Rewards for results, not data

Current practice:
NWP groups at universities use supercomputing time to

Poor reproducibility of data-intensive science
Impact on education and research

Unnecessary duplication of work, steep learning curves

Impaired availability of intermediate results
Limited ability to adopt new technologies

Communities of practice are falling behind

Domain: Numerical Weather Prediction
Informed by EarthCube Users workshops

- Create large ensembles
- Allow scientific reproducibility and

Docker offers open-source container
software for packaging applications
- Container includes minimal OS and all
dependencies needed for an application

WRF, processing, and post-processing precompiled binaries run under any
Docker engine
- Time for new virtual machine (VM) to instantiation: 1 minute
- Full WRF output and graphics completed (macbook pro, 2 cpu): 5
minutes 30 secs

- Allows scientific reproducibility and
application portability from laptops to
severs to clouds. (AWS, Google, Azure)
- Reproducible results across many
platforms
- Control over compiler uncertainty

Big Weather Web Nuclei
1. Large ensemble distributed over 7 universities:
Gretchen Mullendore (UND), Brian Ancell (Texas Tech), William
Capehart (SDSM), Clark Evans (UW Milwaukee), Robert Fowell (U
Albany), Steven Greybush (Penn State), Russ Schumacher (CSU).

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

● Establish “nuclei”: pieces of technology that
- Are easily shareable
- Have the ability to grow &
improve over time
- Ensure “buy-in” from
researchers and students

● Examples:
- Wikipedia
- Linux kernel

● Infrastructures to enable
community-driven review and improvement

Next
- Investigate cloud commons governance models
- Create domain-specific Ansible Roles
- Create tutorials and workshops to achieve
community adoption
- Control everything from Jupyter notebooks!

Web site: http://bigweatherweb.org
Public email list: bww-users@unidata.ucar.edu
Contact: carlosm@ucsc.edu
Funded by NSF Award# 1450488.