Collaborative WRF-based research and education enabled by software containers

J. Hacker, J. Exby, K. Fossell

National Center for Atmospheric Research

Contributions from Tim See (U. North Dakota)

BIG WEATHER WEB

A common and sustainable big data infrastructure in support of weather prediction research and education in universities

1. **Large ensemble distributed over 7 universities:**

2. **Common storage, linking, and cataloging methodology**
   - Permanent naming and high availability of data and experiments
   - Connecting data, platform, tools, analysis

3. **Software Container technologies for easy deployment and reproducibility**
   - Self-contained: software can be instantly deployed in common environments
   - Naming and versioning: compact reference mechanisms for complex environments
   - Good for reproducibility and education

Carlos Maltzahn, **UC Santa Cruz (lead)**
Ivo Jimenez, UC Santa Cruz
Mohan Ramamurthy, **Unidata**,
Josh Hacker, John Exby, **NCAR**
Robert Fovell, Kevin Tyle, SUNY Albany
Gretchen Mullendore, UND
Brian Ancell, Texas Tech
William Capehart, SDSM
Clark Evans, UW Milwaukee
Steven Greybush, Penn State
Russ Schumacher, CSU

Website: bigweatherweb.org
Email list: bww-users@unidata.ucar.edu

Why the Weather Research and Forecasting (WRF) model in Docker?

- WRF is a state-of-the science numerical weather prediction (NWP) model for operations and research
- Compilation and execution can be an intensive effort, slowing time to results
  - Huge complex code
  - Numerous and non-trivial dependencies
  - Inexperienced users can take months to get WRF running for results
- Classroom opportunities for hands-on numerical weather prediction can be intensive to produce
- Research is almost never reproducible
- Collaboration is difficult and cumbersome

Goals of container-wrf

NCAR

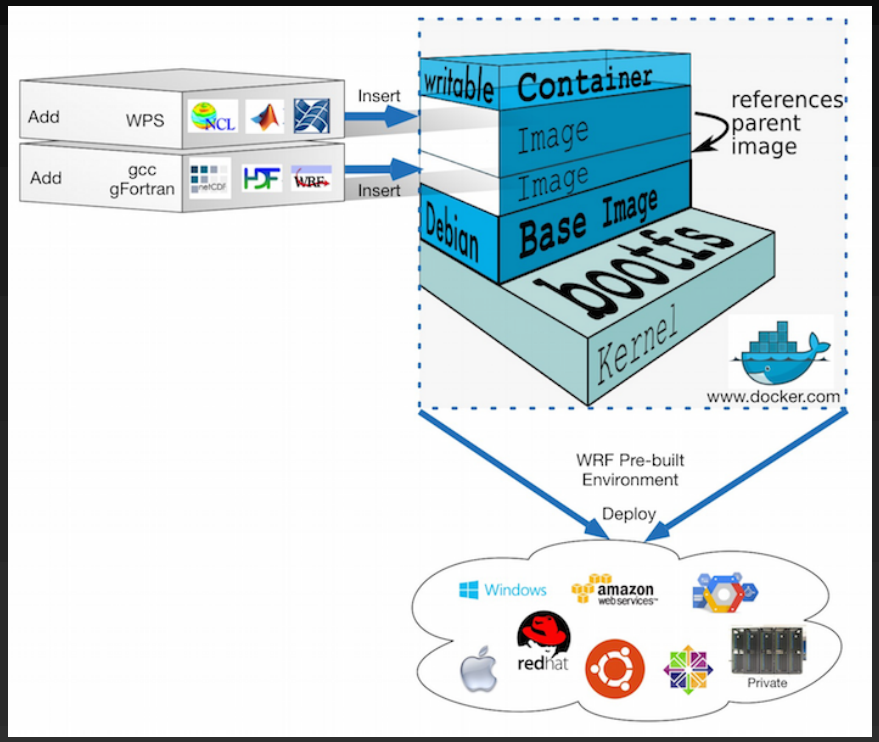
- Lower the technical difficulty for new users
  - Graduate students can accomplish early results simultaneous to learning the Unix/Linux skills needed for more in-depth work
  - Provide a reference context
- Trivialize classroom and lab experimentation
- Provide a platform for reproducible numerical weather prediction research
- Facilitate efficient and easy collaboration

WRF in a container is a grey box, open for changes.

WRF in a container

Build, Deploy, Run

NCAR



- Docker provides stable container software for packaging applications
- Container includes minimal environment, all dependencies needed for an application

Open and Adaptable

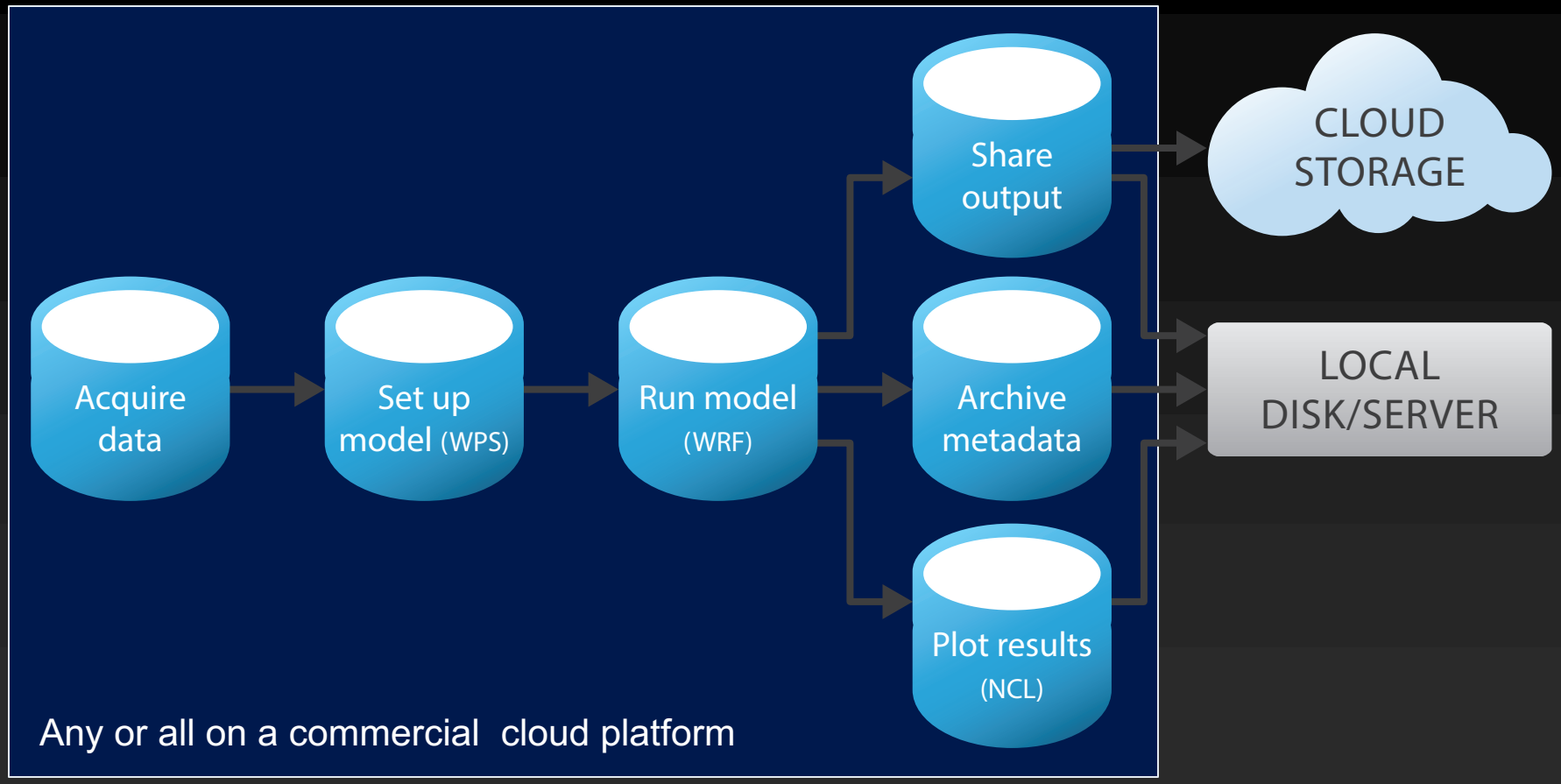
NCAR

- Important use cases in an education
  - Change input data sets for land use
  - Change input data sets for initial and boundary conditions
  - Change physics, diffusion, time steps, parameters
  - Change code and recompile in known environment

Container development allows all of these, including deployments on cloud providers or local compute hardware.

Vision: End to end

NCAR



Run on your command line and link to your filesystem

Bit-wise reproducibility

NCAR

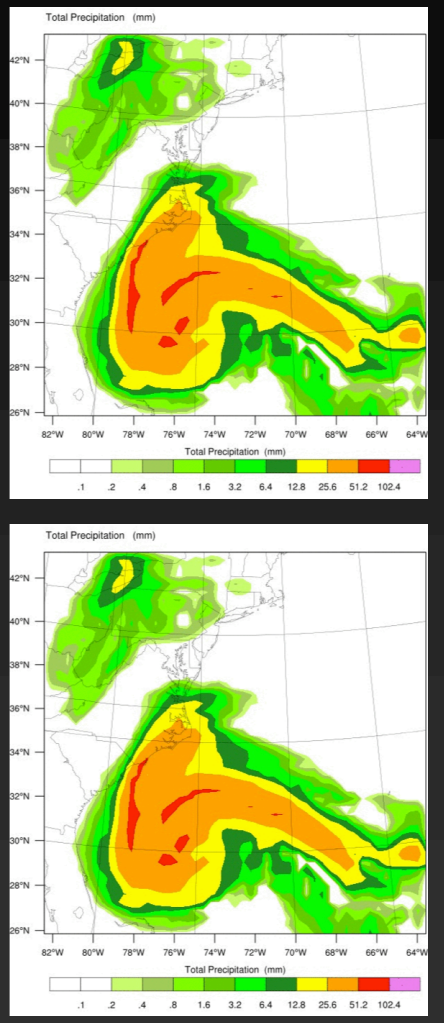
Incomplete list of test platforms

OS	Chip/CPU	Cores
OS X 10.10.5	Intel Xeon E5	6
OS X 10.9.5	Intel Core i7	2
Ubuntu 14.04	AMD Opteron 6320	16
Ubuntu 14.04	Intel Xeon E5	16
RHEL	Intel Xeon X5550	8
*Ubuntu 14.04	Intel Xeon E5-2666 v3	32
**Ubuntu 14.04	Intel Atom C2550	4

AWS EC2

Packet Tiny Atom

Reproducible on all platforms so far

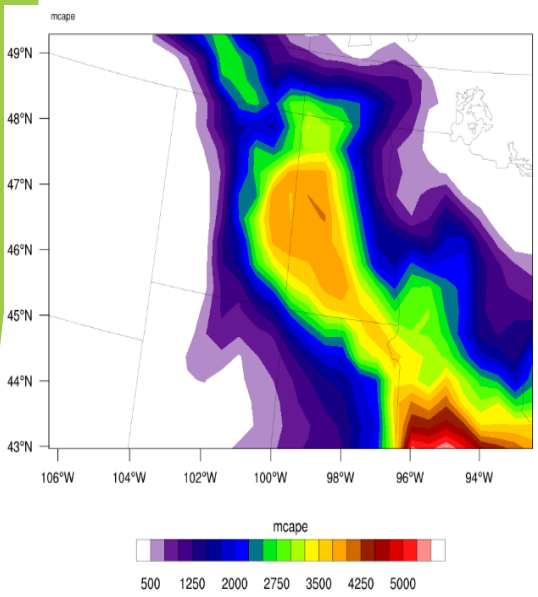


From Tim See, (UND)

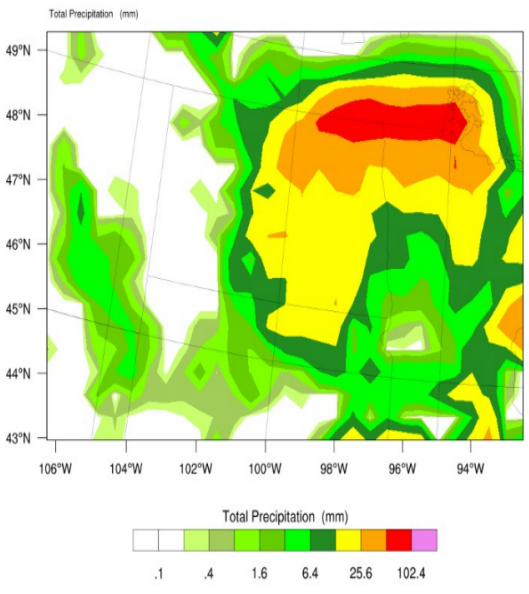
University of North Dakota Classroom Implementation

- Access Docker-WRF through Amazon Web Services.
- Students completed a classroom assignments to create an ensemble output of a tornadic supercell over North Dakota.
- Students personally changed the parameterization schemes within WRF.
- Classroom discussion generated through changing of parameterizations.
- Sample Plots below

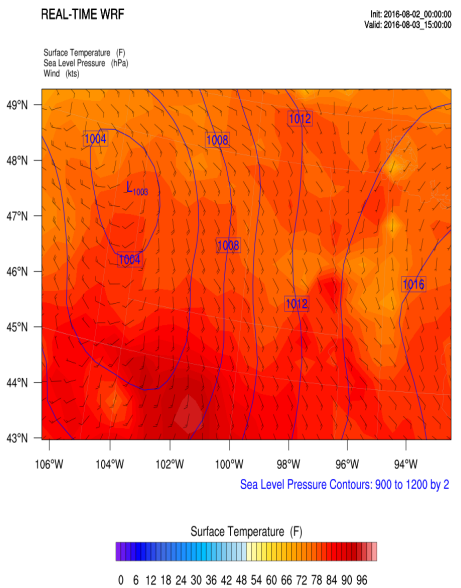
CAPE



Accumulated Precipitation



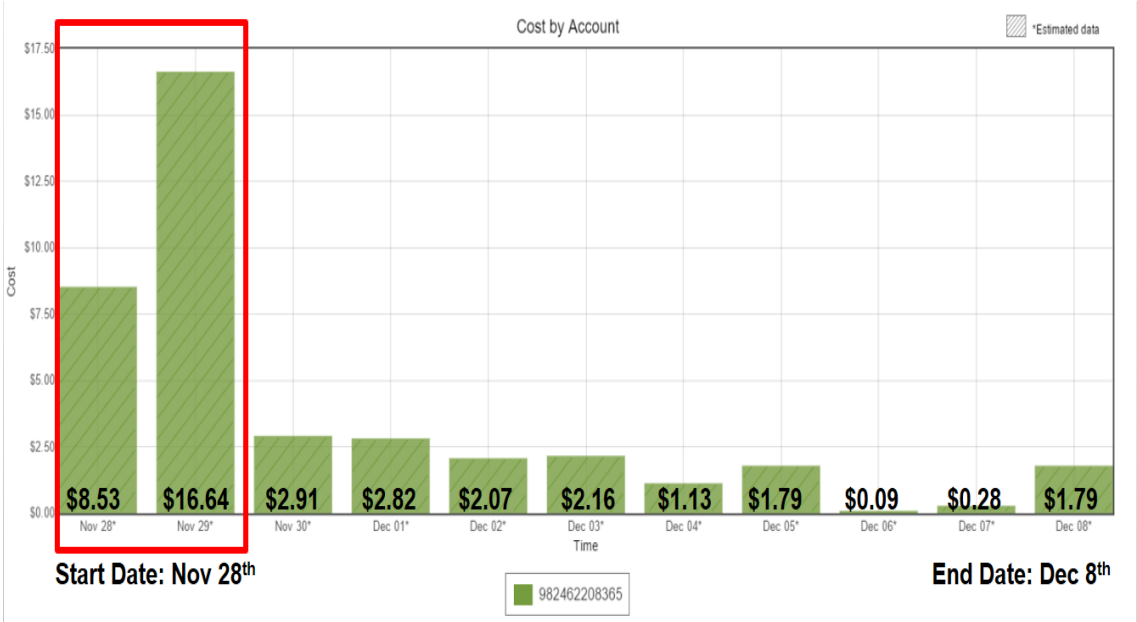
Surface Analysis



University of North Dakota Classroom Implementation

From Tim See, UND)

- Total Cost for Homework Assignment: $40.21 over 11 days
  - Inflated cost due to not shutting down instances properly first day.
- Reproducibility of Docker allows for plots to remain the same across all students.
- Students were able to complete their model runs from personal laptops.



Zero to WRF Tutorial
in 3 Steps

NCAR

Using a relatively modern, multi-core laptop, workstation, server. Linux, MacOS, Windows 10:

- 1. Install docker for free ! https://docker.com/products/
- 2. git clone https://github.com/NCAR/container-wrf
  (Hurricane Sandy and Katrina)
- 3. cd 3.7.1/demos/local ; docker-compose up
  - (psssst. Windows users- please first edit docker-compose.yml for output DIR)

VOILA!

Demo

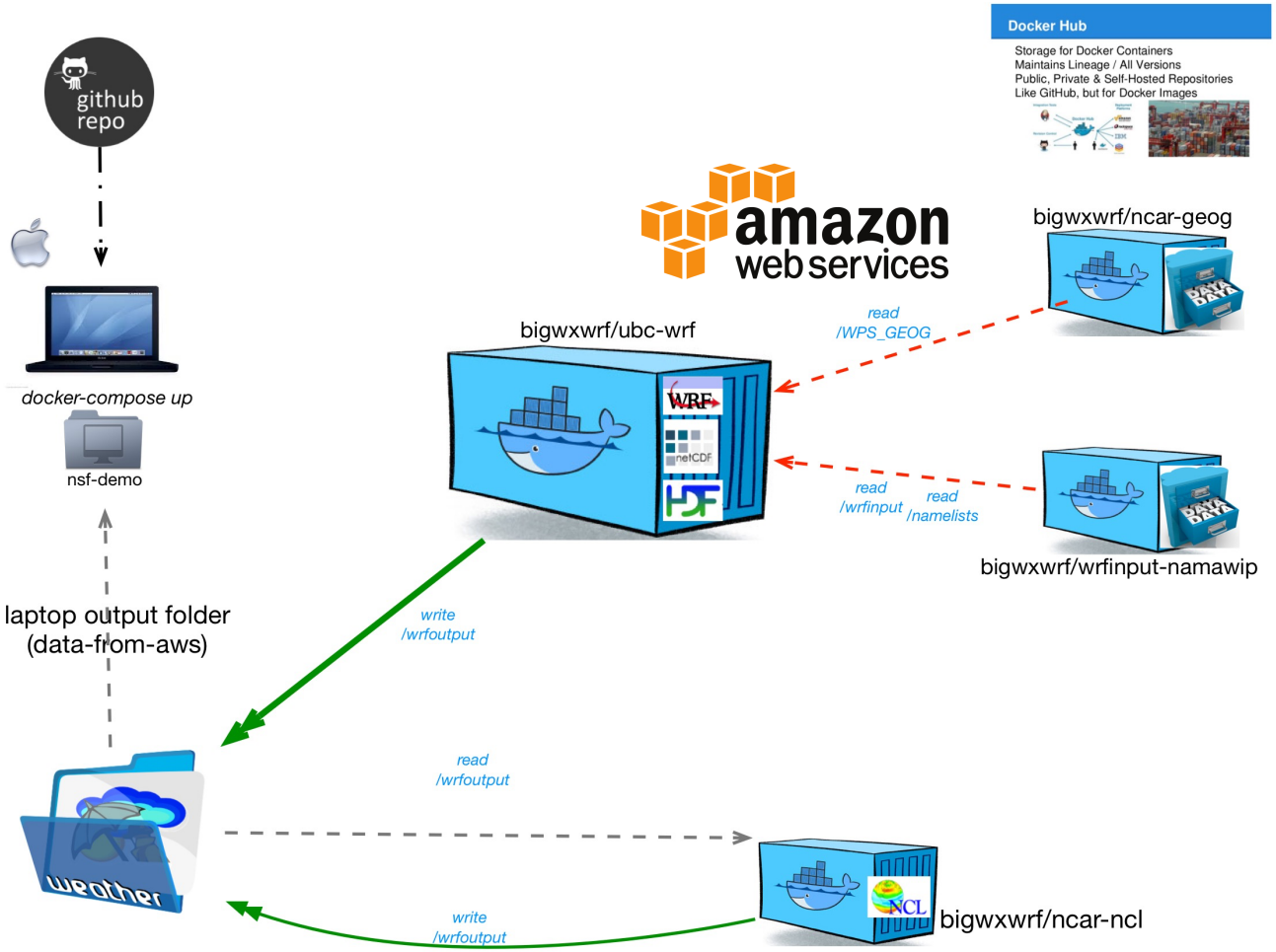
- A University professor gives their student $1 dollar to “do some atmospheric science” in one hour, then write a paper to help toward graduation!

- Is this possible?

- 🏫👨‍🏫🤝$1💵👉👩‍🔬⏩AWS☘️◻️🆔💡🌐🌧️🌤️📈📉🎛️(10🔁⌛) 💾🖋️📄📄📄🏁🎓🍺💤

Tutorial Architecture

NCAR



32 core WRF run on AWS

NCAR

- script to spin up a new AWS resource (spot EC2 under $1)
- our own (ssh secured) compute access

- script to launch the docker-compose.yml elements:
  - two containers holding data
  - a container with wrf executables
  - a container with NCL scripts to post process *.nc
  - copy files from AWS back to macbook

- visualize results on macbook

Live Demo Time

NCAR

- Fingers crossed that wifi is fast!

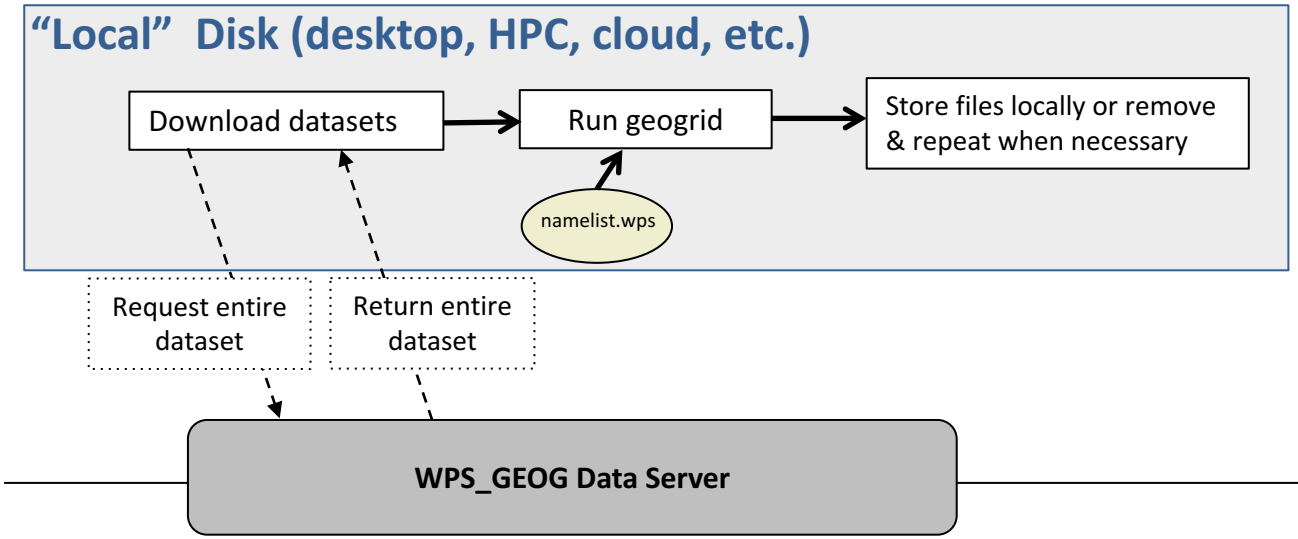
Next Steps

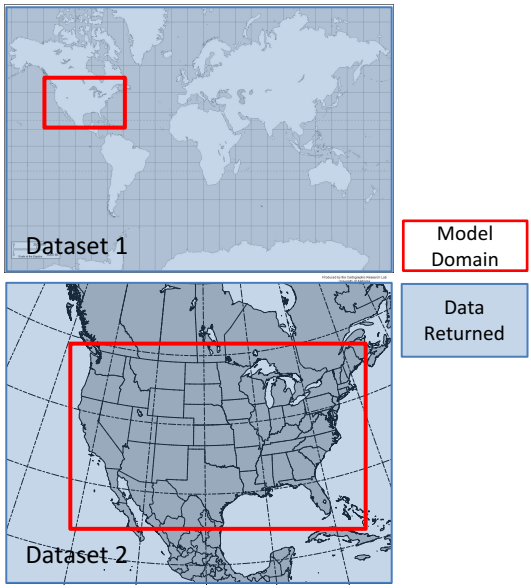
NCAR

- Dynamic data query for smaller downloads
- AWS Batch
- Singularity for performance
- Swarms and queues for scaling, workflow automation.

Dynamic pull of terrestrial data sets (WPS_GEOG) – work in progress

Current practice



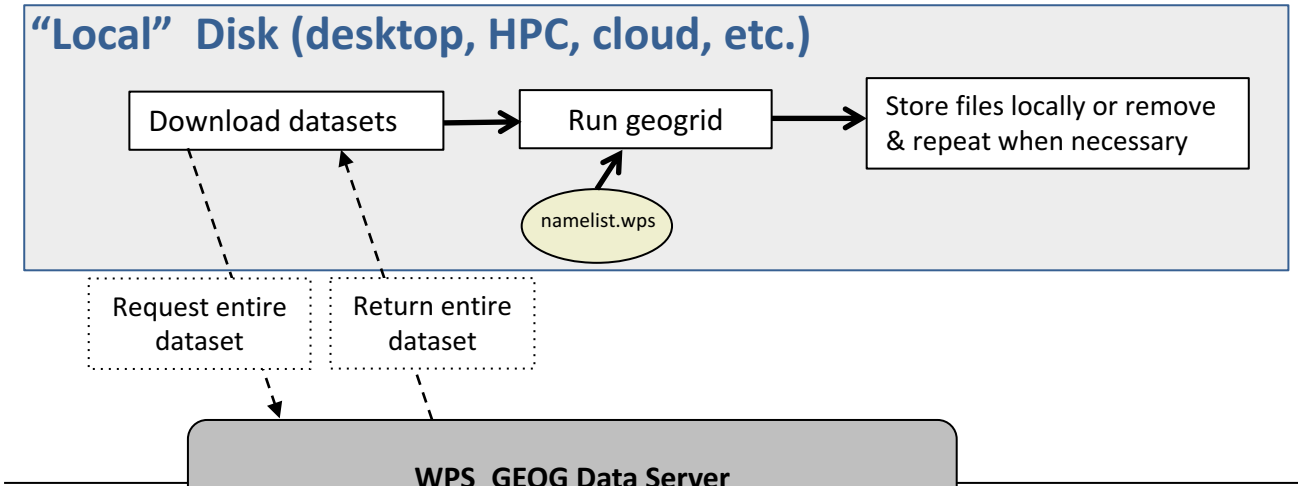


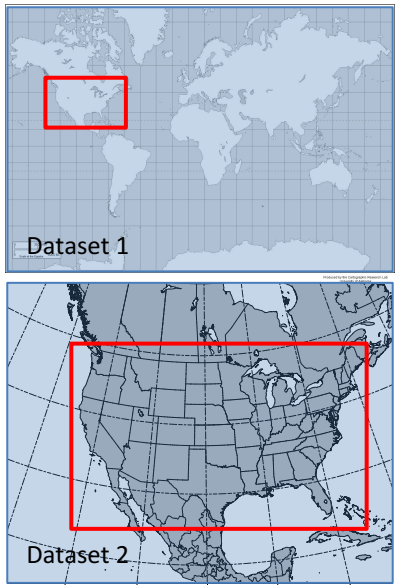
Downsides:

- Required to download and store entire datasets and unnecessary data
- Not conducive to cloud or container environments
- Large files to store and transfer
- Costs ($) associated with storing or downloading/transferring data in cloud
- Computational inefficiencies in containers due to size of files

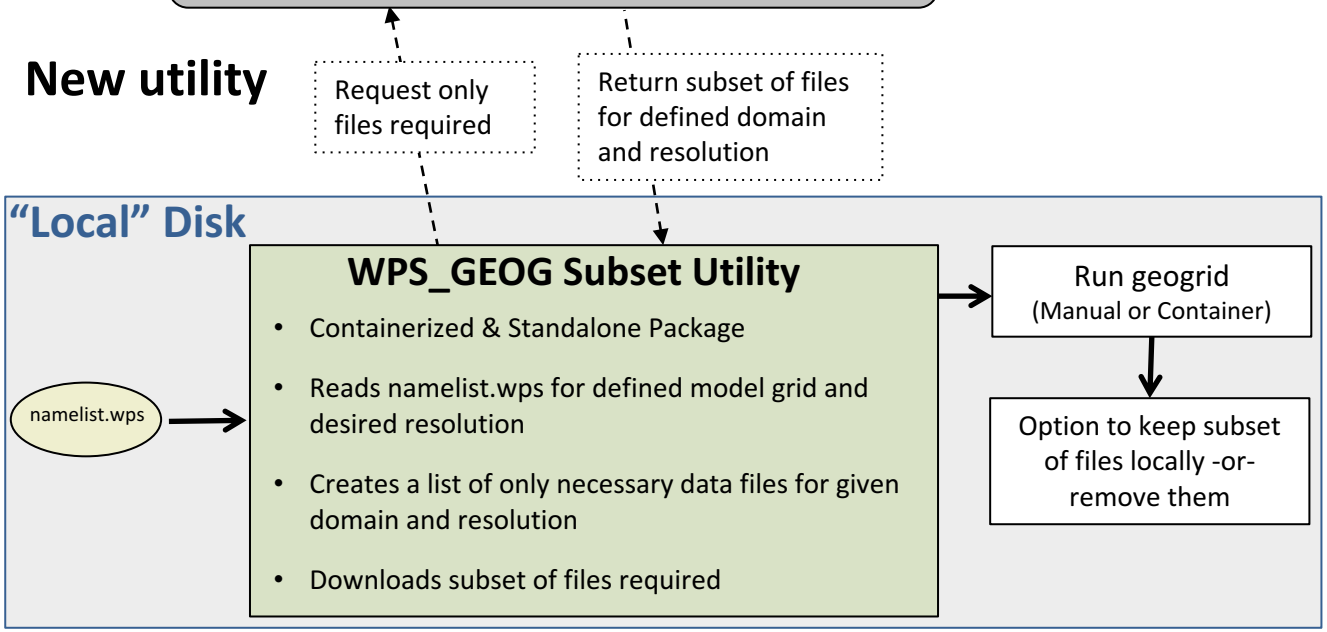
Dynamic pull of terrestrial data sets (WPS_GEOG) – work in progress

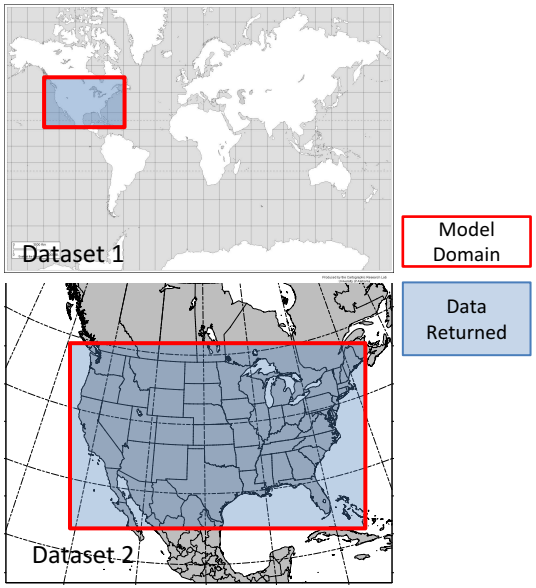
Current practice





New utility





NCAR Resources

NCAR

- NCAR RAL Container-WRF Project Web site:
- https://www.ral.ucar.edu/projects/ncar-docker-wrf

- NCAR Github repo: https://github.com/NCAR/container-wrf

- NCAR Dockerhub repo: https://hub.docker.com/r/bigwxwrf/
  - bigwxwrf/ncar-wrf
  - bigwxwrf/ncar-wpsgeog
  - bigwxwrf/ncar-wrfinputkatrina
  - bigwxwrf/ncar-ncl
  - bigwxwrf/ncar-wrfinputsandy
  - bigwxwrf/ubc-wrf

- Slack channel for docker-wrf community discussion.
- https://ncar-dockerwrf.slack.com