsuk Ji

Scientific Programmer at
IMSG
NOAA/NWS/NCEP/EMC

Dominikus Heinzeller

Scientist and Developer at GSL
NOAA/OAR/ESRL

Ligia Bernardet

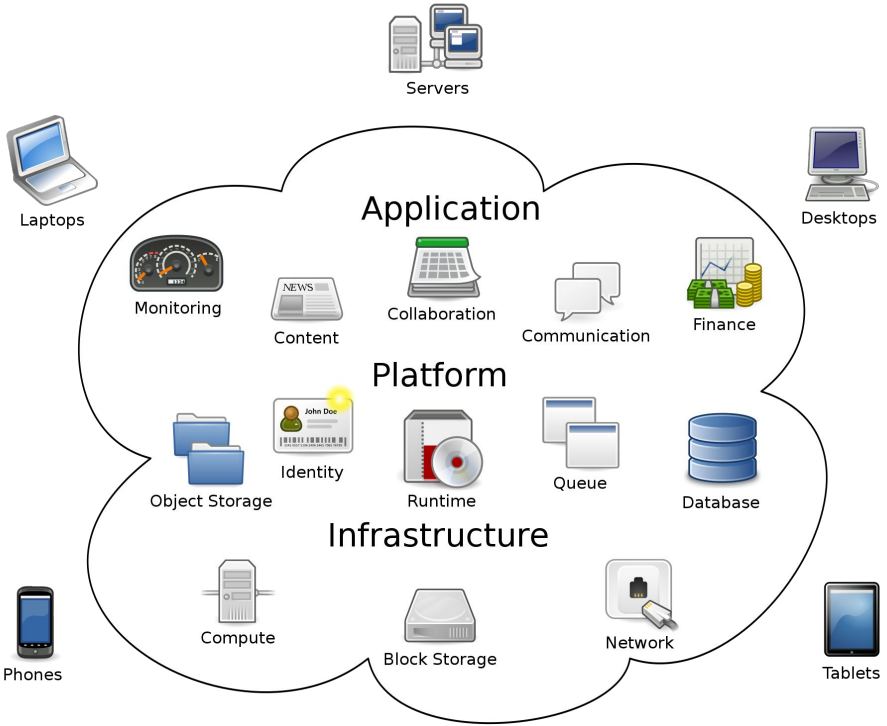
Deputy Chief of Earth Prediction
Advancement Division of GSL
NOAA/OAR/ESRL

CLOUD COMPUTING

What is Cloud Computing?
Advantages of Running UFS GSTs in Cloud

What is Cloud Computing?

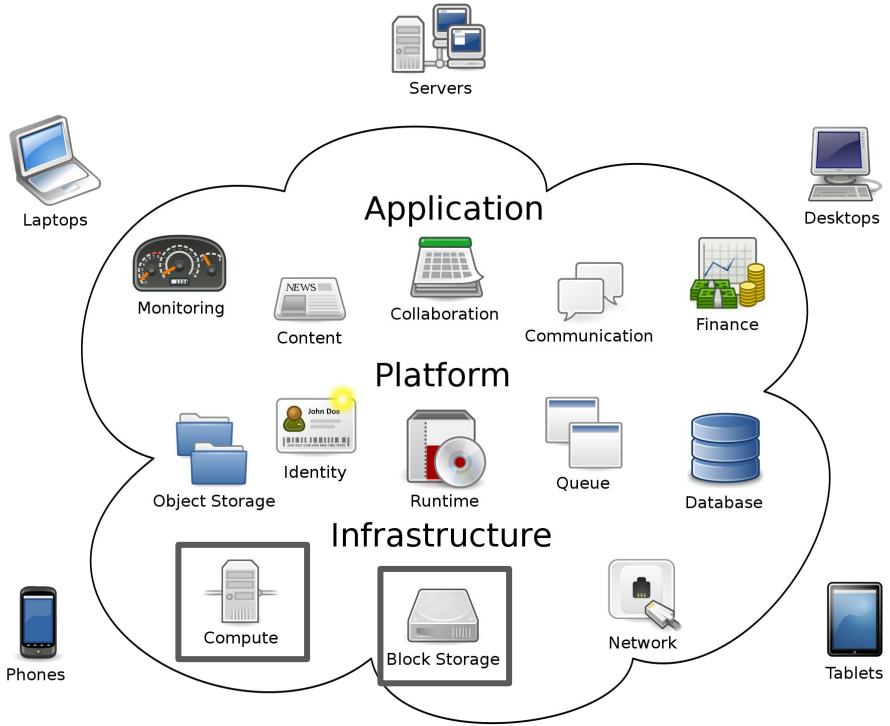
- On-demand availability of computing resources not actively maintained by the user
- Broad network access
- Pay-as-you-go model



What is Cloud Computing?

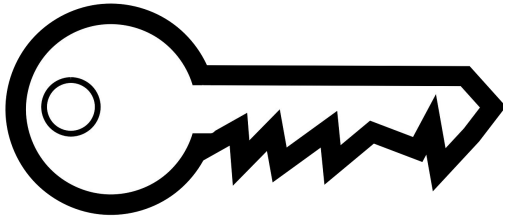
- On-demand availability of computing resources not actively maintained by the user
- Broad network access
- Pay-as-you-go model

There are many cloud service providers, but, for this project, I used Amazon Web Services (AWS)



Advantages of Running GSTs in Cloud

- Accessibility - All that's needed to run the GST is an internet connection. No High Performance Computer (HPC) or spec compliant personal computer needed
- Speed - Computing resources more powerful than personal computers leads to faster runtime
- Cost - Pay as you go pricing using AWS leads to cheap computing resources (can be <$1 per hour)

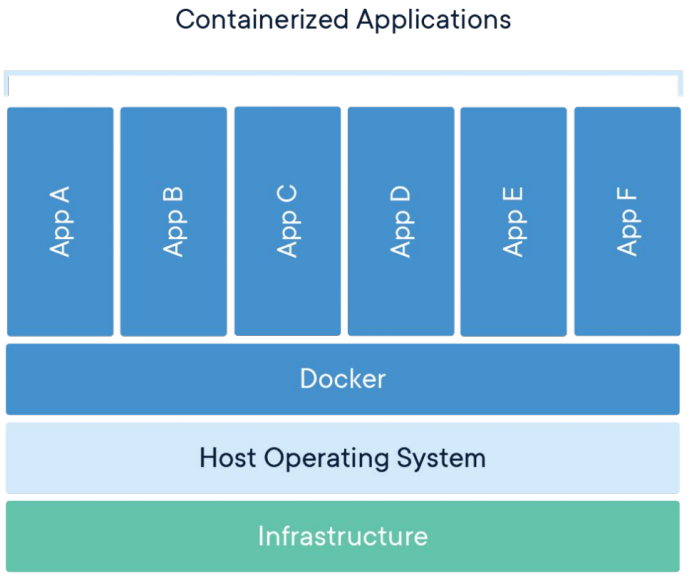


How to run code in the cloud?

One way of running code is inside of containers, a standard unit of software that contains

- All of the code
- System dependencies (libraries)

An application such as the SRW or MRW apps can run inside of a container



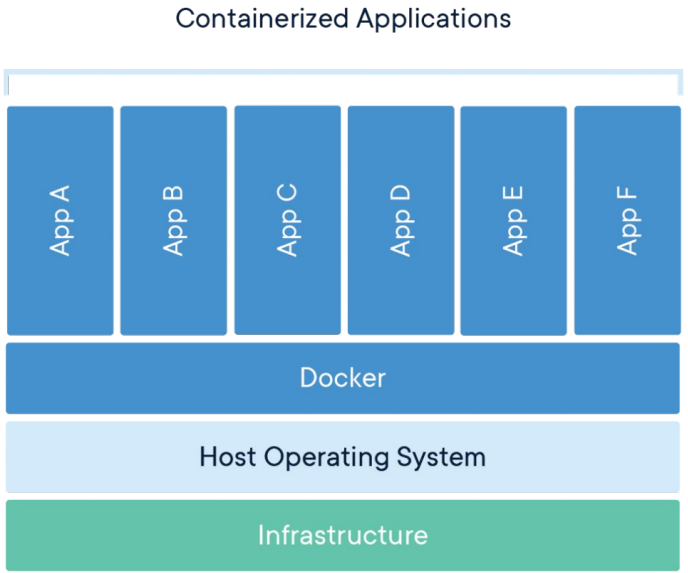
How to run code in the cloud?

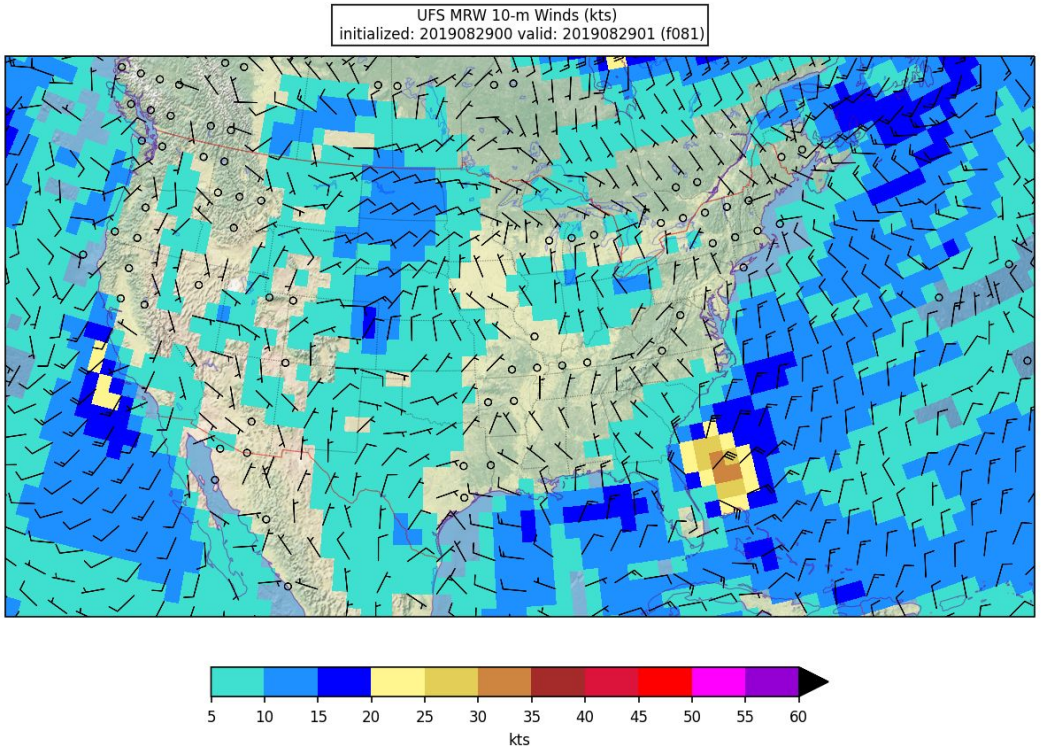
One way of running code is inside of containers, a standard unit of software that contains

- All of the code
- System dependencies (libraries)

An application such as the SRW or MRW apps can run inside of a container

For this project, I used Docker to create and run containers





COMPLETED WORK

Medium Range Weather GST

- Migration to the Cloud
- Documentation and Instructions
- New Portable Visualization Scripts

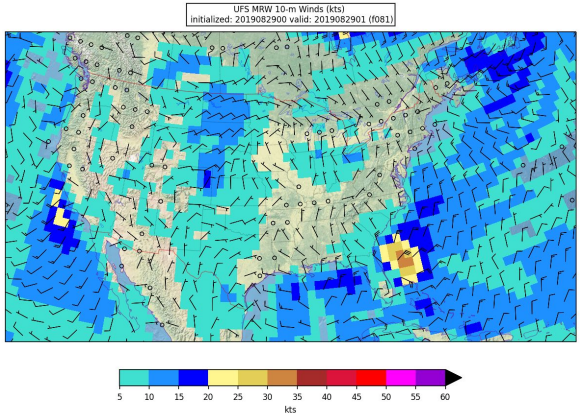
UFS Weather Model Regression Test

- Migration to the Cloud
- Documentation and Instructions

Medium Range Weather App GST

Compares cloud cover forecast during Hurricane Dorian using different cloud condensation nuclei constants

- Created documentation for running a containerized version of MRW GST in AWS
- Wrote FAQ section to help new users resolve issues
- Coded new plotting script for users to visualize experiment results



GRADUATE STUDENT TEST INSTRUCTIONS – MEDIUM-RANGE WEATHER APP IN CLOUD

Test Description

This test assesses how easy it is to access and run a UFS code, modify a physics parameter, re-run the code and then compare results. The test case is a hurricane Dorian simulation using the UFS Medium-Range Weather Application. It is an example of a code usability evaluation following the "Graduate Student Test" protocol developed for the Unified Forecast System (UFS). After completing the technical exercise, participants submit feedback on their experience through a short questionnaire. You don't have to be a graduate student to take the test – all feedback is appreciated. If you would like to take the test and submit your feedback, please register here.

In order to perform the test, you will run the Medium-Range Weather Application for 48 hours, use the workflow software to make a change to the namelist that increases the concentration of cloud condensation nucleii (CCN), and perform another 48 hour run with this change. Then you will compare the total cloud cover field before and after the change.

Cloud condensation nuclei (CCN) are small particles suspended in the atmosphere on which water vapor condenses. The number of nuclei per cubic centimeter is used in the microphysics parameterization. With more CCN in the air, cloud water is distributed over a larger number of smaller drops. This hinders the development of precipitation processes, so droplets stay suspended and cloud coverage is higher. Even after just 48 hours, an increase can be seen in the cloud cover field. As the forecast progresses, nonlinear processes take place and can change this simple interpretation.

The text before the $ in shell commands indicates where the commands should be run:

`host$` - Local computer (ie. laptop)/computer you want to save files to

`ec2$` - Inside EC2 instance

`container$` - Inside container

Medium Range Weather App GST Documentation

Walks users through steps to run the MRW GST

1. Configure Cloud Resources
2. Install Docker
3. Download and Run Container
4. Run Control Case
5. Run Experiment Case
6. Compare Outputs
7. Extra Plotting and Forecast Loop
8. Transfer Files to Local Computer

And contains Sample Output Plots and an FAQ section

Step 2: Install Docker (<5 min)

```
ec2$ sudo yum update -y
ec2$ sudo amazon-linux-extras install docker
ec2$ sudo service docker start
ec2$ sudo usermod -a -G docker ec2-user
```

Step 3: Download and Run Container (<5 min)

Pull MRW container from Docker Hub

```
ec2$ docker pull minsukjinoaa/ufs-mrweather-dorian:v1
```

If you get a permissions error, run this (`ec2$ sudo chmod 777 /var/run/docker.sock`)

Start container

```
ec2$ docker run -it minsukjinoaa/ufs-mrweather-dorian:v1 bash
```

You should now be in a shell inside the container

```
container$ su - ec2-user
container$ source .bashrc
```

Step 4: Run Control Case (3 hrs)

Go to case directory already set up for the control experiment

```
container$ cd my_ufs_sandbox/cime/scripts/DORIAN_C96_GFSv15p2
```

Change the run duration to 48 hours. Then submit the case.

```
container$ ./xmlchange STOP_OPTION=nhours,STOP_N=48
container$ ./case.submit
```

If you want to check on the progress while it is running, create a new shell in your ec2, follow **FAQ Question 1**, and navigate to the directory in the step below.

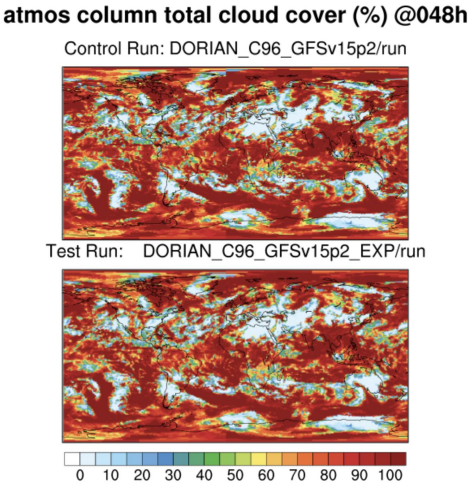
You should see the output files in the case directory for the experiment

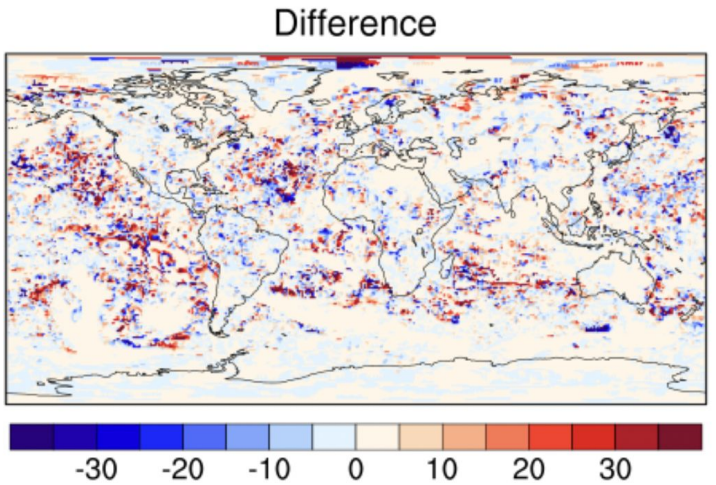
```
container$ export SCRATCH=/home/ec2-user/scratch
container$ cd $SCRATCH/DORIAN_C96_GFSv15p2/run
```

Medium Range Weather App GST Plotting

Old Plotting Routine

- Uses NCL (NCAR Command Language) which is an outdated library that is difficult to download
- Plots cloud fraction for control and experiment runs along with the difference

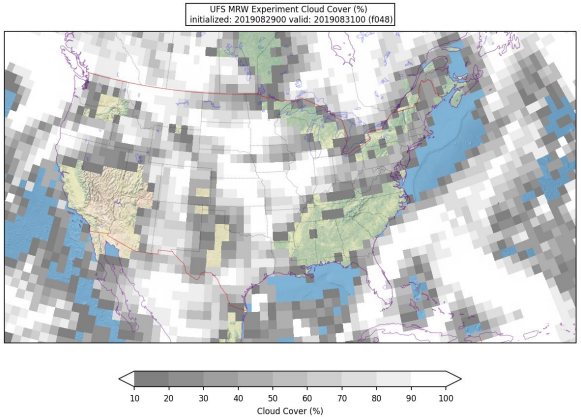


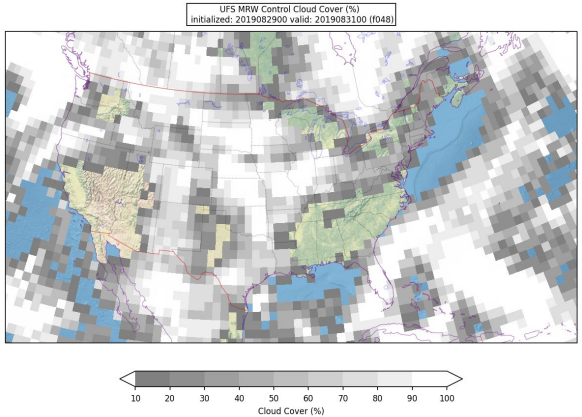


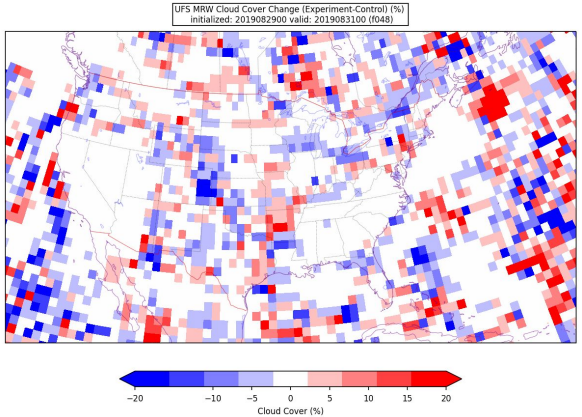
Medium Range Weather App GST Plotting

New Plotting Routine

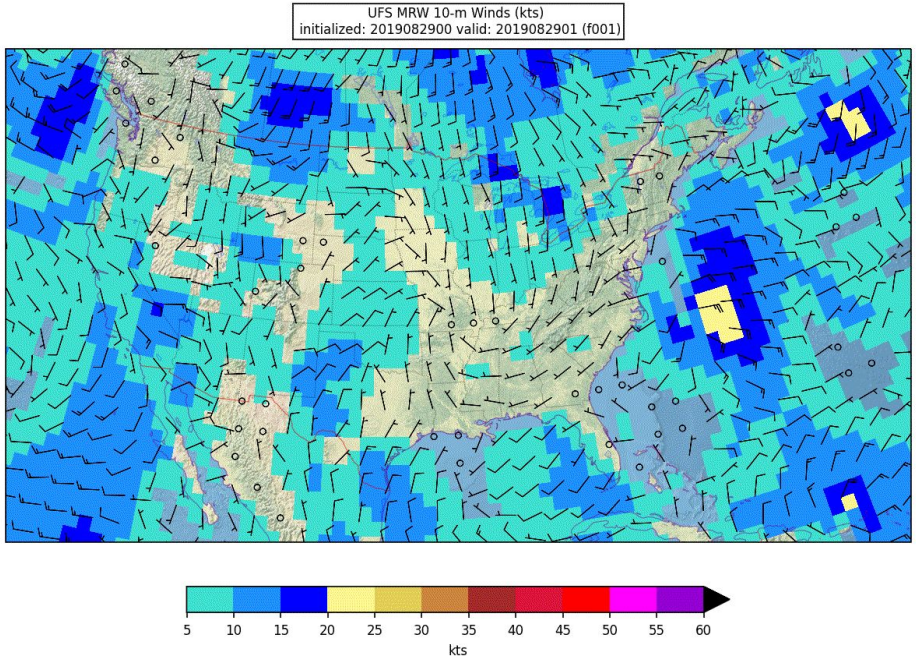
- Uses Python
- Plots cloud fraction for control and experiment runs along with the difference
- Plots 2m temperature, 10m wind, cloud fraction, and QPF along with forecast loop GIFs

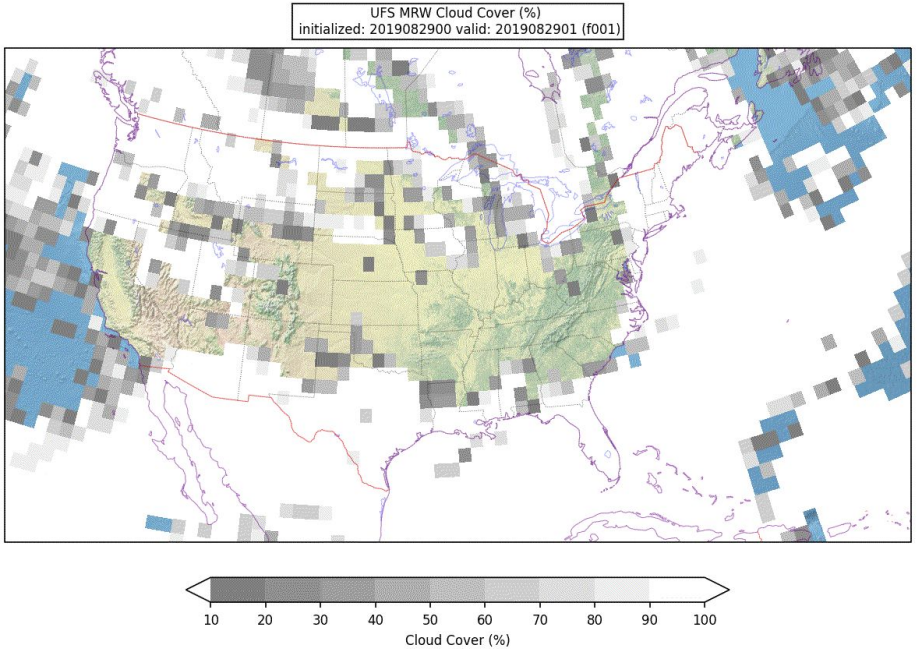




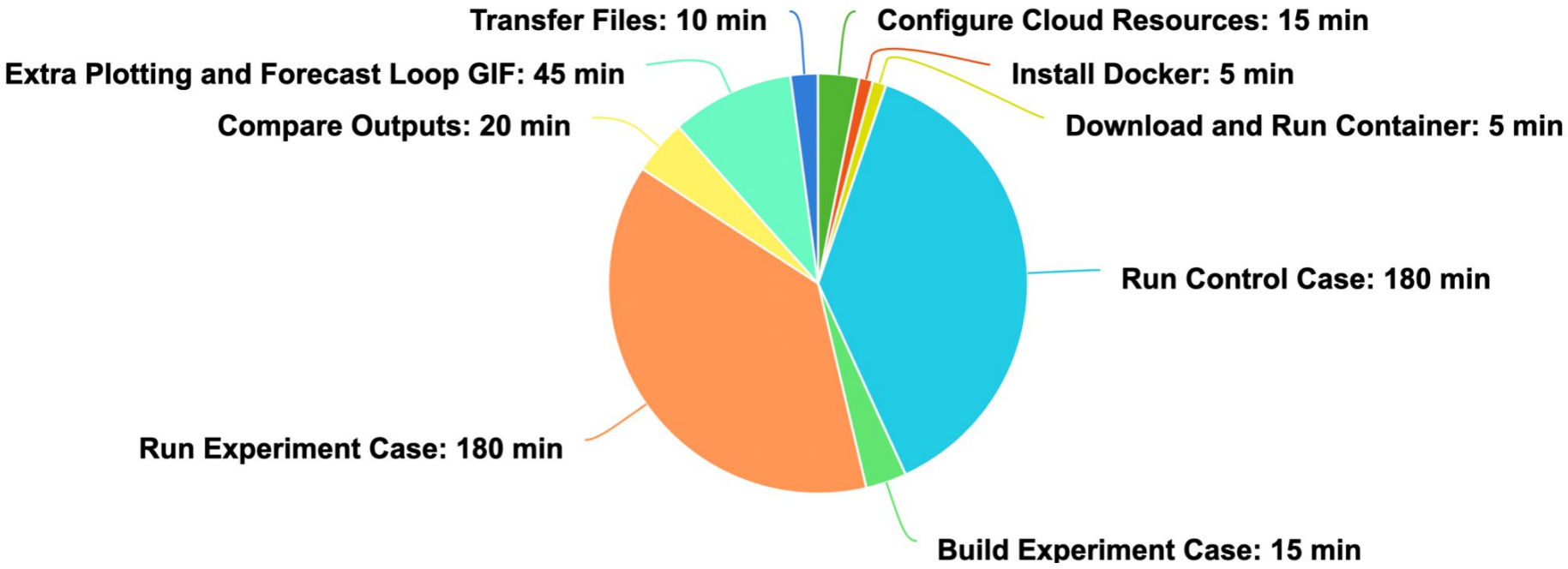


Medium Range Weather App GST Plotting





MRW GST Time Breakdown (< 8 hrs total)



MRW GST Cost Breakdown

EC2

Instance Type

- c5n.2xlarge
- $0.432 per hour
- 8 vCPUs
- 21 Gb Memory

	Hourly Cost	Total (8 hour) Cost
c5n.2xlarge instance	$0.432	$3.46
EC2 NATGateway	$0.045	$0.36
EBS Storage	Free Tier	$0.00
Total	**$0.477**	**$3.82**

NOAA Parallelworks

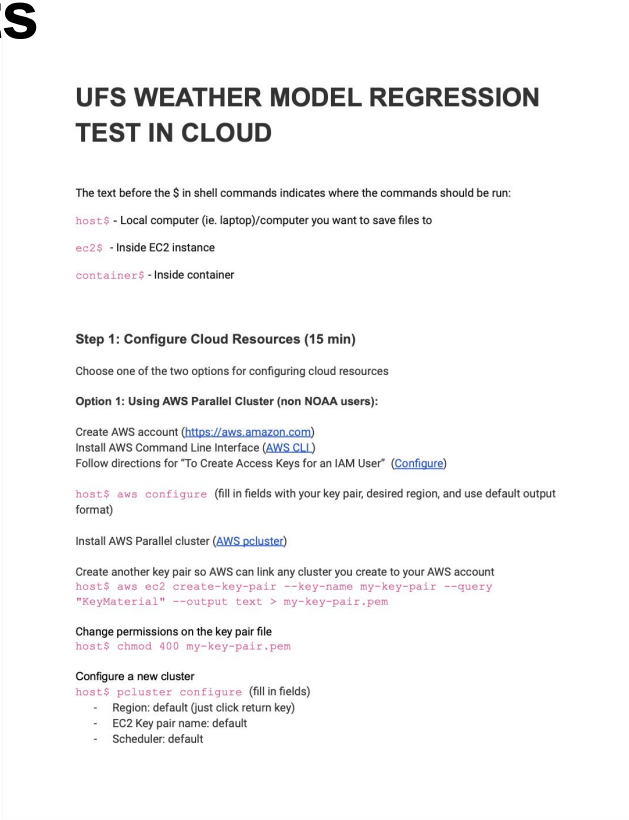
- Cloud platform through NOAA that uses AWS
  - What I used during majority of my internship

- Discovered issue with cost dashboard
  - Said I spent $7 but actually spent a couple thousand, oops
  - Reported issues upwards, that have now been resolved

UFS Weather Model Regression Tests

Runs full test suite for the UFS Weather Model

- Created documentation for running the UFS Regression Tests in the cloud
- Allows users to edit and test the UFS Weather Model in the cloud



UFS WEATHER MODEL REGRESSION TEST IN CLOUD

The text before the $ in shell commands indicates where the commands should be run:

`host$` - Local computer (ie. laptop)/computer you want to save files to

`ec2$` - Inside EC2 instance

`container$` - Inside container

Step 1: Configure Cloud Resources (15 min)

Choose one of the two options for configuring cloud resources

Option 1: Using AWS Parallel Cluster (non NOAA users):

Create AWS account (https://aws.amazon.com)
Install AWS Command Line Interface (AWS CLI)
Follow directions for "To Create Access Keys for an IAM User" (Configure)

`host$ aws configure` (fill in fields with your key pair, desired region, and use default output format)

Install AWS Parallel cluster (AWS pcluster)

Create another key pair so AWS can link any cluster you create to your AWS account

```
host$ aws ec2 create-key-pair --key-name my-key-pair --query
"KeyMaterial" --output text > my-key-pair.pem
```

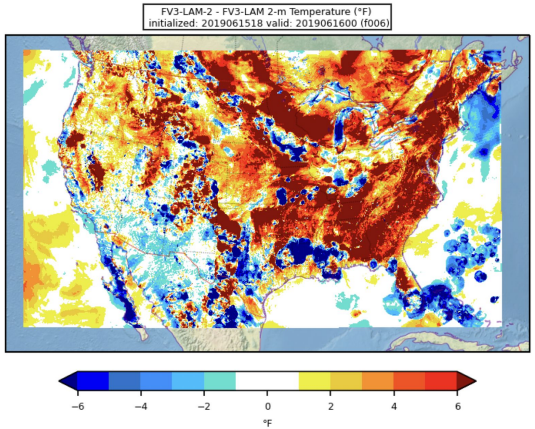
Change permissions on the key pair file

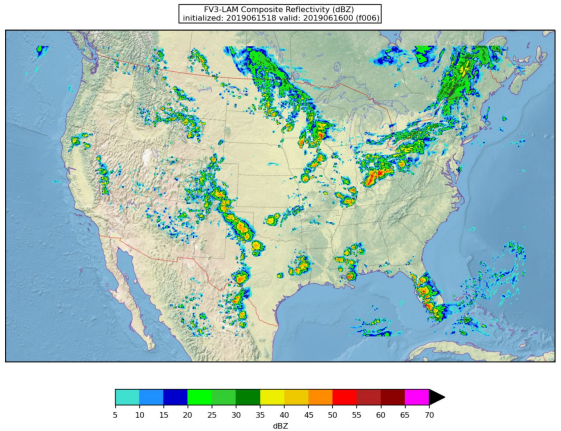
```
host$ chmod 400 my-key-pair.pem
```

Configure a new cluster

`host$ pcluster configure` (fill in fields)

- Region: default (just click return key)
- EC2 Key pair name: default
- Scheduler: default





IN PROGRESS WORK

Short Range Weather GST

- Migration to the Cloud
- Documentation and Instructions

Short Range Weather App GST

Compares the model runs of the UFS using different grid resolutions (25km and 3km) and physics schemes (GFSv15p2 and RRFSv1alpha)

- Created Docker container for the SRW GST 25km resolution control case
- Created documentation for running a containerized version of SRW GST in AWS for the control case
- Wrote FAQ section to help new users resolve issues

GRADUATE STUDENT TEST INSTRUCTIONS – SHORT-RANGE WEATHER APP IN CLOUD

This test assesses how easy it is to access and run a UFS SRW Application configuration, modify the configuration to run a getting started step and two experiments:

1. Different horizontal resolution, and
2. New physics suite definition file (SDF).

You will be rerunning the code each time, and then comparing the results.

The test case used here is a two-consecutive day period with significant severe weather across the Contiguous United States (CONUS) from 15-16 June 2019. After completing the technical exercise, participants submit feedback on their experience through a short questionnaire. You don't have to be a graduate student to take the test – all feedback is appreciated!

If you decide to take the test and have not yet registered, please register and submit your feedback here.

In order to perform the test, you will start by running the default SRW Application configuration using the GFSv15p2 suite definition file (SDF) for 48 hours starting at 00 UTC on 15 June 2019 on a 25-km predefined CONUS domain to establish a control experiment. Once that is successful, two additional experiments will be conducted.

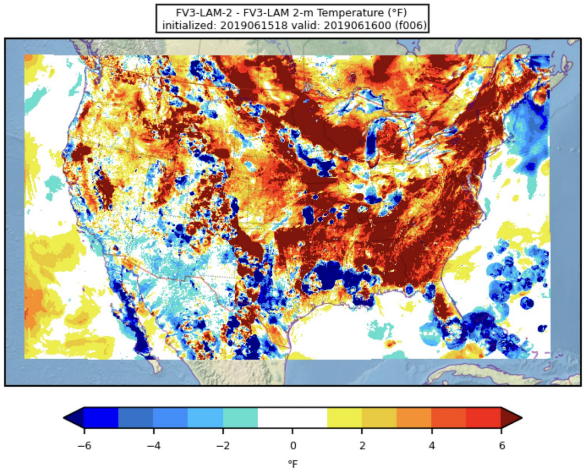
First, you will change the configuration to run a 12-hour forecast starting at 18 UTC on 15 June 2019 on a 3-km predefined CONUS domain, still using the GFSv15p2 SDF. Next, you will run the same 12-hour forecast on the 2-km predefined domain, but this time you will use the RRFSv1alpha SDF. Working your way through these examples will assist you with learning how to change these settings and conduct new experiments. Python scripts are available to plot a variety of files from each run as well as difference those fields between two runs.

Important Note: If you are using a laptop for running the SRW App you will need a minimum of 4 Gb of memory and at least 40 Gb of disk space to run a 25 km resolution CONUS case. Running a 48 hour simulation will likely take several hours and thus you may wish to run 12 hour simulations instead. In addition, if you would like to run a 3 km resolution CONUS case you will need at least 24 Gb of memory and additional disk space depending on your forecast output frequency.

Short Range Weather App GST Challenges

Unable to run the experiment cases that use 3km resolution

- Segfault due to too little computational power of using a single node
  - Using Docker container constraints model run to be on a single node
  - Hypothesized solution is to use a container service such as Docker Swarm that can run on multiple nodes



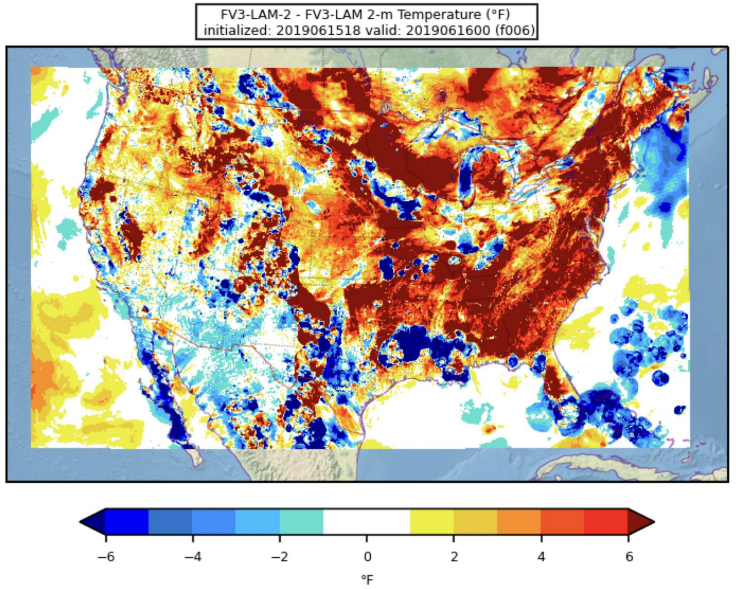
Created documentation of the problem along with a compilation of forum post and email chain responses from a variety of experts on the UFS SRW

Accomplishments

- First person to complete the MRW GST in the cloud
- Created documentation to run MRW GST in the cloud
- Created documentation to run the UFS Weather Model regression tests in the cloud
- Created documentation to run the SRW GST control case in the cloud
- Documented challenges and potential solutions to running experiment case (3km resolution) from SRW GST in the cloud

Future Work

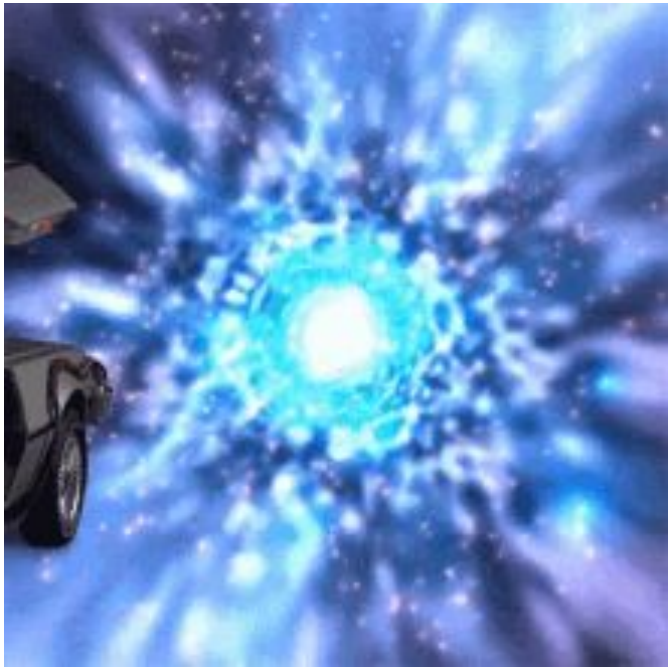
- Finish SRW GST to work for 3km resolution
- Test running the GST on various cloud service providers such as Azure and Google



Azure

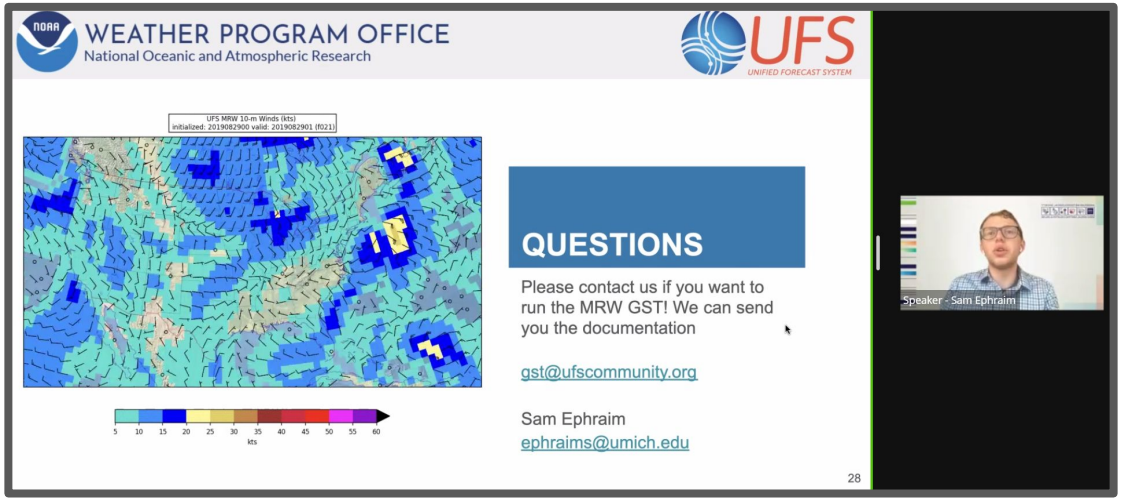
Google Cloud

Fast Forward 1 Year

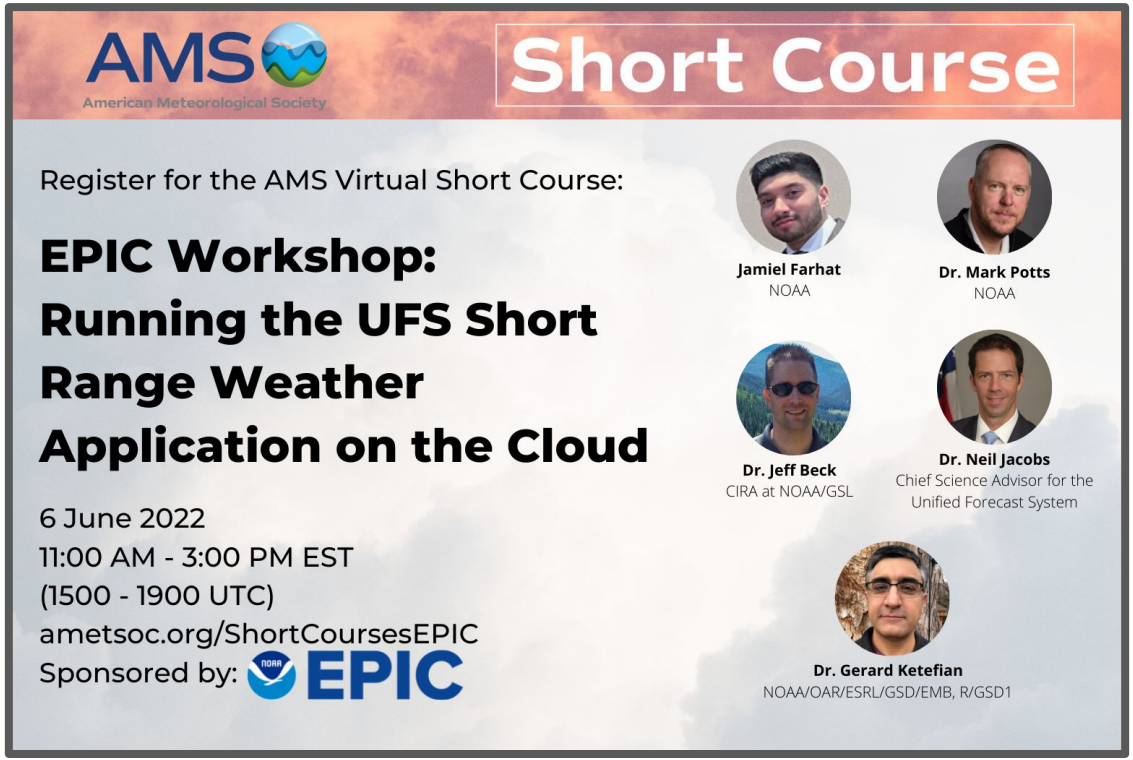


AMS Presentation

- Presented abbreviated version of this presentation to the First Symposium on Earth Prediction Innovation and Community Modeling



AWS Short Course



AWS Short Course

EPIC
EARTH PREDICTION INNOVATION CENTER

Objective: I can use the Short-Range Weather Application to run, modify, and compare forecast outputs.

Jamiel Farhat - Technical Training Specialist - ?'s about EPIC

Dr. Neil Jacobs - Chief Science Advisor for the UFS - ?'s about UFS

Dr. Mark Potts - Lead Cloud Computing Engineer - ?'s about Accessing the Cloud

Dr. Jeff Beck and Dr. Gerard Ketefian - SRW App Developers - ?'s about the SRW App

UFS Short-Range Weather Application 2.0

New Capabilities

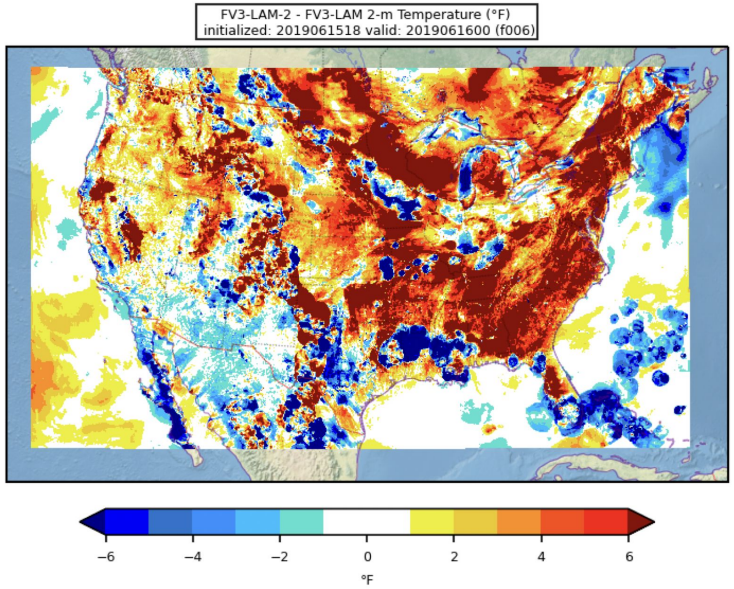
- Verification package (METplus) for deterministic and ensemble simulations
- Support for four Stochastically Perturbed Perturbation (SPP) schemes

Future Releases

- Data assimilation
- Forecast restart/cycling

Future Work

- ~~Finish SRW GST to work for 3km resolution~~
- Test running the GST on various cloud service providers such as Azure and Google

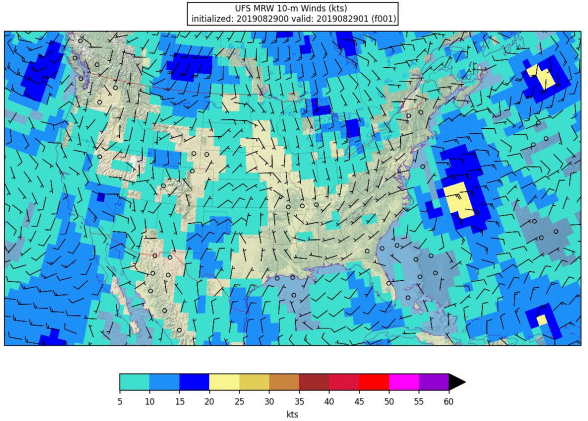


Azure

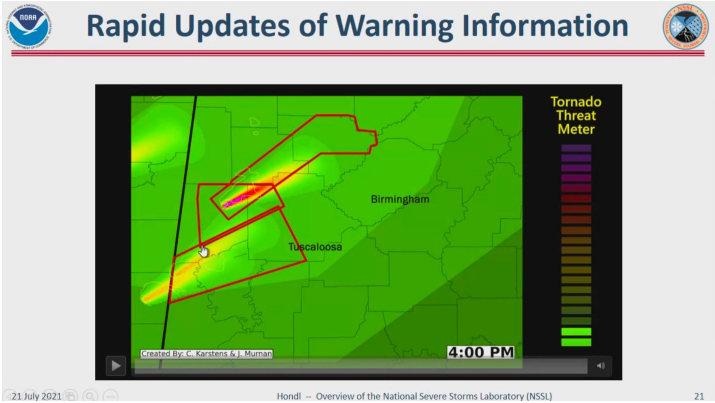
Google Cloud

My Future

- Favorite tasks are creating plots and analyzing data
  - Within EPIC/my project, that was creating the new MRW plotting script



  - My favorite office visit during internship was by NSSL
    - Talked about developing new and testing performance of new forecast products



MRW Links

Medium Range Weather App Documentation

Medium Range Weather GST (in AWS)

- Link shared with all NOAA users, by request for others

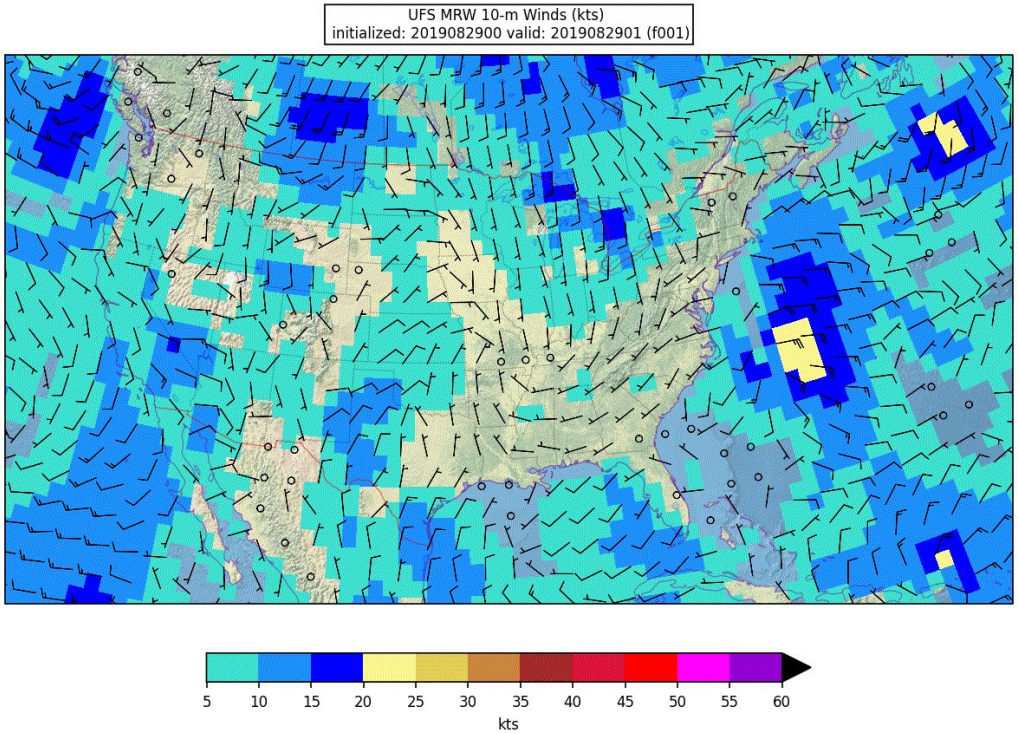
Medium Range Weather GST (not in cloud)

SRW Links

Short Range Weather App Documentation

Short Range Weather App Workshop

- Includes GST walkthrough
- EPIC running workshop at AGU Fall 2022



QUESTIONS

Please contact us if you want to run the MRW GST! We can send you the documentation

support.epic@noaa.gov

Sam Ephraim
ephraims@umich.edu
samuel.ephraim@rsmas.miami.edu