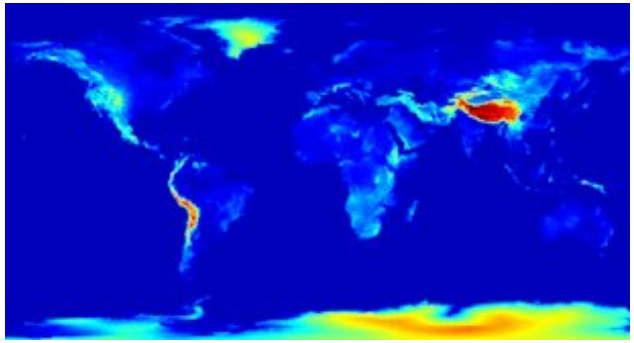
# NOAA Modeling Portal

Jeff Smith and Mark Govett
Earth System Research Lab / NOAA / USA
November 6, 2008

**Earth System Research Lab in Boulder, Colorado**

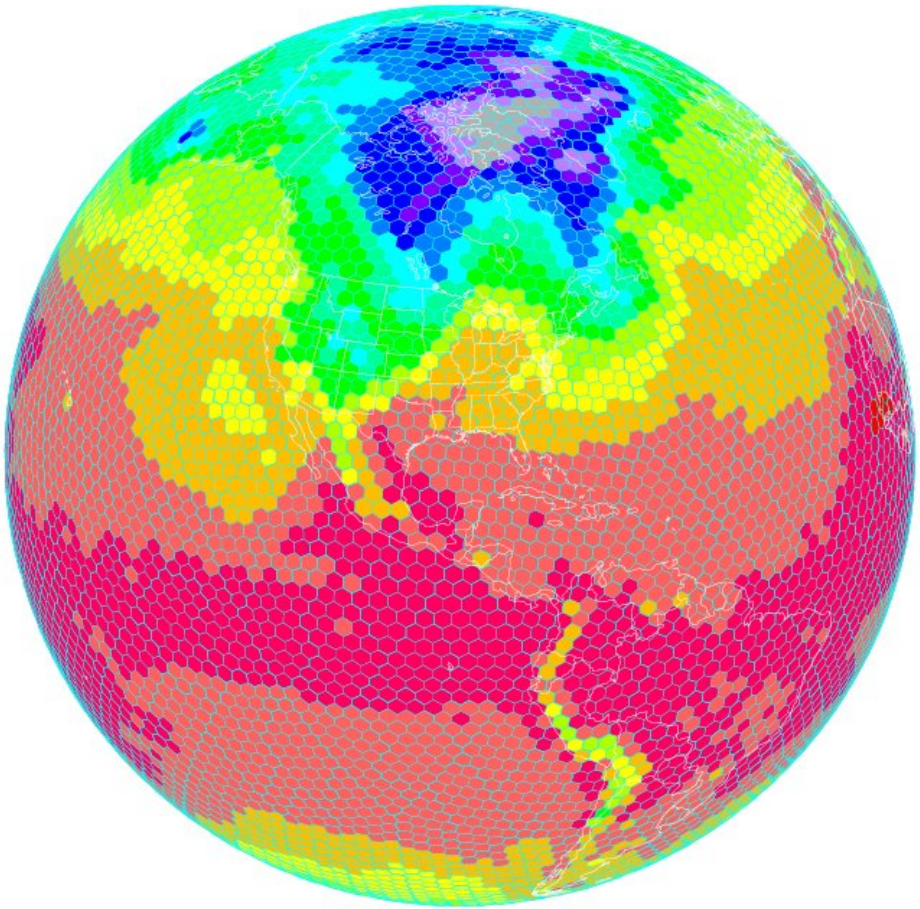
# What is WRF?



- WRF is NCAR's mesoscale and global Weather Research and Forecasting model
- Designed for both operational forecasters and atmospheric researchers
- Features multiple dynamical cores, a 3-dimensional variational (3DVAR) data assimilation system, and an extensible software architecture that supports parallel computing
- WRF is currently in operational use at NCEP (National Centers for Environmental Prediction)

# What is FIM?

- FIM (Flow-following, finite-volume, Icosahedral Model)
- NOAA's global model that employs adaptive isentropic-sigma hybrid vertical coordinate accurate finite-volume horizonta advection, and use of an icosahedral horizontal grid
- FIM was designed to run efficiently on parallel computer systems



# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of
  Global Systems Division (GSD) of

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of
  Global Systems Division (GSD) of
  Earth System Research Lab (ESRL) of

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of
  Global Systems Division (GSD) of
  Earth System Research Lab (ESRL) of
  National Oceanic and
  Atmospheric Administration (NOAA)

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of
  Global Systems Division (GSD) of
  Earth System Research Lab (ESRL) of
  National Oceanic and
  Atmospheric Administration (NOAA)
  part of Department of Commerce (DOC)

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of
  Global Systems Division (GSD) of
  Earth System Research Lab (ESRL) of
  National Oceanic and
  Atmospheric Administration (NOAA)
  part of Department of Commerce (DOC)
  of the United States of America (USA)

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of
  Global Systems Division (GSD) of
  Earth System Research Lab (ESRL) of
  National Oceanic and
  Atmospheric Administration (NOAA)
  part of Department of Commerce (DOC)
  of the United States of America (USA)
  of Planet Earth (PE)

# About Our Little Section of the U.S. Gov't

- We are a group of 4 people in the

  Advanced Computing Section (ACS) of
  Aviation Division (AD) of
  Global Systems Division (GSD) of
  Earth System Research Lab (ESRL) of
  National Oceanic and
  Atmospheric Administration (NOAA)
  part of Department of Commerce (DOC)
  of the United States of America (USA)
  of Planet Earth (PE)

- Or, turning it into the world's longest domain name:
  jeff@ACS.AD.GSD.ESRL.NOAA.DOC.USA.PE

# Group Activities -1

- **Code Parallelization** - the Scalable Modeling System (SMS) is a tool for directive-based code parallelization for Fortran codes. The user inserts directives which are converted into parallel code which relies both on SMS libraries and the MPI libraries

# Group Activities -1

- **Code Parallelization** - the Scalable Modeling System (SMS) is a tool for directive-based code parallelization for Fortran codes. The user inserts directives which are converted into parallel code which relies both on SMS libraries and the MPI libraries
- **GPU** (graphical processing units) Research & Develop – because of CPU limitations (power & performance), we're investigating running as much of the code on the GPU as possible, while communications will be done on the CPU. We've written a Fortran to CUDA-C code translator and initial FIM tests yielded a 17 times performance improvement

# Group Activities -1

- **Code Parallelization** - the Scalable Modeling System (SMS) is a tool for directive-based code parallelization for Fortran codes. The user inserts directives which are converted into parallel code which relies both on SMS libraries and the MPI libraries
- **GPU** (graphical processing units) Research & Develop – because of CPU limitations (power & performance), we're investigating running as much of the code on the GPU as possible, while communications will be done on the CPU. We've written a Fortran to CUDA-C code translator and initial FIM tests yielded a 17 times performance improvement
- **Grid Computing** - explored the feasibility of using for ensembles, but security issues, and lack of maturity has hindered its use at NOAA

# Group Activities -2

- **FIM Development** – FIM is being run at ESRL and producing 10 day forecasts at 30km resolution.
  - 30 km FIM runs producing 10 day forecasts were run at the Texas Advanced Computing Center (TACC) and provided to the NOAA National Hurricane Center for information on hurricane tracks. Also testing a 20 member ensemble of 60 km FIM runs (240 cores/ensemble), with plans to increase the resolution to 15 km (1680 cores)

# Group Activities -2

- **FIM Development** – FIM is being run at ESRL and producing 10 day forecasts at 30km resolution.
  - 30 km FIM runs producing 10 day forecasts were run at the Texas Advanced Computing Center (TACC) and provided to the NOAA National Hurricane Center for information on hurricane tracks. Also testing a 20 member ensemble of 60 km FIM runs (240 cores/ensemble), with plans to increase the resolution to 15 km (1680 cores)
- **Data Locator -** an html based search engine for finding and viewing meteorological data sets (also includes web services callable from other applications)

# Group Activities -2

- **FIM Development** – FIM is being run at ESRL and producing 10 day forecasts at 30km resolution.
  - 30 km FIM runs producing 10 day forecasts were run at the Texas Advanced Computing Center (TACC) and provided to the NOAA National Hurricane Center for information on hurricane tracks. Also testing a 20 member ensemble of 60 km FIM runs (240 cores/ensemble), with plans to increase the resolution to 15 km (1680 cores)
- **Data Locator** - an html based search engine for finding and viewing meteorological data sets (also includes web services callable from other applications)
- **NOAA Modeling Portal** - the subject of this presentation

# What is NOAA Modeling Portal? -1

- A graphical user interface for running weather models such as WRF and FIM

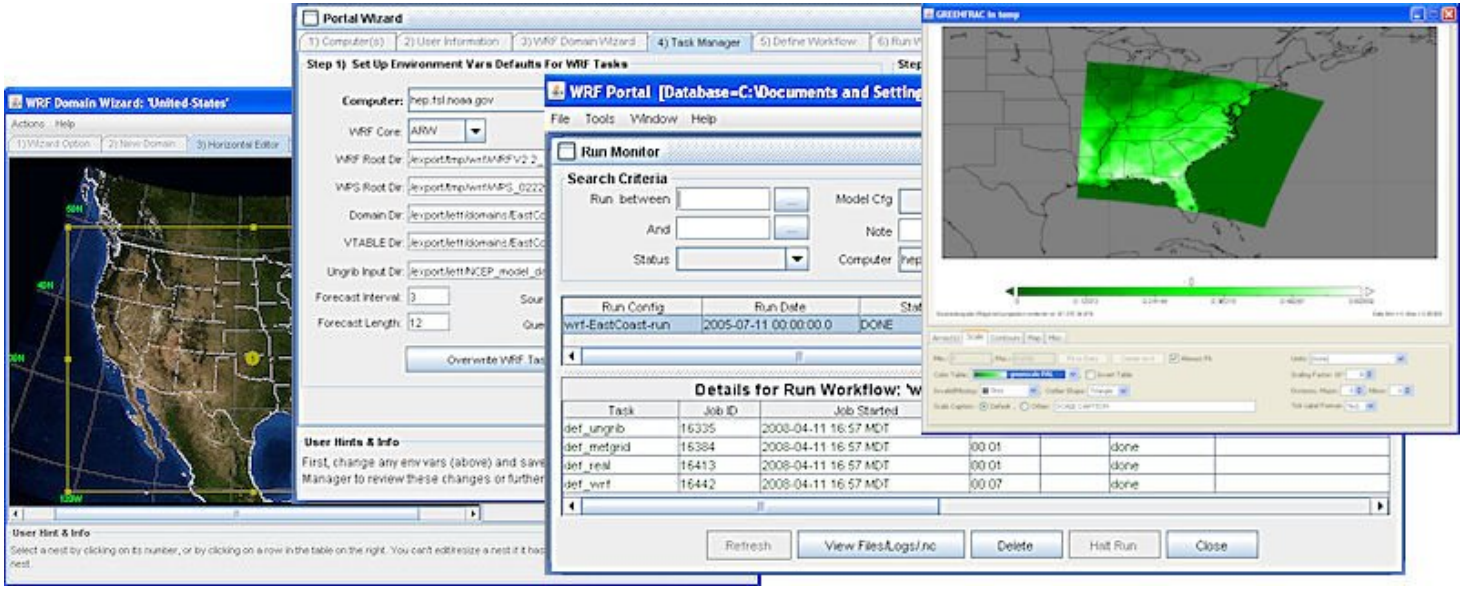
# What is NOAA Modeling Portal? -1

- A graphical user interface for running weather models such as WRF and FIM
- This application runs on all platforms and can be launched from a standard web browser as a Java Web Start program

# What is NOAA Modeling Portal? -1

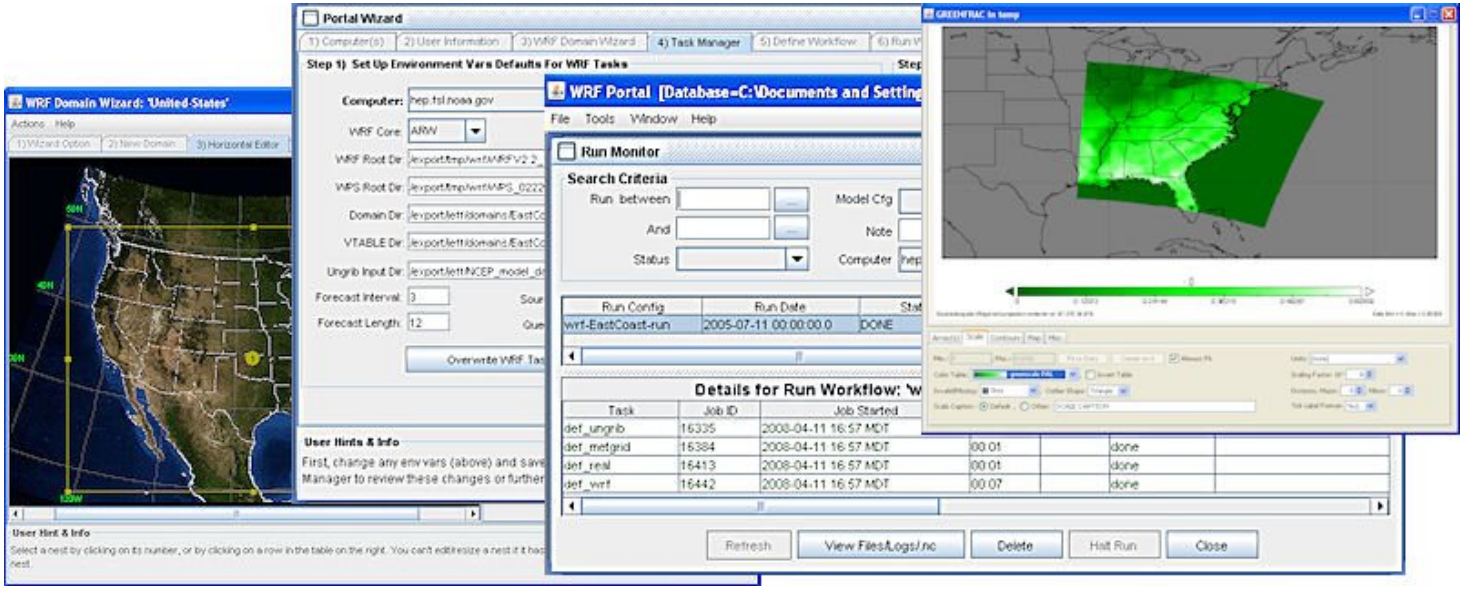
- A graphical user interface for running weather models such as WRF and FIM
- This application runs on all platforms and can be launched from a standard web browser as a Java Web Start program
- It simplifies and automates:
  - configuring and running of model workflows
  - selection/localization of your domain (WRF)
  - launching and monitoring runs
  - Halting or canceling runs/jobs
  - visualization of your model's output

# What is NOAA Modeling Portal? -2



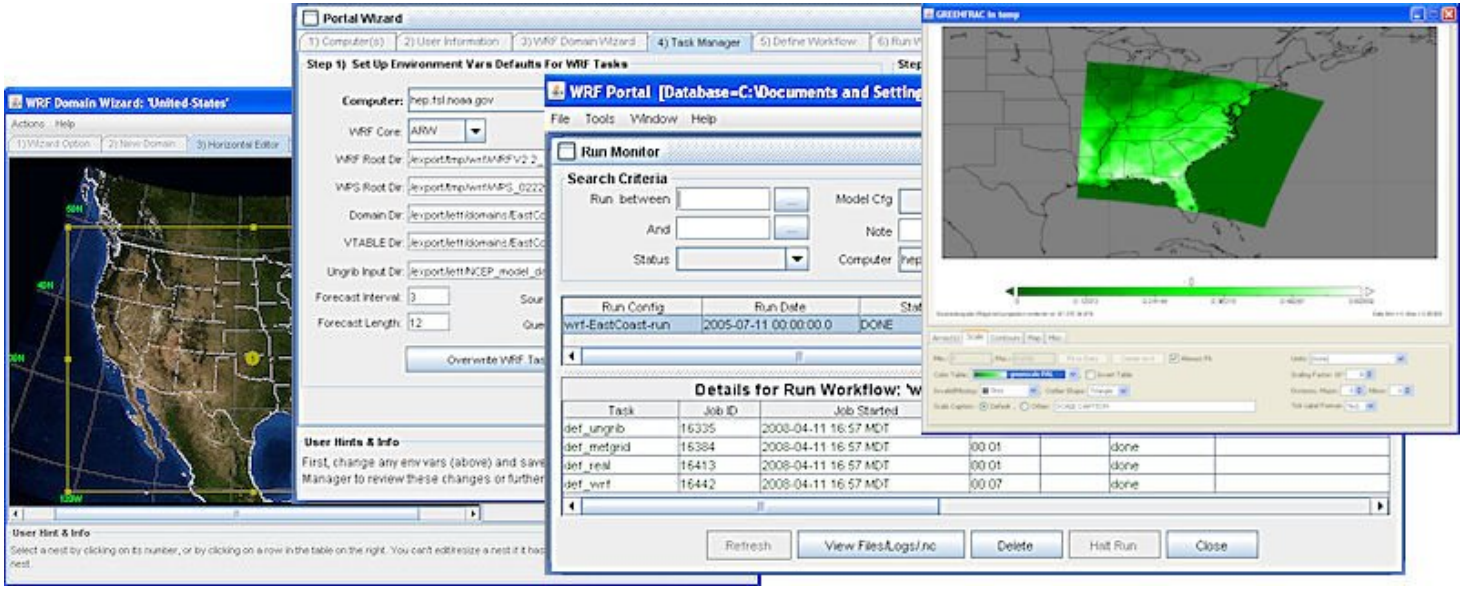

- It does not include the models (WRF and FIM) themselves—this software must be installed separately.
  - Portal is a GUI front end for running WRF or FIM models

# What is NOAA Modeling Portal? -2




- It does not include the models (WRF and FIM) themselves—this software must be installed separately.
  - Portal is a GUI front end for running WRF or FIM models
- Supports batch queue systems such as LSF and SGE and includes job managers

# What is NOAA Modeling Portal? -2




- It does not include the models (WRF and FIM) themselves—this software must be installed separately.
  - Portal is a GUI front end for running WRF or FIM models
- Supports batch queue systems such as LSF and SGE and includes job managers
- Version 1.00 was released on Oct 10, 2008
- Requires Java 1.5 or later, runs on most platforms

# Why Use NOAA Portal?

- Saves user's time by automating tedious and repetitive tasks and providing time saving features
- Portal Wizard that walks the user through the steps of configuring computers, user preferences, and tasks

# Why Use NOAA Portal?

- Saves user's time by automating tedious and repetitive tasks and providing time saving features
- Portal Wizard that walks the user through the steps of configuring computers, user preferences, and tasks
- "Diff" tool for comparing different workflows and runs
- Graphical file browsers to quickly locate files
- Robust job managers for running and managing tasks
- Progress monitor for tracking the progress of runs

# Why Use NOAA Portal?

- Saves user's time by automating tedious and repetitive tasks and providing time saving features
- Portal Wizard that walks the user through the steps of configuring computers, user preferences, and tasks
- "Diff" tool for comparing different workflows and runs
- Graphical file browsers to quickly locate files
- Robust job managers for running and managing tasks
- Progress monitor for tracking the progress of runs
- Graphical netcdf viewers to visualize model input/output
- Stores its information in a database so you can easily search and retrieve your information without the tedium of hunting through a myriad of files in directories directories

# Two Main Categories of NOAA Portal Users

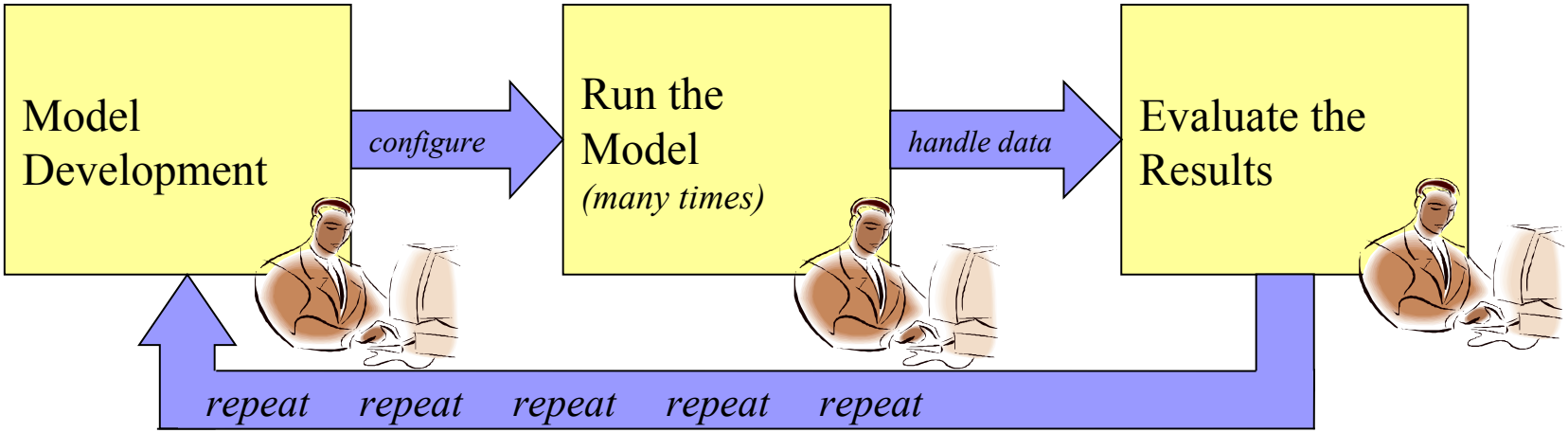
- **"Black box" users** who know very little about the details of running a model and want an quick and easy way to run their model without having to master long and complex instruction manuals or tutorials. They may run their models on a Linux desktop, a local cluster, or perhaps a super-computer

# Two Main Categories of NOAA Portal Users

- **"Black box" users** who know very little about the details of running a model and want an quick and easy way to run their model without having to master long and complex instruction manuals or tutorials. They may run their models on a Linux desktop, a local cluster, or perhaps a super-computer
- **Model developers and testers** who, while familiar with the details of configuring namelists and creating run scripts, want a tool to simplify the process. Managing and making dozens or hundreds of model runs can be tedious, time-consuming, and prone to error. The portal automates many tedious tasks, freeing the developer or tester to focus on the science of their model runs

# Modelers and Testers

Model development is an iterative process



# Configuring/Running Models is Complicated!

- Typical workflow entails editing multiple configuration files (namelist text files), setting environment variables, running executable files in the correct order, checking the output to make sure the workflow can proceed to the next step

# Configuring/Running Models is Complicated!

- Typical workflow entails editing multiple configuration files (namelist text files), setting environment variables, running executable files in the correct order, checking the output to make sure the workflow can proceed to the next step
- Files are scattered across many different directories
- Editing the configuration files by hand in a text editor such as **vi** is tedious and mistake prone

# Configuring/Running Models is Complicated!

- Typical workflow entails editing multiple configuration files (namelist text files), setting environment variables, running executable files in the correct order, checking the output to make sure the workflow can proceed to the next step
- Files are scattered across many different directories
- Editing the configuration files by hand in a text editor such as **vi** is tedious and mistake prone
- Selecting a domain (geographic region) in WRF, for example, without visual feedback can be a frustrating exercise in trial and error

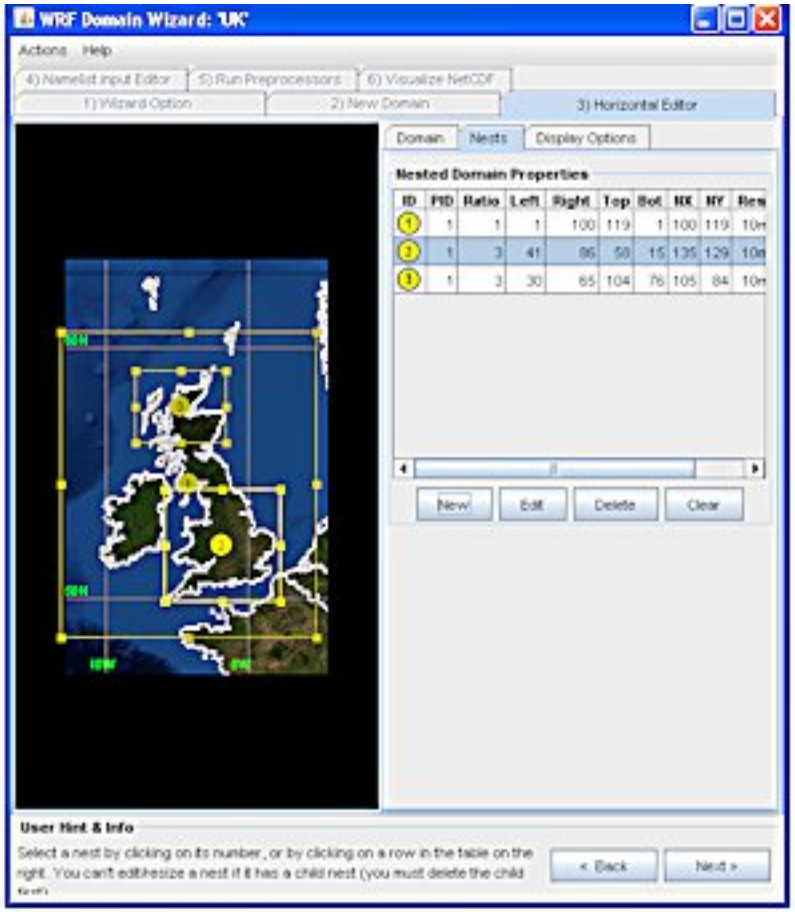
# Example: WRF Workflow comparisons

| WRF from command line Hard! | WRF using portal Easy! |
|---|---|
| Lats/Lons are looked up in an atlas, grid points are computed with a calculator, user manually types all this data into namelist text files | User draws a box around a region of the Earth on the screen, grid points are automatically calculated with minimal user input, namelists are automatically written |
| Namelists and scripts are scattered in multiple directories, manually edited in vi | Namelists and scripts are stored in database and edited in WRF Portal visual editors |
| New runs involve copying lots of files, editing them with vi, set environment vars, launching scripts manually. | New runs are launched from the Run Workflow window. Most of the info is already filled out for you. |
| Runs are monitored by tailing certain log files in specific directories. Requires detailed knowledge of WRF. | Runs are monitored from a window which lists all tasks, how long they've been running and the estimated time to finish |
| Comparing two different model runs involves viewing multiple namelists files for each model and hunting for the differences. | Comparing two different model runs is done from a "diff" window that highlights all differences between the models. |
| Finding an old run involves hunting through directories on different computers/file systems | Finding an old run is done by typing in search criteria into a window |
| Runs are canceled by killing jobs from the command line | Runs are canceled with a mouse click |
| Visualization is done with external tools | Visualization is done within WRF Portal with a mouse click |

# Graphical Editors -1

From namelist.wps configuration file (domain and nests)

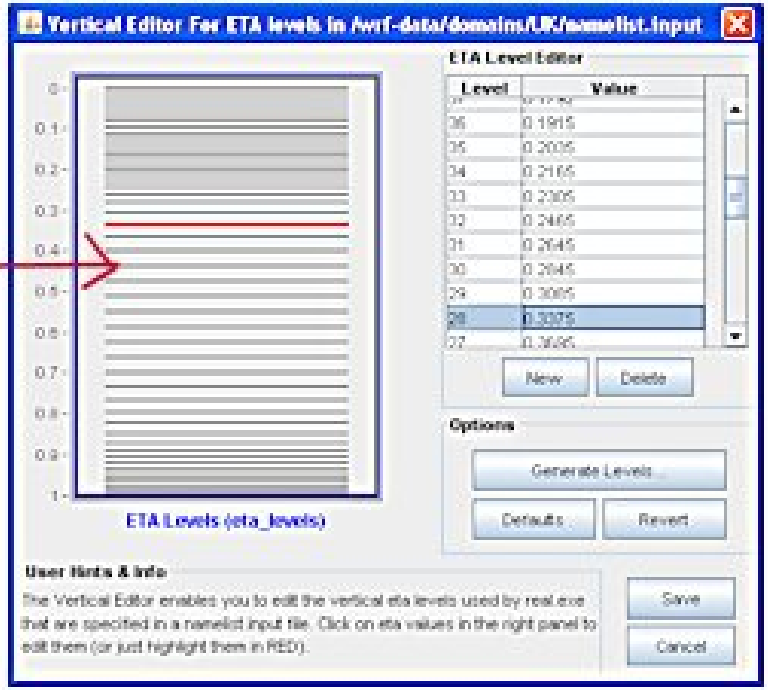
```
 interval_seconds = 10800,
 io_form_geogrid = 2,
 opt_output_from_geogrid_path = '/wrf-data/domains/UK/',
 debug_level = 0,
/

&geogrid
 parent_id         = 1,1,1,
 parent_grid_ratio = 1,3,3,
 i_parent_start    = 1,41,30,
 j_parent_start    = 1,15,76,
 e_we          = 100,136,106,
 e_sn          = 119,130,85,
 geog_data_res = '10m','10m','10m',
 dx = 11400,
 dy = 11400,
 map_proj =  'mercator',
 ref_lat   = 54.804,
 ref_lon   = -4.195,
 truelat1  = 54.804,
 truelat2  = 0,
 stand_lon = -4.195,
 geog_data_path = '/wrf-data/geog10m',
 opt_geogrid_tbl_path = '/wrf-data/domains/UK/',
 ref_x = 50.0,
 ref_y = 59.5,
/
```



# Graphical Editors -2

From namelist.input configuration file (ETA levels)

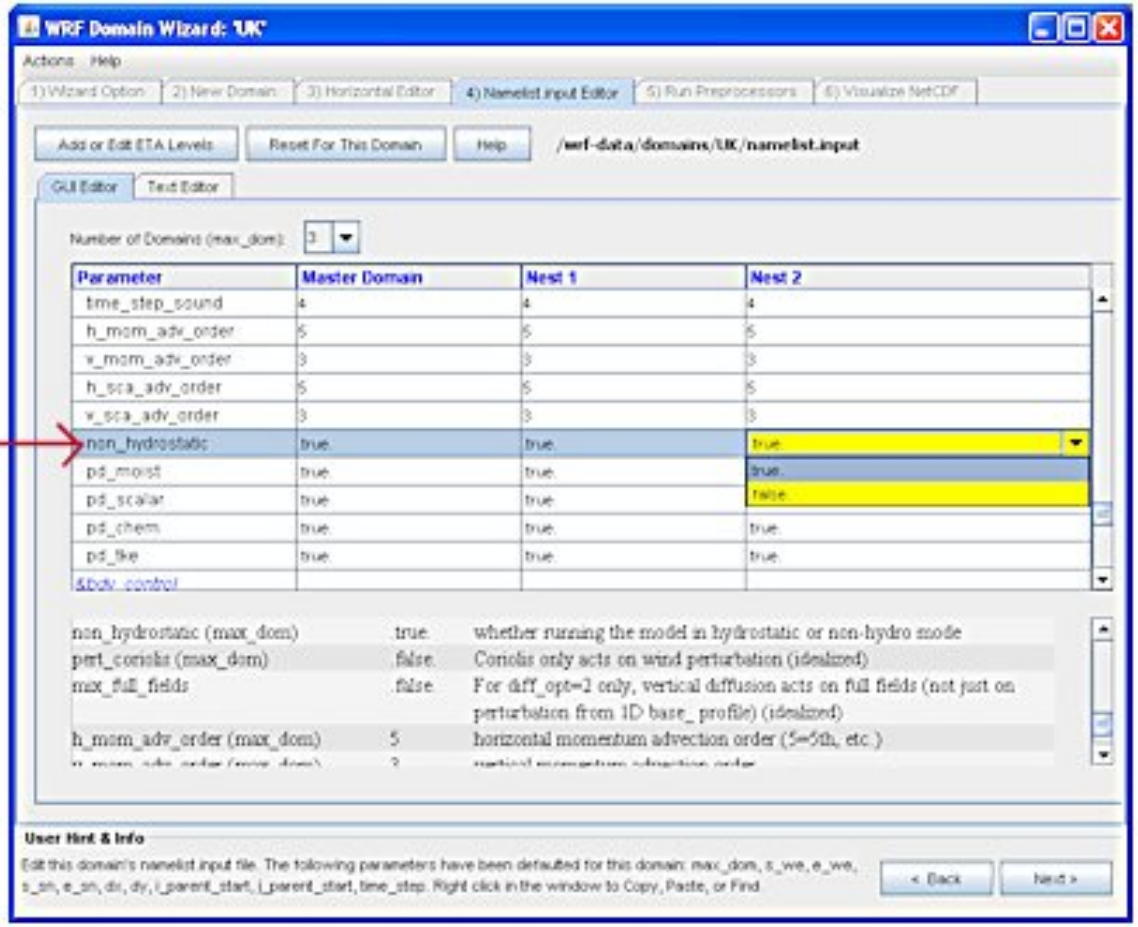
```
&domains
eta_levels   = 1.000, 0.994, 0.987, 0.979, 0.97,
               0.96, 0.949, 0.937, 0.924, 0.909,
               0.892, 0.873, 0.851, 0.826, 0.798,
               0.768, 0.736, 0.702, 0.666, 0.629,
               0.5915, 0.5536, 0.5153, 0.4773, 0.44,
               0.404, 0.3695, 0.3375, 0.3085, 0.2845,
               0.2645, 0.2465, 0.2305, 0.2165, 0.2035,
               0.1915, 0.1792, 0.1667, 0.1539, 0.1407,
               0.1272, 0.1134, 0.0995, 0.0855, 0.0713,
               0.0571, 0.0429, 0.0287, 0.0145, 0.000,
time_step              = 68,
time_step_fract_num     = 0,
time_step_fract_den     = 1,
max_dom                 = 3,
s_we                   = 1,      1,      1,
e_we                   = 100,    136,    106,
```



# Graphical Editors -3

From namelist.input configuration file (all parameters)

```
&dynamics
w_damping              = 0,
diff_opt               = 1,
km_opt                 = 4,
base_temp              = 290.,
damp_opt               = 0,
zdamp                  = 5000.,   5000.,   5000.,
dampcoef               = 0.01,    0.01,    0.01,
khdif                  = 0,       0,       0,
kvdif                  = 0,       0,       0,
smdiv                  = 0.1,     0.1,     0.1,
emdiv                  = 0.01,    0.01,    0.01,
epssm                  = 0.1,     0.1,     0.1,
time_step_sound        = 4,       4,       4,
h_mom_adv_order        = 5,       5,       5,
v_mom_adv_order        = 3,       3,       3,
h_sca_adv_order        = 5,       5,       5,
v_sca_adv_order        = 3,       3,       3,
non_hydrostatic        = .true.,  .true.,  .true.,
pd_moist               = .true.,  .true.,  .true.,
pd_scalar              = .true.,  .true.,  .true.,
pd_chem                = .true.,  .true.,  .true.,
pd_tke                 = .true.,  .true.,  .true.,
/
```



# NOAA Portal Design -1

- Written in Java so it runs on virtually all platforms (e.g. Linux, Windows, Unix, Mac)

# NOAA Portal Design -1

- Written in Java so it runs on virtually all platforms (e.g. Linux, Windows, Unix, Mac)
- Stores most of the user's work and information in a SQL database

# NOAA Portal Design -1

- Written in Java so it runs on virtually all platforms (e.g. Linux, Windows, Unix, Mac)
- Stores most of the user's work and information in a SQL database
- Standard version of the portal (the single-user or "desktop version") is a Java Web Start application that employs an embedded HSQL database.
  - All a user has to do is click the web link and the software downloads (if the user hasn't downloaded it before) and runs. No other installation steps are required.

# NOAA Portal Design -2

- Portal also supports a MySQL database, enabling all users at a site to share the same database instance
  - Developmental Testbed Center (DTC) is using this setup

# NOAA Portal Design -2

- Portal also supports a MySQL database, enabling all users at a site to share the same database instance
  - Developmental Testbed Center (DTC) is using this setup
  - Administrator can create users and configure the computing resources available at the site
  - Portal users are prompted for a username and password

# NOAA Portal Design -2

- Portal also supports a MySQL database, enabling all users at a site to share the same database instance
  - Developmental Testbed Center (DTC) is using this setup
  - Administrator can create users and configure the computing resources available at the site
  - Portal users are prompted for a username and password
  - Workflows are associated with a username
  - Users can, however, open up read-only instances of other user's workflows and their runs

# NOAA Portal Design -3

- Advantages of storing this data in a SQL database instead of in multiple disk files
  - Easy to back up the data (archive it) and query it based on multiple criteria

# NOAA Portal Design -3

- Advantages of storing this data in a SQL database instead of in multiple disk files
  - Easy to back up the data (archive it) and query it based on multiple criteria
  - Eliminated the clutter and complication of the user maintaining many files in multiple directories on potentially multiple file systems

# NOAA Portal Design -3

- Advantages of storing this data in a SQL database instead of in multiple disk files
  - Easy to back up the data (archive it) and query it based on multiple criteria
  - Eliminated the clutter and complication of the user maintaining many files in multiple directories on potentially multiple file systems
  - Enabled the user to run NOAA Portal and access most of his data from any computer, mounted to any file system

# NOAA Portal Design -4

- After creating a workflow in NOAA Portal, the user can choose to run it on the desired computer (which was previously configured in the portal software)

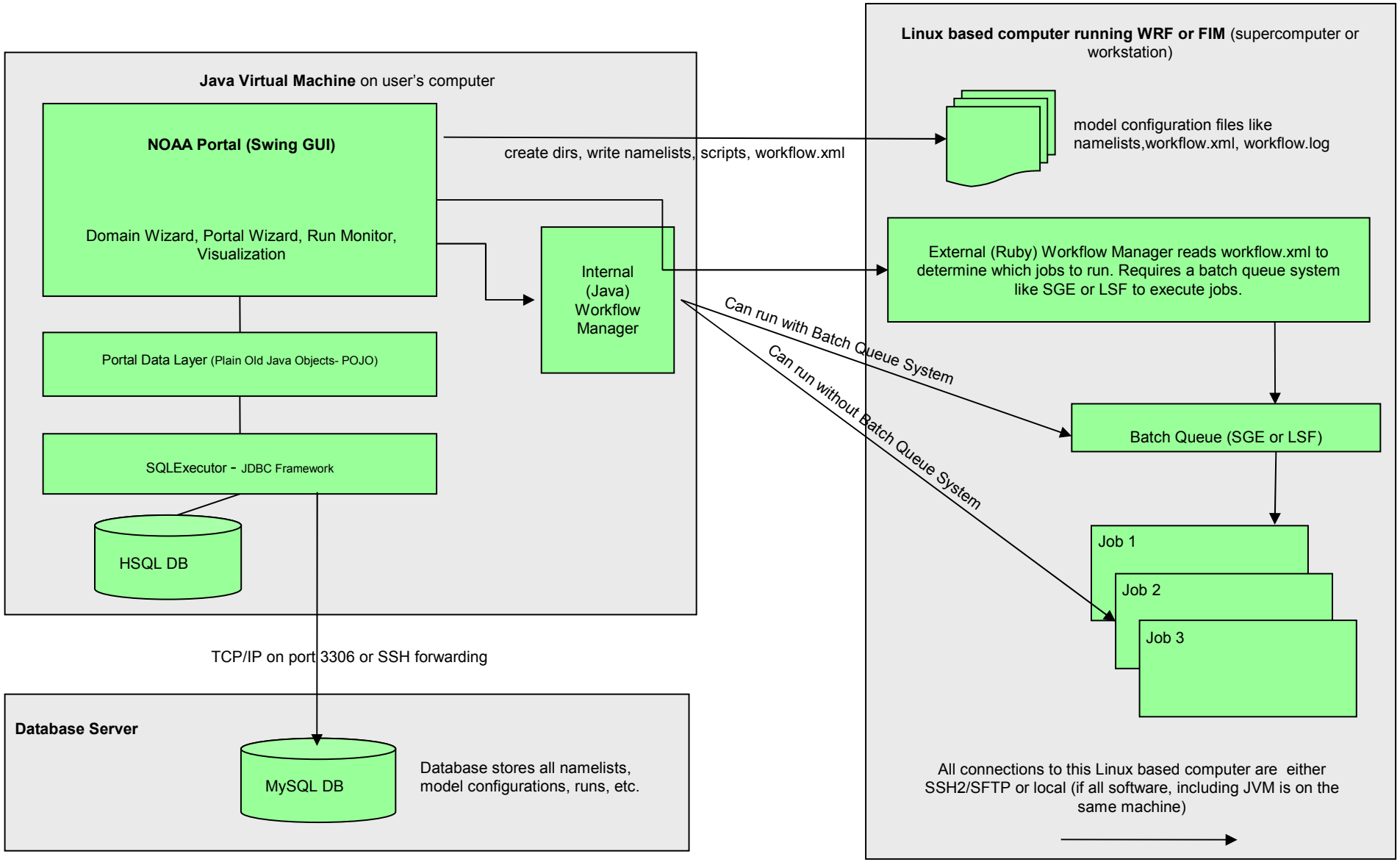
# NOAA Portal Design -4

- After creating a workflow in NOAA Portal, the user can choose to run it on the desired computer (which was previously configured in the portal software)
- Portal can be run
  - on a local computer (direct access to files and exes)
  - on a remote computer using SSH2/SFTP

# NOAA Portal Design -4

- After creating a workflow in NOAA Portal, the user can choose to run it on the desired computer (which was previously configured in the portal software)
- Portal can be run
  - on a local computer (direct access to files and exes)
  - on a remote computer using SSH2/SFTP
- User can also run portal “locally” on a remote computer if they first SSH to the remote computer and then launch the portal in an X windows session.

# NOAA Portal Design -5



# NOAA Portal Security

- User logs in with credentials that grants him the limited database access required to run the software

# NOAA Portal Security

- User logs in with credentials that grants him the limited database access required to run the software
- Database does not store any sensitive information such as usernames or passwords to other computers.

# NOAA Portal Security

- User logs in with credentials that grants him the limited database access required to run the software
- Database does not store any sensitive information such as usernames or passwords to other computers.
- If using an external MySQL database, all database traffic is transmitted via TCP/IP to port 3306 (or another port designated by the administrator), and the password is encrypted during all transfers
  - HSQL database version runs in same VM as portal, so password isn't transferred over the network

# NOAA Portal Security

- User logs in with credentials that grants him the limited database access required to run the software
- Database does not store any sensitive information such as usernames or passwords to other computers.
- If using an external MySQL database, all database traffic is transmitted via TCP/IP to port 3306 (or another port designated by the administrator), and the password is encrypted during all transfers
  - HSQL database version runs in same VM as portal, so password isn't transferred over the network
- We are also considering the possibility of further strengthening database security by using SSH Port forwarding to encrypt all incoming and outgoing data

# NOAA Portal Security

- User logs in with credentials that grants him the limited database access required to run the software
- Database does not store any sensitive information such as usernames or passwords to other computers.
- If using an external MySQL database, all database traffic is transmitted via TCP/IP to port 3306 (or another port designated by the administrator), and the password is encrypted during all transfers
  - HSQL database version runs in same VM as portal, so password isn't transferred over the network
- We are also considering the possibility of further strengthening database security by using SSH Port forwarding to encrypt all incoming and outgoing data
- Portal uses SSH2/SFTP to connect to remote computers

# NOAA Portal Job Management

- Two ways to execute jobs with the portal.
  - **External (Ruby) workflow manager** developed by Chris Harrop. This workflow manager must be installed on the same computer as WRF or FIM, and it runs independently. It provides robust job management including job restart on failure, job monitoring, and supports running multiple tasks simultaneously.
    - Requires that either SGE or LSF be present

# NOAA Portal Job Management

- Two ways to execute jobs with the portal.
  - **External (Ruby) workflow manager** developed by Chris Harrop. This workflow manager must be installed on the same computer as WRF or FIM, and it runs independently. It provides robust job management including job restart on failure, job monitoring, and supports running multiple tasks simultaneously.
    - Requires that either SGE or LSF be present
  - **Internal (Java) workflow manager**. This workflow manager is built into the portal software, and thus runs on your desktop computer and can only monitor jobs while the portal itself is running.
    - SGE and LSF are optional

# Modelers Working in Groups

- Version 1.01 of the portal has limited support for working in groups when using the site-wide MySQL database option.

# Modelers Working in Groups

- Version 1.01 of the portal has limited support for working in groups when using the site-wide MySQL database option.
- Multiple users can log into the portal at the same time and can access and edit the same workflows and runs, as long as they share a group login.
  - DTC used this setup when using the portal to do dozens of retrospective runs for the 2007 13-km Core Test (NMM-WRF versus ARW-WRF).

# Modelers Working in Groups

- Version 1.01 of the portal has limited support for working in groups when using the site-wide MySQL database option.
- Multiple users can log into the portal at the same time and can access and edit the same workflows and runs, as long as they share a group login.
  - DTC used this setup when using the portal to do dozens of retrospective runs for the 2007 13-km Core Test (NMM-WRF versus ARW-WRF).
- We are planning to add a more flexible group management approach in a future release of the portal software

# Testing

- Testing is such a laborious process for modelers that sometimes it receives short shrift
- Setting up and re-running dozens of regression test runs for seemingly minor code changes can be so time consuming that users often do not do it.
- Since the portal makes it so easy to repeat workflow runs (or subsets of workflows), it encourages modelers to perform rigorous testing

# DTC 2007 WRF Core Tests -1

- In 2007 (and into 2008), the DTC ran side by side comparison of both WRF cores: NMM and ARW
  - Tests were run on two supercomputers: 'bluevista' at NCAR (National Center for Atmospheric Research) and 'wjet' at Earth System Research Lab
  - Purpose was to compare the efficacy of each core at predicting various meteorological phenomena

# DTC 2007 WRF Core Tests -1

- In 2007 (and into 2008), the DTC ran side by side comparison of both WRF cores: NMM and ARW
  - Tests were run on two supercomputers: 'bluevista' at NCAR (National Center for Atmospheric Research) and 'wjet' at Earth System Research Lab
  - Purpose was to compare the efficacy of each core at predicting various meteorological phenomena
- For each core on 'bluevista' (an AIX system), thirty runs were performed for each season (winter, spring, summer, and fall), and for platform comparisons, an additional ten runs per season were conducted on 'wjet' (a Linux system)

# DTC 2007 WRF Core Tests -1

- In 2007 (and into 2008), the DTC ran side by side comparison of both WRF cores: NMM and ARW
  - Tests were run on two supercomputers: 'bluevista' at NCAR (National Center for Atmospheric Research) and 'wjet' at Earth System Research Lab
  - Purpose was to compare the efficacy of each core at predicting various meteorological phenomena
- For each core on 'bluevista' (an AIX system), thirty runs were performed for each season (winter, spring, summer, and fall), and for platform comparisons, an additional ten runs per season were conducted on 'wjet' (a Linux system)
- Each run consisted of executing dozens of model tasks, including moving model data from a mass store to a network file system, using the portal's external workflow manager

# DTC 2007 WRF Core Tests -2

- Portal was used for reruns of the WRF NMM core (120 runs in all) using a MySQL (portal) database.

# DTC 2007 WRF Core Tests -2

- Portal was used for reruns of the WRF NMM core (120 runs in all) using a MySQL (portal) database.
- Since the portal doesn't support (yet) group accounts, we ran into file protection problems
  - the file system, by default, wouldn't allow other users to rerun another person's tests
  - As a workaround solution to the problem, we had to manually change file permissions so other users could overwrite files

# DTC 2007 WRF Core Tests -2

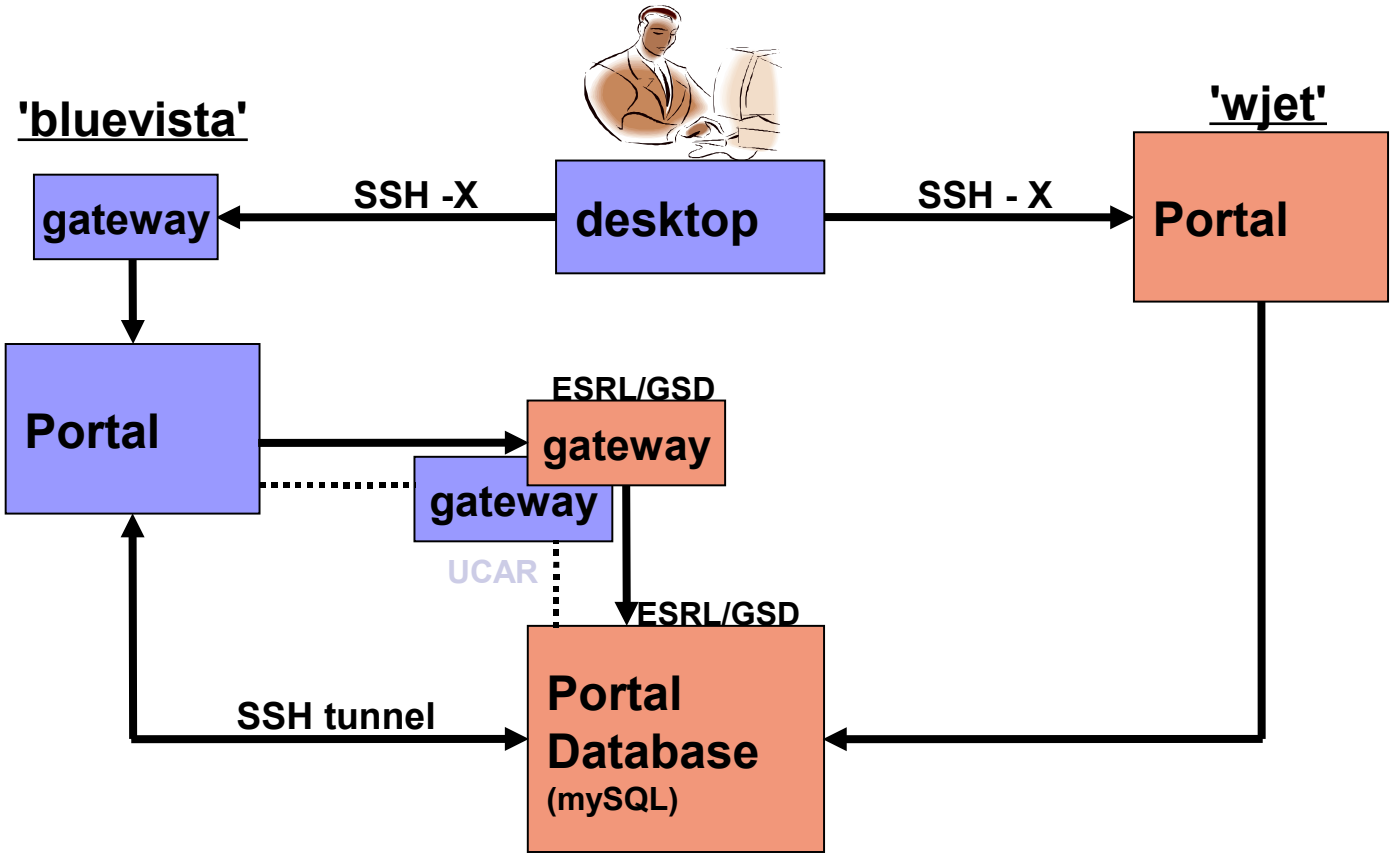
- Portal was used for reruns of the WRF NMM core (120 runs in all) using a MySQL (portal) database.
- Since the portal doesn't support (yet) group accounts, we ran into file protection problems
  - the file system, by default, wouldn't allow other users to rerun another person's tests
  - As a workaround solution to the problem, we had to manually change file permissions so other users could overwrite files
- Despite problems, the portal managed 10-15 runs every twelve hours and the portal's Run Monitor made it easy to check the job(s) status.

# DTC 2007 WRF Core Tests -2

- Portal was used for reruns of the WRF NMM core (120 runs in all) using a MySQL (portal) database.
- Since the portal doesn't support (yet) group accounts, we ran into file protection problems
  - the file system, by default, wouldn't allow other users to rerun another person's tests
  - As a workaround solution to the problem, we had to manually change file permissions so other users could overwrite files
- Despite problems, the portal managed 10-15 runs every twelve hours and the portal's Run Monitor made it easy to check the job(s) status.
- Each workflow took about 8 hours to run, 4 hours for WRF
  - All 120 runs launched and monitored by the portal were successfully completed within eight days

# DTC 2007 WRF Core Tests -3

- Because of the complex security of 'bluevista' and 'wjet', we copied the portal software to each system and then ran the portal locally on the supercomputers, just forwarding the display back to our desktops (using the SSH -X) option



# Managing Complex Workflows -1

- The portal simplifies the management of complex workflows with numerous tasks
- The next slide illustrates a FIM workflow with
  - 45 tasks (.csh and .ksh scripts)
  - 892 environment variables (under the Settings tabs) in all
    - about 20 of them were common to all tasks with around 3 unique environment variables per task
  - Note that scripts, environment variables, and settings for this large workflow can be managed from this single window

# Managing Complex Workflows -2

FIM Portal [Database=wrfportal] [User=govett]

File Tools Window Help

Workflow: retroFIM [Model=FIM] [wjet.rdhpcs.noaa.gov] [User=govett] [CanEdit=false]

Actions

Configuration Files | General Settings | Task Settings

Show run scripts

fimprep_50_8.csh
FIM_50_8.csh
pop_PRS.csh
pop_NAT.csh
interp_220_NAT.csh
interp_220_PRS.csh
interp_236_NAT.csh
interp_236_PRS.csh
interp_africa_NAT.csh
interp_africa_PRS.csh
interp_219_NAT.csh
interp_219_PRS.csh
interp_244_NAT.csh
interp_244_PRS.csh
interp_w_pacific_NAT.csh
interp_w_pacific_PRS.csh
grib12_fim_PRS.csh
grib12_fim_NAT.csh
grib12_220_NAT.csh
grib12_220_PRS.csh
grib12_236_NAT.csh
grib12_236_PRS.csh
grib12_africa_NAT.csh
grib12_africa_PRS.csh
grib12_219_NAT.csh

Tasks: Edit

fimprep_50_8
FIM_50_8
pop_PRS
pop_NAT
interp_220_NAT
interp_220_PRS
interp_236_NAT
interp_236_PRS
interp_africa_NAT
interp_africa_PRS
interp_219_NAT
interp_219_PRS
interp_244_NAT
interp_244_PRS

File: fimprep_50_8.csh Task: fimprep_50_8

```
#!/bin/csh
#
#  Make sure that the .e and .o file arrive in the working director
#$ -cwd
#
#  Set the account
#$ -A fim
#
#  My code is NOT re-runnable
#$ -r n
#
#  The max walltime
#$ -l h_rt=00:40:00
#$ -p 24

# Source setup file for the Lahey compiler if we are using it
if(${FC} == 'L') then
  source /opt/lahey/lf6480/csh_setup
endif

module switch intel intel/9.1-openmpi-1.2.6

# Make sure data directories exist
if(!(-e $DATADIR)) then
  echo $DATADIR  'does not exist'
  exit 1
endif
if(!(-e $DATADR2)) then
  echo $DATADR2  'does not exist'
  exit 1
endif

if(! $?ID) then
  set ID=""
endif

# $ID is only set by test scripts to avoid file system race conditi
# concurrent tests.  By default it is set to "" .
set fimnamelist="${FIM_HOME}/FIMwfm/FIMnamelist${ID}"
set thetacoor="${FIM_HOME}/FIMrun/theta_coor${NVL}.txt${ID}"
set dpsig="${FIM_HOME}/FIMrun/dpsig${NVL}.txt${ID}"
```

Import Multiple | Import | Export

Connected to local computer: Jeff-M2400

Scripts and namelists associated with tasks

Tasks in this workflow

# Future Work

- Improve the workflow managers to support complex dependencies and running multiple instances of the same task simultaneously.

# Future Work

- Improve the workflow managers to support complex dependencies and running multiple instances of the same task simultaneously.
- Extend NOAA Portal to support WRF-Chem and other models

# Future Work

- Improve the workflow managers to support complex dependencies and running multiple instances of the same task simultaneously.
- Extend NOAA Portal to support WRF-Chem and other models
- Improve the support for groups of modelers to work together on a project (so as to avoid the file permission problems we experienced during the DTC 2007 Core Tests).

# Future Work

- Improve the workflow managers to support complex dependencies and running multiple instances of the same task simultaneously.
- Extend NOAA Portal to support WRF-Chem and other models
- Improve the support for groups of modelers to work together on a project (so as to avoid the file permission problems we experienced during the DTC 2007 Core Tests).
- Supporting the visualization of additional file formats (e.g. GRIB visualization is currently under development)

# Future Work

- Improve the workflow managers to support complex dependencies and running multiple instances of the same task simultaneously.
- Extend NOAA Portal to support WRF-Chem and other models
- Improve the support for groups of modelers to work together on a project (so as to avoid the file permission problems we experienced during the DTC 2007 Core Tests).
- Supporting the visualization of additional file formats (e.g. GRIB visualization is currently under development)
- Supporting a data discovery feature (through ongoing work on the Data Locator project) so the portal can automatically search for and find data required to run a model

# http://wrfportal.org website

WRF Portal - Mozilla Firefox

File Edit View History Bookmarks Tools Help

WRF Portal

Home | Contact Us

## Download WRF Portal

- Home
- WRF Portal
- Domain Wizard
- Tutorials (HTML)
- Tutorials (Video)
- F.A.Q.
- About

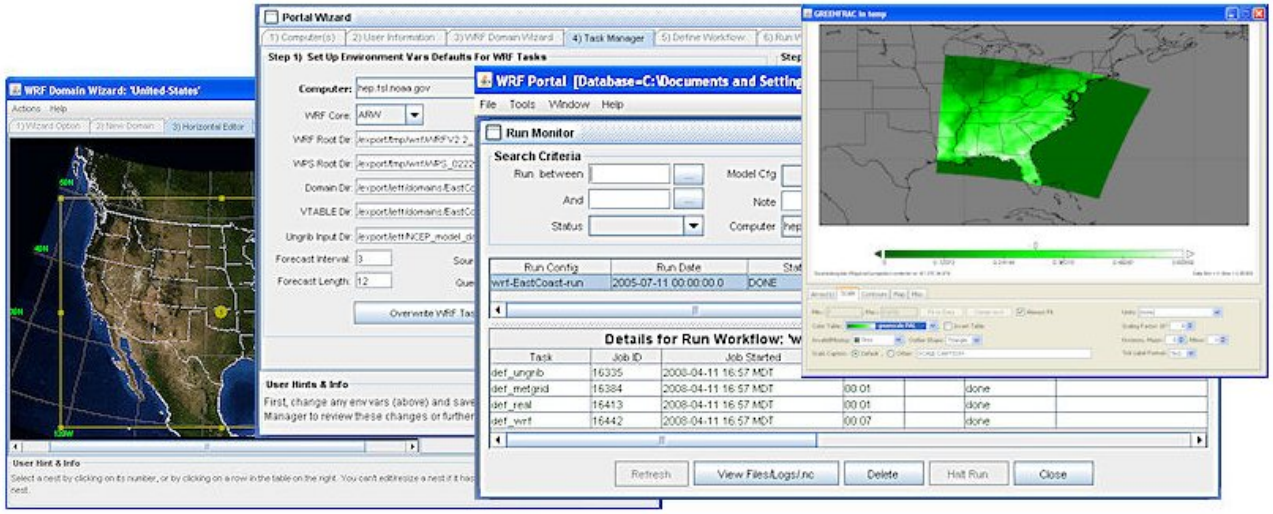
Dislaimer
Privacy Policy

NOAA website
ESRL website
FSL website

Accessibility statement

### WRF Portal: A GUI For Running WRF

Version 1.01 for Linux, AIX, Mac, and Windows - released October 15, 2008




WRF Portal is the GUI that takes you through the entire process of running WRF: creating a domain (using the built-in WRF Domain Wizard component that generates your namelist.wps and namelist.input files), creating and running workflows, monitoring the progress of your runs, diffing workflows and files. Supports WRF version 2.x and the new WRF version 3.

Source Code is available here.
What's new in version 1.01

**Run WRF Portal using Java Web Start (recommended, no installation required)**

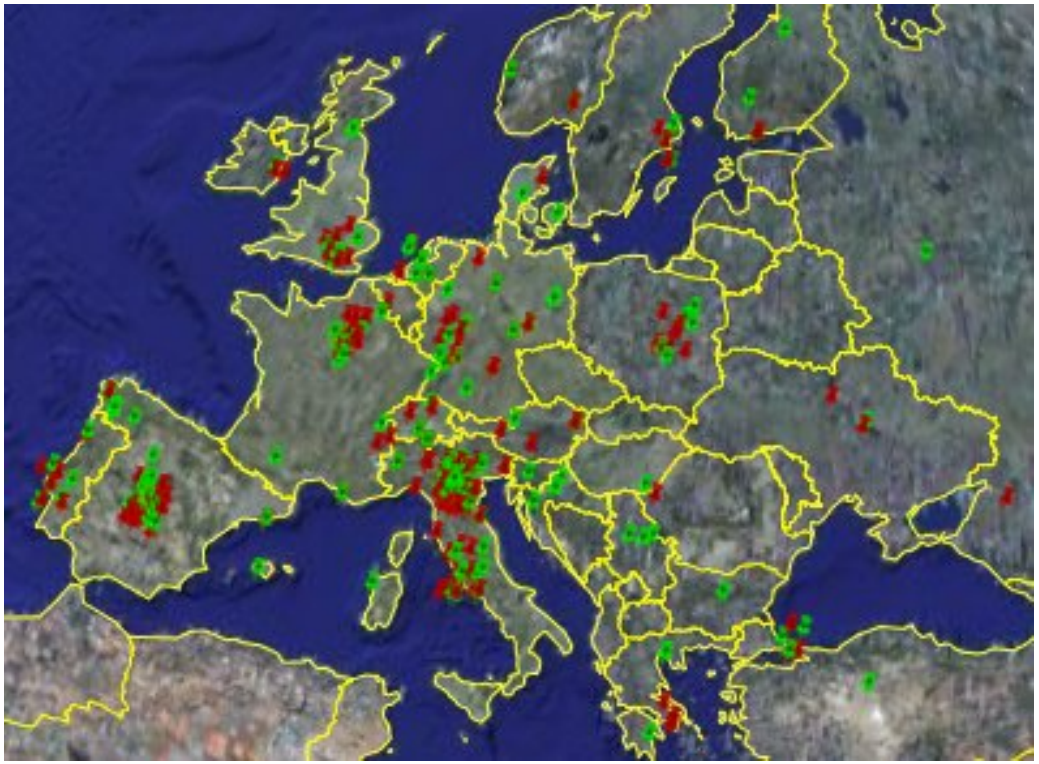
**Click here to launch WRF Portal 1.01**

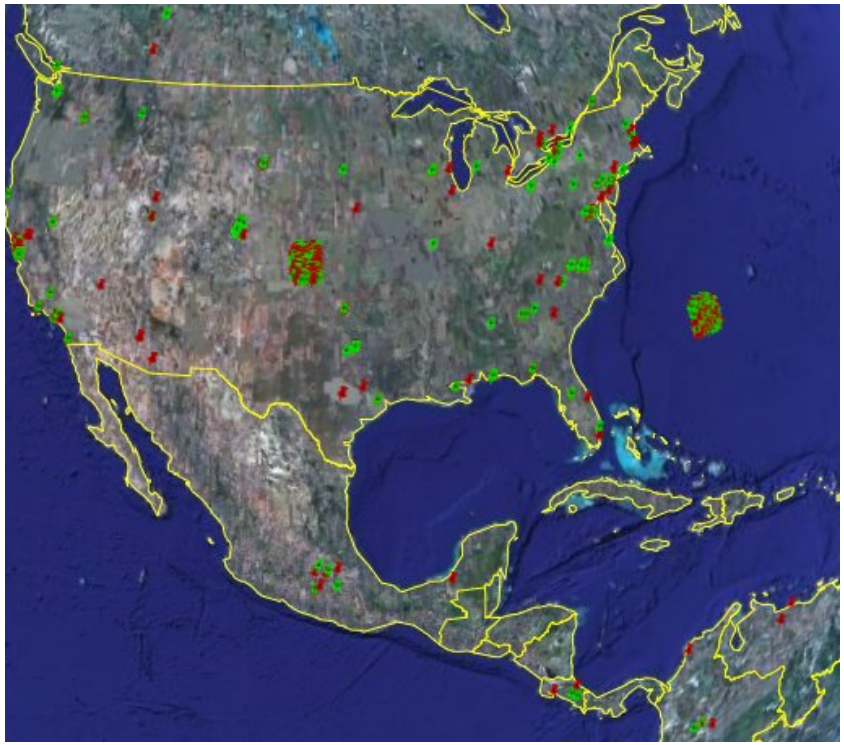
Done

# Worldwide Users Of This Software -1

| Software | Est. Users | Countries | Google Earth |
|---|---|---|---|
| WRF Portal | 623 | 58 | open in Google Earth |
| WRF Domain Wizard | 961 | 59 | open in Google Earth |
| Ext. Workflow Mgr | 49 | 14 | open in Google Earth |

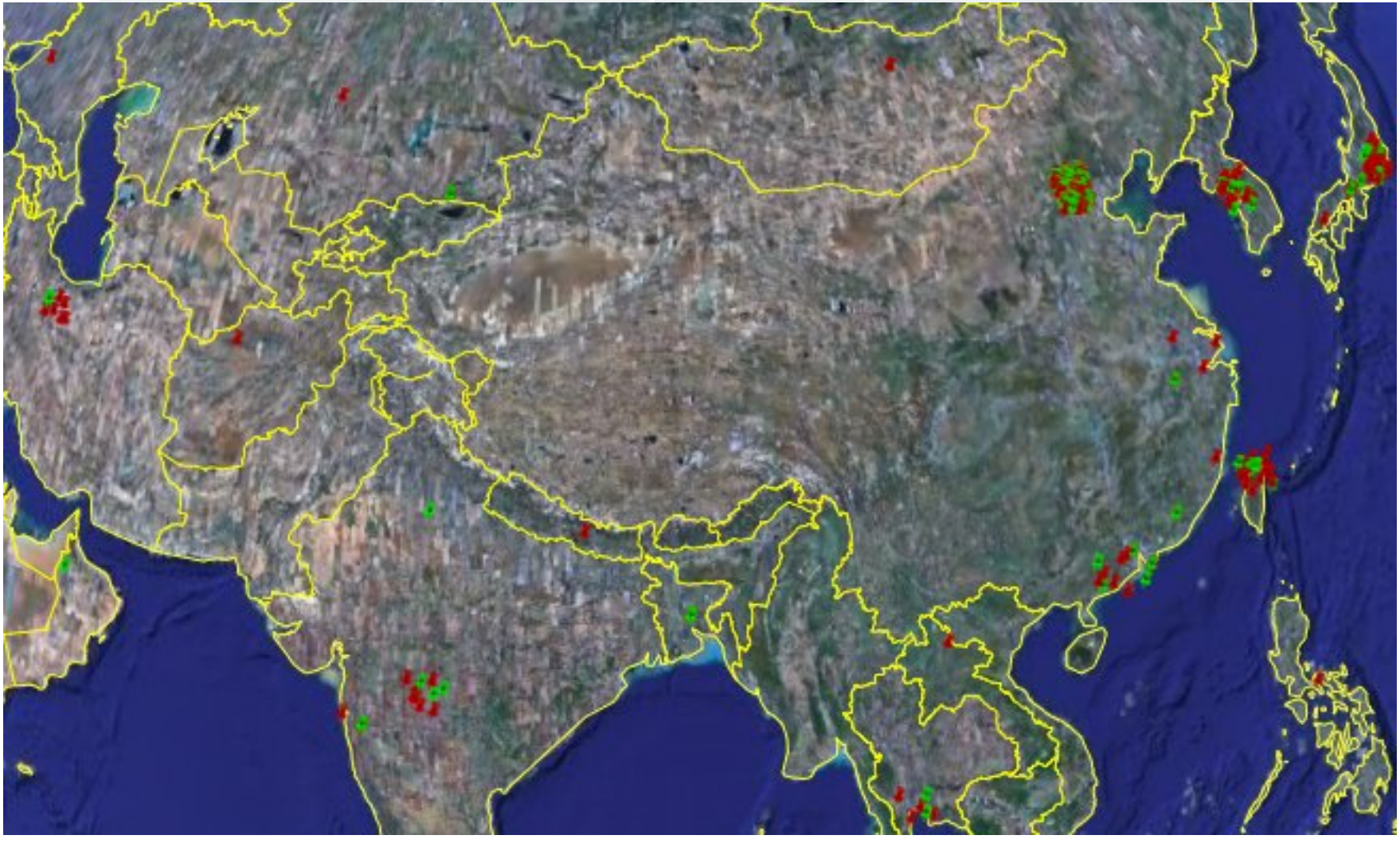
These users have downloaded or run the software from unique IP addresses during these time spans in 2008: May 4 - July 15 and Sep 14 - Oct 20.
http://wrfportal.org/about.html





# Worldwide Users Of This Software -2

| Software | Est. Users | Countries | Google Earth |
|---|---|---|---|
| WRF Portal | 623 | 58 | open in Google Earth |
| WRF Domain Wizard | 961 | 59 | open in Google Earth |



# Portal Support

- Web form for submitting bug reports or to ask questions: http://wrfportal.org/RequestInfoOrBugReport.html

# Portal Support

- Web form for submitting bug reports or to ask questions: http://wrfportal.org/RequestInfoOrBugReport.html
- Forums online here:
  http://forum.wrfforum.com/viewforum.php?f=48
  http://forum.wrfforum.com/viewforum.php?f=24

# Portal Support

- Web form for submitting bug reports or to ask questions: http://wrfportal.org/RequestInfoOrBugReport.html
- Forums online here:
  http://forum.wrfforum.com/viewforum.php?f=48
  http://forum.wrfforum.com/viewforum.php?f=24
- Tutorials on our website
  - 18 video (flash) tutorials here: http://wrfportal.org/flash-tutorial.html
  - 8 html tutorials here: http://wrfportal.org/html-tutorial.html

# In Conclusion

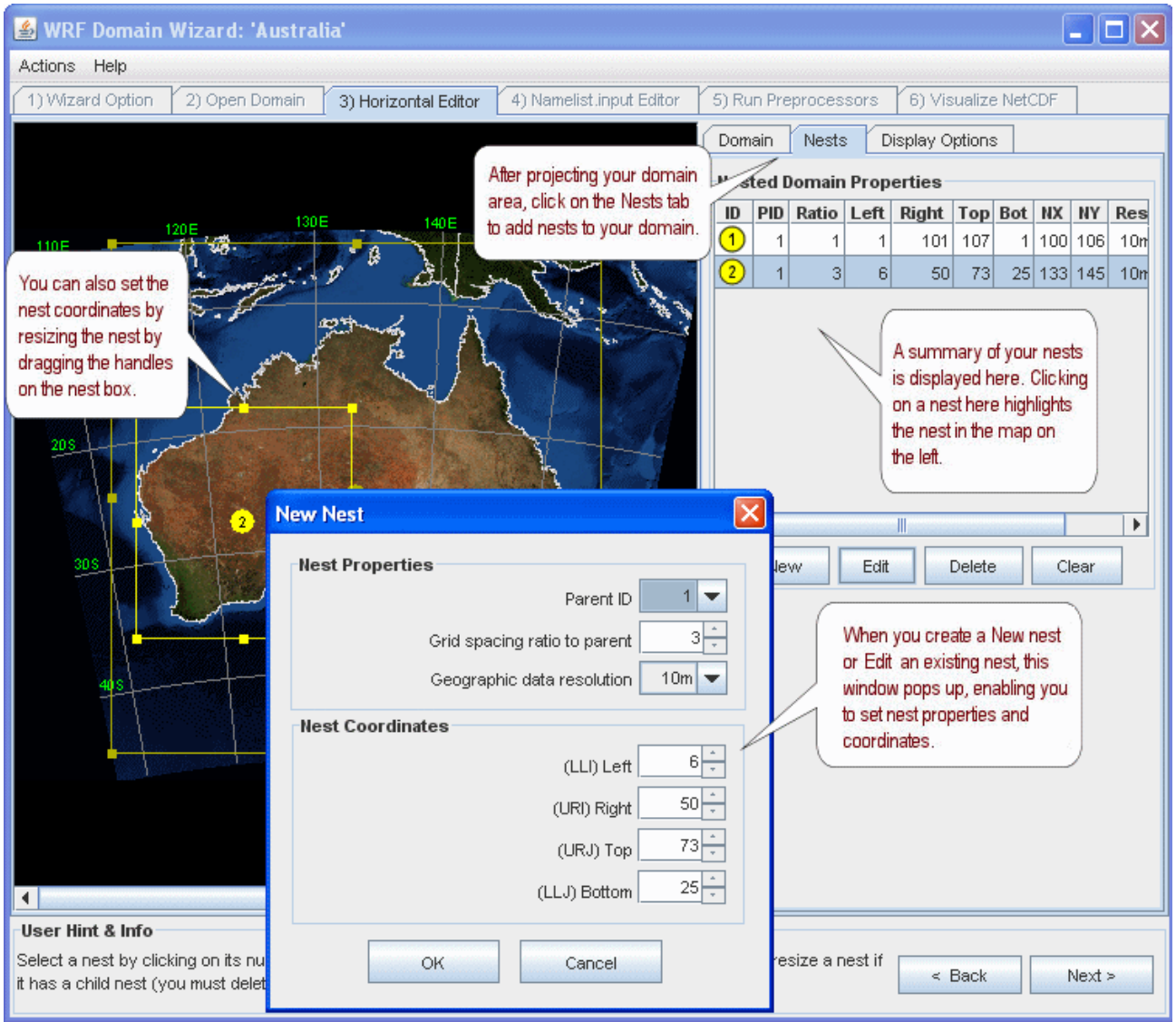
- NOAA Portal simplifies the configuration of domains, model workflows, runs, and monitoring runs, while providing many useful tools to make the modeler's life easier.

# In Conclusion

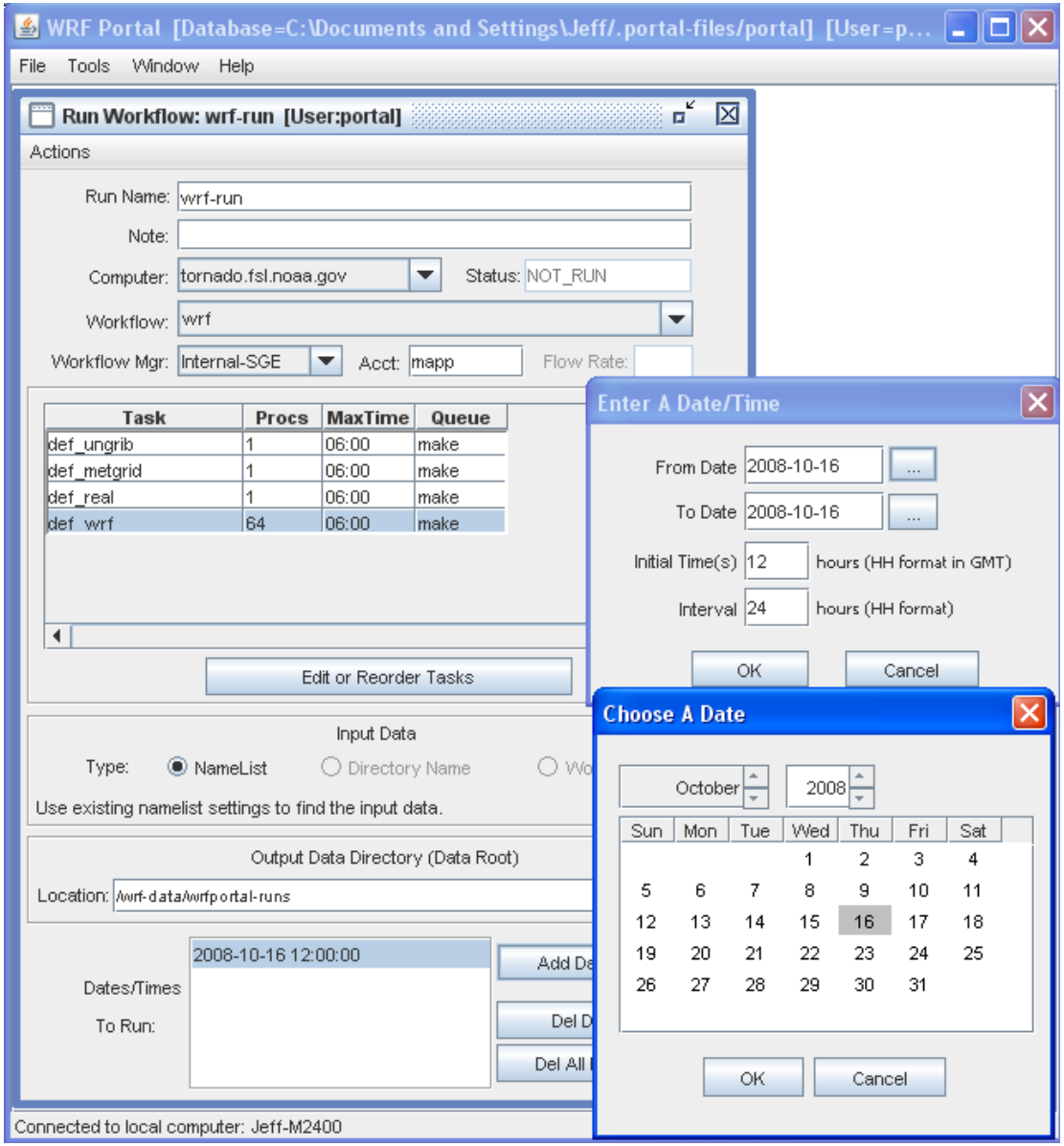
- NOAA Portal simplifies the configuration of domains, model workflows, runs, and monitoring runs, while providing many useful tools to make the modeler's life easier.
- The software is under active development at ESRL and is fully compatible with the latest versions of FIM and WRF3.

*The following slides show portal screen images...*

- **WRF Domain Wizard** screen enables users to easily define and localize domains (for WRF) by selecting a region of the Earth and choosing a map projection, and defining nests



- **Run Workflow** window is where a user selects the workflow, computer, tasks, nbr of procs allocated to each task, & dates



# **Run Monitor** enables a user to follow the progress of runs

WRF Portal [Database=C:\Documents and Settings\Jeff/.portal-files/portal] [User=portal]

File Tools Window Help

**Run Monitor**

**Search Criteria**

Run between ... Model Cfg

And ... Note

Status Computer tornado.fsl.noaa.gov

Search

| Run Config | Run Date | Status | Run Time | Date Started | Elapsed Time | Model |
|---|---|---|---|---|---|---|
| wrf-run1 | 2005-07-11 00:00:00.0 | RUNNING | 00:02 | 2008-10-16 15:25:16.0 | 00:02 | wrf |
| hello-test-run1 | 2008-10-09 12:00:00.0 | DONE | 00:00 | 2008-10-09 17:36:21.0 | 00:00 | hello-test |
| wrf-run | 2008-10-16 12:00:00.0 | ERROR | 00:00 | | | wrf |

**Details for Run Workflow: 'wrf-run1' on 2005-07-11 00:00:00.0**

| Task | Job ID | Job Started | Run Time | Est. Time | Status |
|---|---|---|---|---|---|
| def_ungrib | 186 | 2008-10-16 15:25 MDT | 00:01 | | done |
| def_metgrid | 187 | 2008-10-16 15:26 MDT | 00:01 | | done |
| def_real | 188 | 2008-10-16 15:26 MDT | 00:02 | | running |
| def_wrf | | | | | |

Refresh | View Files/Logs/.nc | Delete | Halt Run | Close

**File Browser**

Local Files (Jeff-M2400) | tornado.fsl.noaa.gov

/wrf-data/wrfportal-runs/wrf-run1/2005071100/logs/

| File Name | Attr | File Size | File Date | Owner |
|---|---|---|---|---|
| def_metgrid_200507110000.log | -rw-r--r-- | 1 KB | 2008-10-16 15:26:46 | smith |
| def_real_200507110000.log | -rw-r--r-- | 7 KB | 2008-10-16 15:27:30 | smith |
| def_ungrib_200507110000.log | -rw-r--r-- | 92 KB | 2008-10-16 15:25:36 | smith |
| workflow.log | -rw-r--r-- | 1 KB | 2008-10-16 15:28:17 | smith |

View as Text | View NetCDF | Close

Connected to remote computer: tornado.fsl.noaa.gov

# NetCDF Viewer

WRF Portal [Database=C:\Documents and Settings\Jeff/.portal-files/portal] [User=portal]

File Tools Window Help

File Browser

Local Files (Jeff-M2400) tornado.fsl.noaa.gov

/wrf-data/wrfportal-runs/wrf-run1/2005071100/wps-output/

| File Name | Attr | File Size | File Date | Owner |
|---|---|---|---|---|
| namelist.wps | -rw-rw-r-- | 2 KB | 2008-10-16 15:26:29 | smith |
| geo_em.d01.nc | -rw-r--r-- | 3,146 KB | 2008-10-10 16:26:25 | smith |
| geo_em.d02.nc | -rw-r--r-- | 5,508 KB | 2008-10-10 16:26:27 | smith |
| NAM:2005-07-11_00 | -rw-rw-r-- | 20,187 KB | 2008-10-16 15:25:32 | smith |
| NAM:2005-07-11_03 | -rw-rw-r-- | 20,187 KB | 2008-10-16 15:25:33 | smith |

View as Text | View NetCDF | Close

Datasets Browser

File Edit Plot Window Help

Create Plot | Target | <new plot window>

Datasets & Variables

| Name | Long Name | Type |
|---|---|---|
| temp.nc | temp.nc | Local Fi |
| ALBEDO12M | - | [lon][la |
| CLAT | - | [lon][la |
| CLONG | - | [lon][la |
| COSALPHA | - | [lon][la |
| E | - | [lon][la |
| F | - | [lon][la |
| GREENFRAC | - | [lon][la |
| HGT_M | - | [lon][la |
| HGT_U | - | [lon][la |

List: Only Plottable Variables

GREENFRAC in temp

File Edit Plot Window Help

Plot 1: GREENFRAC

Array(s) Scale Map Contours Colors Captions

Plot Map of Array 1 Only Interpolate

Array 1: GREENFRAC

Synthesized time coordinate from Times(time): 1 of 1 = 1905-05-03 18:35:44

Month 5 of 12 = 5 of 12

Connected to remote computer: tornado.fsl.noaa.gov

# **Diff Tool** compares workflows, runs, text files (e.g. namelists)

WRF Portal [Database=C:\Documents and Settings\jssmith/.portal-files/portal] [User=portal]

File Tools Window Help

Diff Tool

Type
- Workflow
- Workflow Run
- Text files from disk

Workflows To Diff

| | | | |
|---|---|---|---|
| 1 | portal | hep.fsl.noaa.gov | wrf-EastCoast |
| 2 | portal | hep.fsl.noaa.gov | wrf-EastCoast2 |

Diff

Clear

```
namelist.input

FILES ARE IDENTICAL


namelist.wps
&share
wrf_core = 'ARW',
max_dom = 1,
start_date = '2006-01-10_00:00:00',
end_date = '2006-01-10_12:00:00',
interval_seconds = 10800,
io_form_geogrid = 2,
opt_output_from_geogrid_path = '/export/jeff/domains/E
debug_level = 0,
/

&geogrid
parent_id = 1,
parent_grid_ratio = 1,
```

```
namelist.input

FILES ARE IDENTICAL


namelist.wps
&share
wrf_core = 'ARW',
max_dom = 1,
start_date = '2005-07-11_00:00:00',
end_date = '2005-07-11_12:00:00',
interval_seconds = 10800,
io_form_geogrid = 2,
opt_output_from_geogrid_path = '/export/jeff/domains/EastCoast/'
debug_level = 0,
/

&geogrid
parent_id = 1,
parent_grid_ratio = 1,
```

Connected to local computer: Tornado2

Thank you!