

Enabling Open Access to High-Resolution Global Models Running in Real-time Research Mode

The HIWPP Open Data Initiative

Bonny Strong, NOAA, Boulder, CO, and CIRA/CSU, Fort Collins, CO,
T. Schneider, R. C. Lipschutz, J. Stewart, S. G. Benjamin, C. DeLuca, J.
Valdez, and C. MacDermaid



Context: HIWPP

High Impact Weather Prediction Project

A “Sandy Supplemental”
project



Hurricane Sandy, October 28, 2012; image captured by
GOES-13 satellite.
NASA Earth Observatory image by Robert Simmon

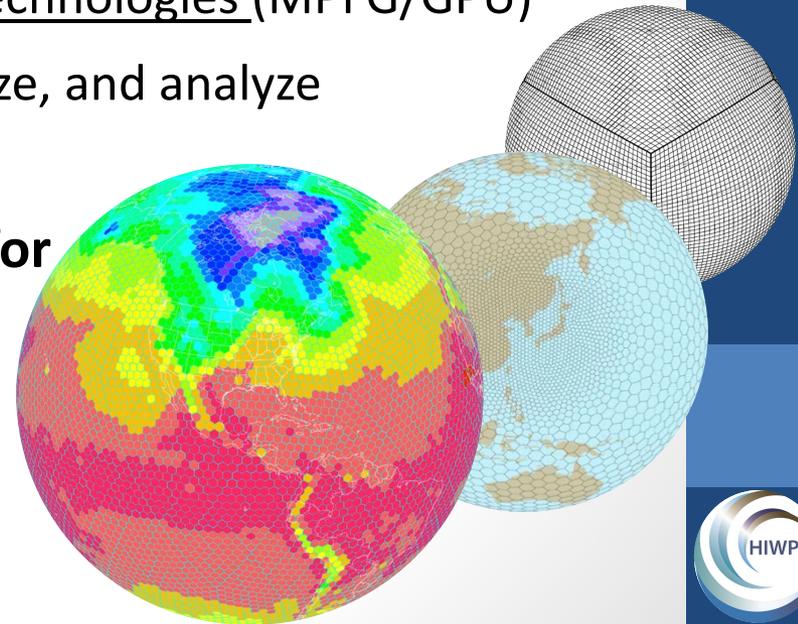
Goal: develop the world’s best medium-range weather forecast model by the end of the decade to improve our time-zero to two-week prediction of nature's most dangerous storms such as hurricanes, floods, and blizzards, over the whole globe.

See: hiwpp.noaa.gov



HIWPP Six-Point Strategy

1. Drive the current generation (hydrostatic) of global NWP models to max performance & resolution (10km-20km), to medium-range and beyond
2. In parallel, accelerate the development of the next generation (non-hydrostatic) of high-resolution (3km) medium-range global and nested NWP models
3. Utilize the latest hybrid assimilation techniques (4D-En-Var) and scale-aware physical parameterizations
4. Migrate models to newer and faster HPC technologies (MPFG/GPU)
5. Provide new tools to quickly access, visualize, and analyze amounts of gridded data
6. **Engage with the weather community for feedback and an open process**



Open Data Initiative created

- Aligns with NOAA and U.S. government Open Data policies

“Making information resources accessible, discoverable, and usable by the public can help fuel entrepreneurship, innovation, and scientific discovery”

- 1) First, provide a mechanism for interaction between public, private, and academic sectors as we develop and explore models that are relatively mature in their developmental life cycle
- 2) Second, the community can help us to improve these models, leading to a more robust and effective product development cycle.

[Real-time Data Feed](#)

[NEIS Visualization Tool](#)

Live: February 9, 2015



HIWPP

HIGH IMPACT WEATHER PREDICTION PROJECT

FUNDED BY HURRICANE SANDY DISASTER RELIEF SUPPLEMENTAL APPROPRIATIONS

[HOME](#)

[ABOUT US](#)

[CONTACT US](#)

[RESOURCES](#)



Included Models

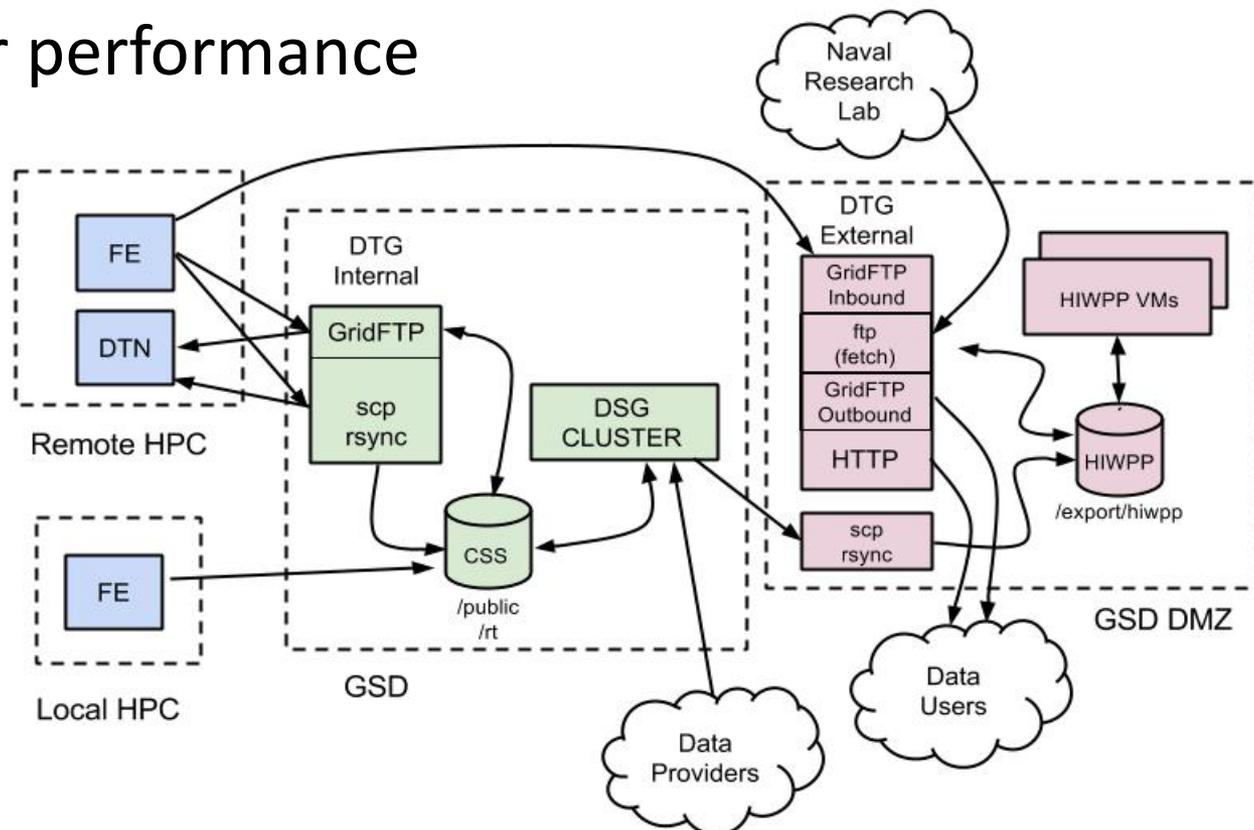
Model	Model Resolution	Output Resolution	Frequency	Notes
FIM9	13 km	1/8 deg	Hrly to 14 days	GFS 2014 physics
FIM9-alternate	13 km	1/8 deg	Hrly to 14 days	GFS 2011 physics (inactive)
GFS	13/27 km	1/4 deg	Hrly to 12 hours, 3-hrly to 16 days	
NAVGEM	21 km	1/4 deg	Hrly to 7 days	
Multi-model		1/2 deg	6-hrly to 7 days	Post-processed

“Alternate” allows for running different physics or data assimilation



Challenges and concerns

- Volume of data
- Distribution of data
- Network bandwidth
- Transfer performance
- User authentication
- Resources needed for user support



Solutions: New Hardware

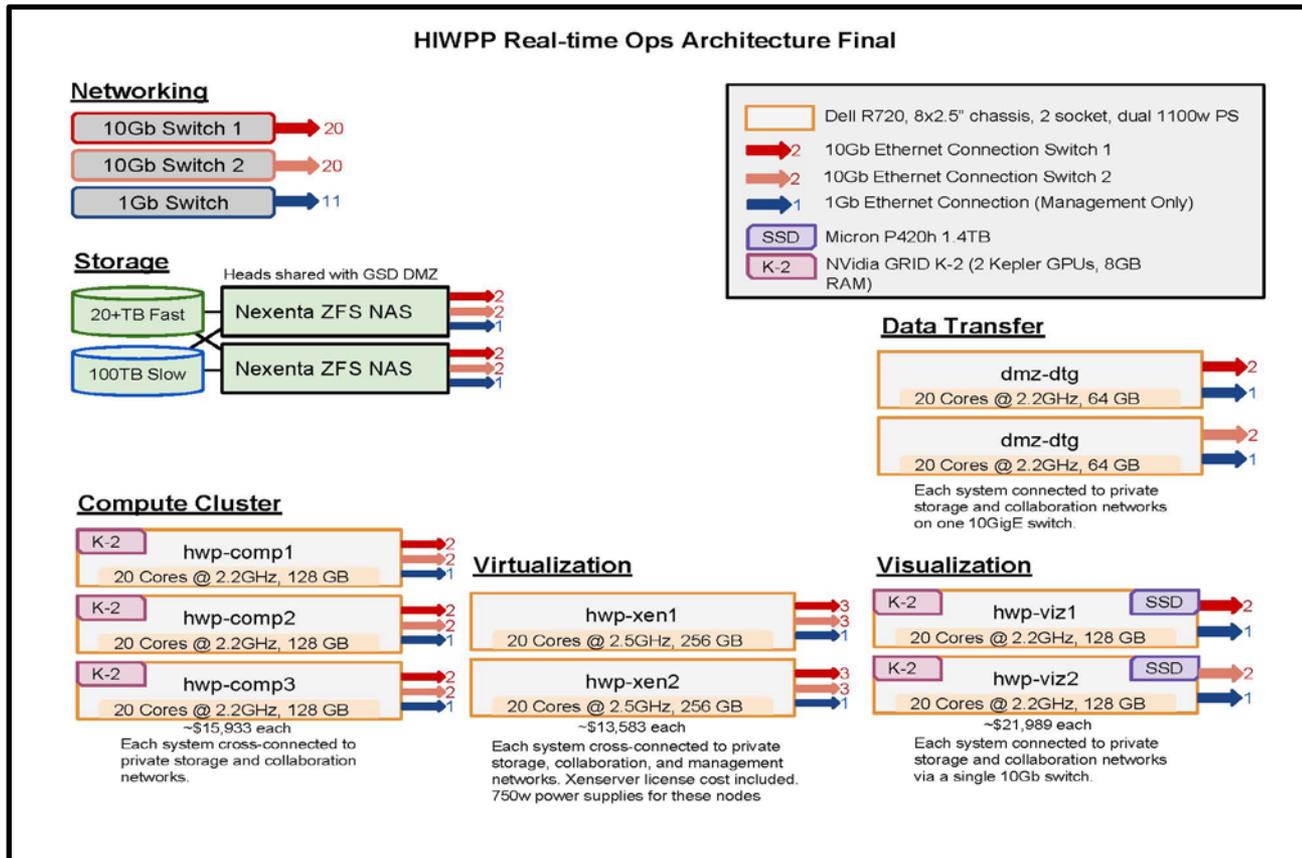
Data storage system:

- Fast disk: 20 TB
- Standard disk: 100 TB

Networking Switches

Processing nodes:

- Visualization: 7 nodes
- Data transfer : 2 nodes



Solutions: Bandwidth Management

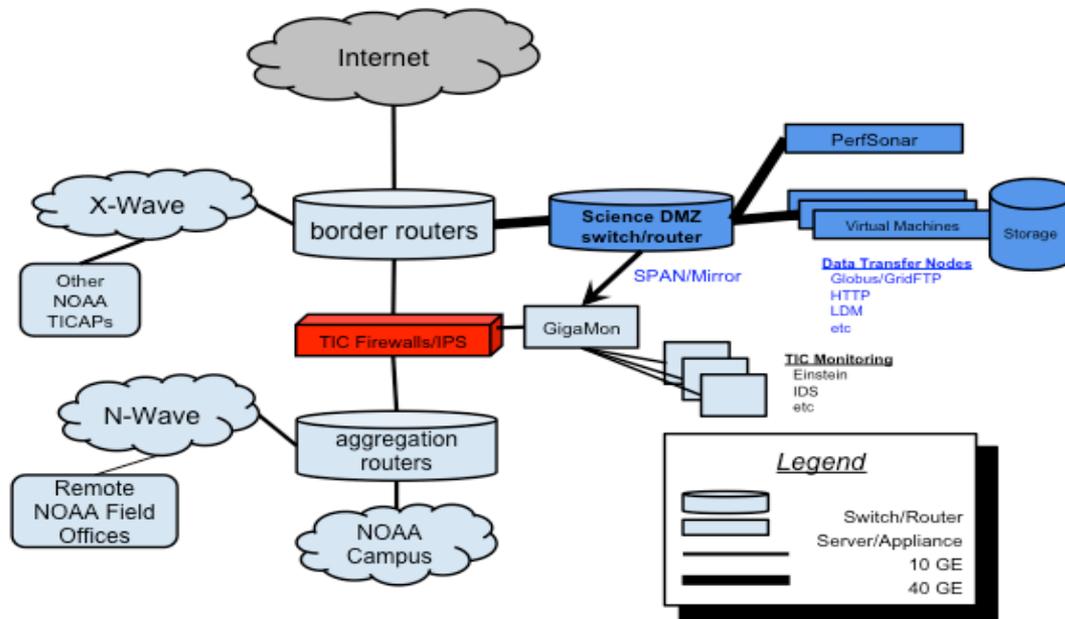
Options:

1) Cloud

2) Science DMZ

3) Restrict bandwidth

Typical Science DMZ



In the end: Not a Problem

Solutions: Software Tools

- For managing internal data transfers between sites: **GridFTP** from **Globus**

<http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/>

- For managing user data access with ability to select data sub-sets: **THREDDS** from **Unidata**

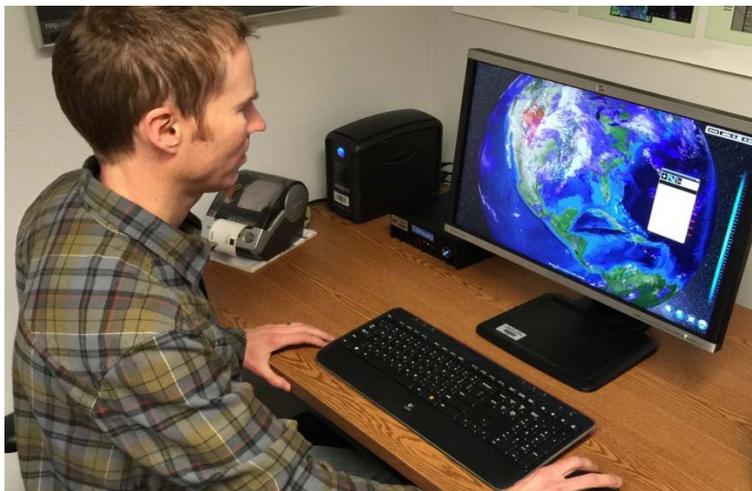
<http://www.unidata.ucar.edu/software/thredds/current/tds/>

- For managing user registration and authentication: **CoG** from **NOAA/NESII**

<https://www.earthsystemcog.org/projects/cog/>



Open Data Initiative Users



Total Registered Users	56
Commercial users	18
NOAA users	13
Academic users	12
Other gov't agencies, CIs	13

Count	gbytes	Org name
5,644	1,119	Minnesota education cooperative
5,528	804	Meteociel
10,615	577	Weatherbell
30,295	74	Weather Company
52,082	2,577	Total

**Usage 9/10/15
- 10/10/15**

User Feedback

Users Meeting Sep 2015

- It has been extremely helpful to have forecasts at 1-hr intervals out to 14 days.
- Early access is “huge”, especially to a third party distributor .
- Would really like to know if there will be early access to the new dycores, MPAS and FV3. They need years in advance to prepare for a new operational model.
- Even relatively low quality data is useful
- Would be reluctant to participate further if no data will be available beyond Feb 2016.

What Did We Learn?

Users:

- Early access to research model data was highly welcomed by private enterprise
- Reliability was not a concern
- Most active users were companies with a web-based product that included weather information

Engagement with modelers:

- Users did not engage with modelers or provide any feedback about the models through the forum
- Some engagement did occur directly with some researchers

Support required:

- User support was not a significant issue for real-time data
- High-resolution visualization was difficult to support for bandwidth available to many users

Unexpected outcomes

Benefit for research collaboration:

- Infrastructure for public access was very useful for sharing dycore test output for researchers

Issues with volume of data:

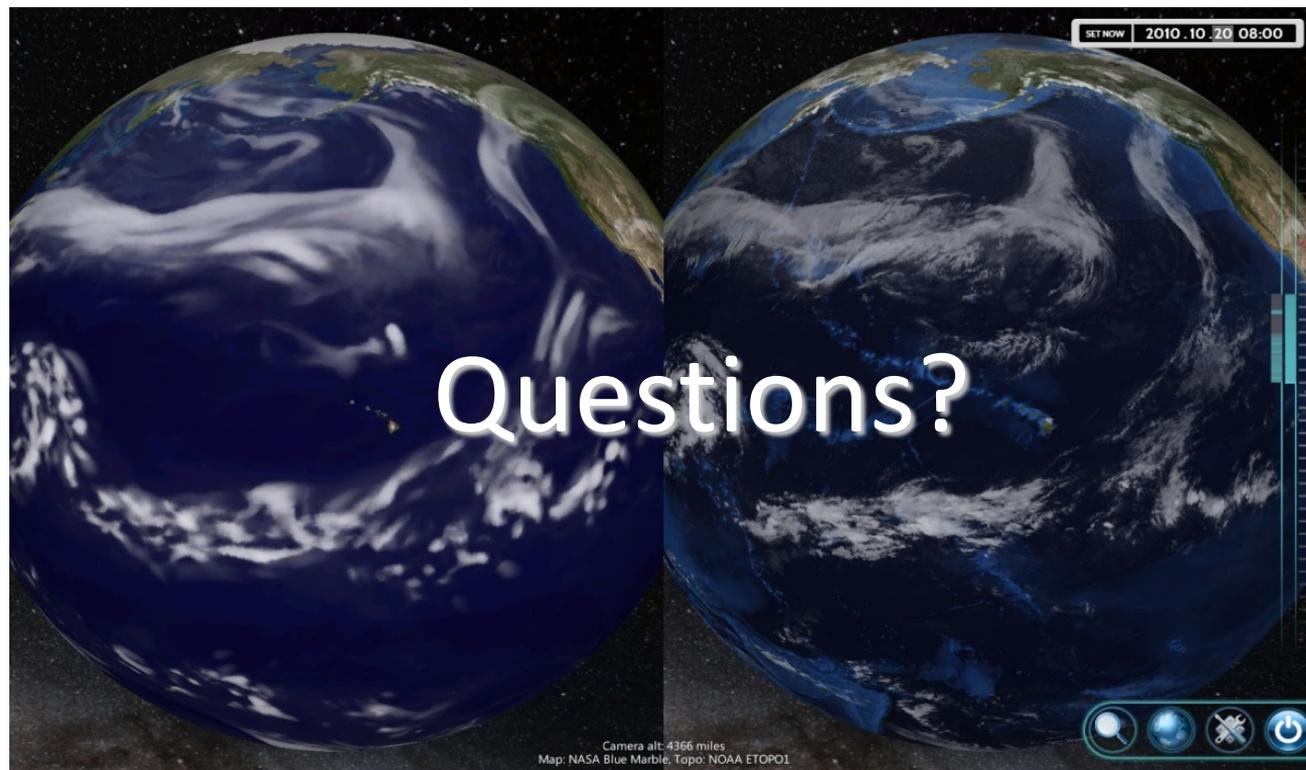
- Issues with quantity of data produced by high resolution non-hydrostatic models had not been fully anticipated
- This is likely to be a significant issue in many aspects of model research in the near future

What Next?

- HIWPP project is in its final year and Open Data Initiative will conclude on **Feb 8, 2016**
- **BUT** – access to real-time research model data will carry on under **NOAA's Big Data Partnership**
 - For further info, email Amy.Gaskins@noaa.gov
- Infrastructure will continue to be used to support NOAA research collaboration
- FIM data is expected to remain available as long as:
 - 1) the model runs in real-time at NOAA/ESRL/GSD, and
 - 2) resources are available to support it

For Further Info:

- Website: hiwpp.noaa.gov
- Email: hiwpp.support@noaa.gov
- NEIS: contact Jebb.Stewart@noaa.gov
- Big Data Partnership: Amy.Gaskins@noaa.gov



AMS Forecast Improvement Group



Key recommendations:

4) *“Support active collaboration between public, private, and academic sectors with models particularly in their final stages of development”*

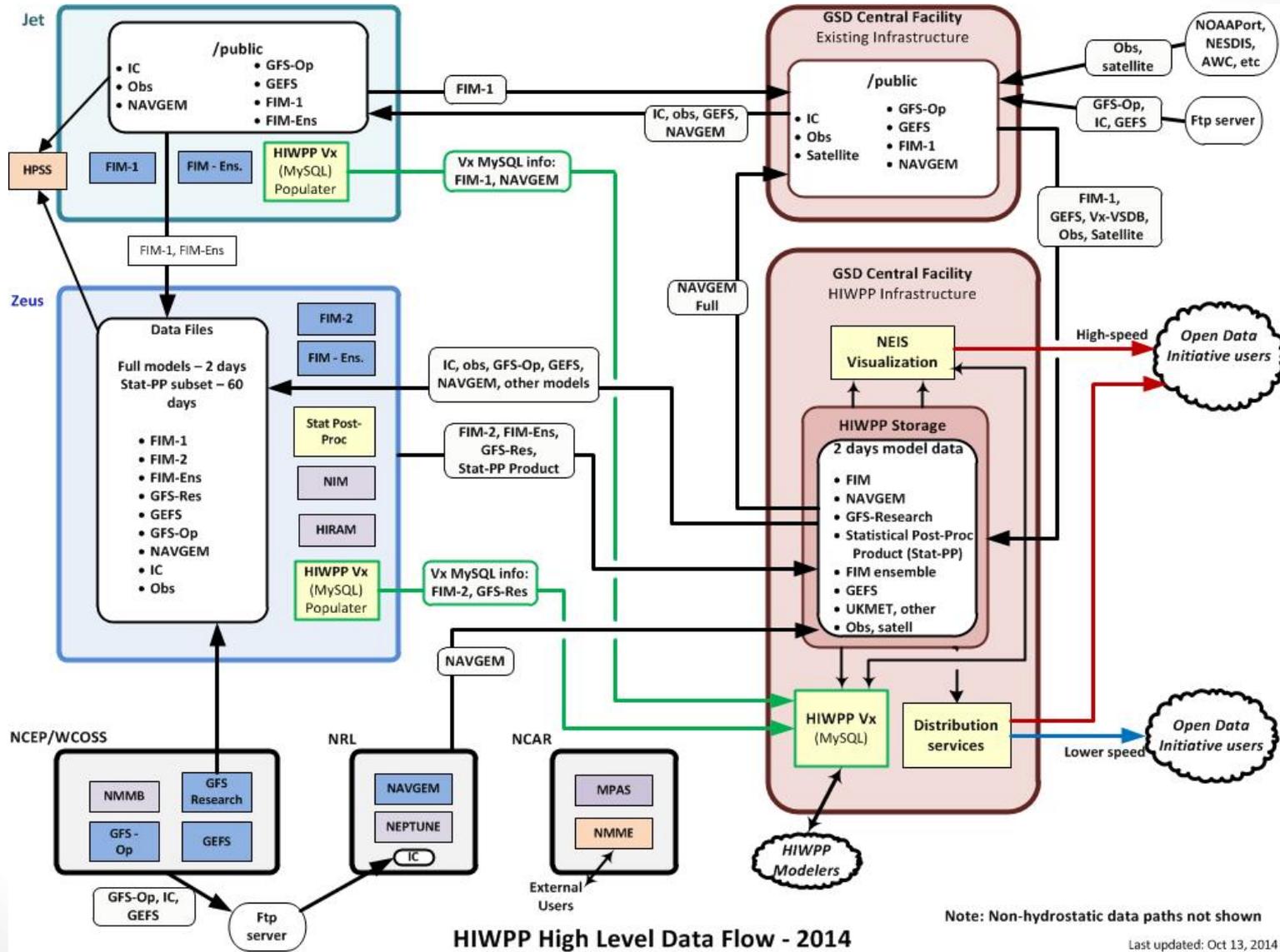
From <http://www2.ametsoc.org/cwwce>

Included Models

Model Name	Agency	Description	Model Resolution (40o lat)	Output Resolution and Freq.	Vert. Levels	Model Init Times
FIM9Zeus	NOAA/ESRL	Icosahedral grid; hybrid isentropic-sigma vertical coordinate; finite volume core. GFS 2011 physics, with updates. More info at http://fim.noaa.gov/	14 km	1/8° 1-hrly to 14 days	64	00Z 12Z
FIM9	NOAA/ESRL	Same as FIM9Zeus with GFS 2015/T1534 physics See http://fim.noaa.gov/ for latest config.	14 km	1/8° 1-hrly to 14 days	64	00Z 12Z
GFS	NOAA/NWS	Latest NWS global forecast model, scheduled to become operational on Jan 14, 2015. For details see the NWS Technical Implementation Notice Spectral core (spherical harmonic basis functions w/transformation to a Gaussian grid); vertical sigma pressure hybrid coordinates; GFS 2014 physics.	13 km (T1534) out to 10 days 27 km (T574) days 10-16	1/4° 1-hrly to 12 hours 3-hrly to 16 days	64	00Z 06Z 12Z 18Z
NAVGEN	Navy NRL	Semi-Lagrangian, semi-implicit core; vertical sigma pressure hybrid coordinates; NAVGEN physics package.	21 km	1/4° 1-hrly to 7 days	64	00Z 12Z
Stat-PP	HIWPP	Weighted mean of bias-removed members from HIWPP deterministic models, with probability distributions at each grid point.	NA	1/2° 6-hrly to 7 days	Surface variables and 500mb height	00Z 12Z

All models use GFS T1534 initial conditions.

Data Flow



Estimating Data Storage Requirements

Parameters:

- GRIB2 output
- 1/8th degree lat/lon
- 43 Isobaric levels (1000 - 50 mb, by 25 mb)
- 3-D variables: HGT, TMP, RH, U, V, VVEL
- 2-D variables: 22 fields (precip, sfc T, sfc fluxes, cloud base/top, etc.)
- 2 cycles/day, hourly forecasts 0-168 hrs (7 days), 6-hourly 174-336 hrs (7-14 days)
- This gives an approximate file size of 530 MB.

Total for 1 day for one model each system:

- FIM 208 GB
- GFS 24 GB to 52 GB
- NAVGEM 52 GB

Further assumptions for estimates include:

- Design to handle 200 GB/day for hydrostatic models.
- Runs for 2 days will be kept online in HIWPP main storage.
- Number of hydrostatic models will be at least 3, possibly 6
- Non-hydrostatic model data to be kept online will be evaluated at the beginning of year 2
- Statistical post-processing product will encompass 5 2-D vars for year 1, but may expand



Data Volumes Planned

Hydrostatic models:	2.4 TB
6 models x 200GB/day x 2 days	
Non-hydrostatic models	2.0 TB
Ensembles and Statistical Post-Processing	< 0.1 TB
Obs, satellite, other	< 0.5 TB
Retrospective working space	100 TB
Total	110 TB

Solutions: User Authentication – CoG

NOAA's NESII team, headed by Cecelia DeLuca, developed new features in their CoG tool to support user authentication via OpenID, forums, and other enhancements .

Hosted by  National Oceanographic and Atmospheric Administration
Earth System Research Laboratory

Powered by  ESGF and CoG

Welcome, **Bonny-test** | [Join the HIWPP_OpenData project](#) | [Register a New Project](#) | [My Profile](#) | [Log out](#)

HIWPP Open Data Initiative

You are at the NOAA ESRL ESGF-CoG node

[Home](#) [Forum](#) [Contact Us](#) [Technical Support](#)

[My Data Cart \(0\)](#)

HIWPP_OpenData

- Home
- Open Data Initiative Info
 - Info & Instructions
 - Real-time Data Access
 - FAQ
 - Model Change Log
 - Stat. Post-Processing
 - Feedback
- Visitors
 - List All News
 - List All Files
- Contributors
 - Add Page
 - Publish News
 - Add File
 - Add Resource

HIWPP Open Data Initiative

Welcome to the HIWPP Open Data Initiative

Important Announcements:

Please see the [Info & Instructions](#) page for announcements such as upcoming down times.

Real-Time Research

The HIWPP Open Data Initiative provides a mechanism for interaction between public, private, and academic sectors with our project as we develop and explore models that are relatively mature in their developmental life cycle.

Feedback for Model Development

We hope that you, the community, will help us to improve these models. Feedback on the model forecasts, the post-processed data, and the data delivery and visualization tools will be used to improve models and products.

[Provide Feedback](#)

Data Use

First and foremost, these are not the official products and forecasts of NOAA – those products and services are issued by the National Weather Service, and can be found at:

- Official Forecasts (NWS): <http://www.weather.gov>
- Official Model Products (NWS/NCEP): <http://mag.ncep.noaa.gov>

Read News

First Open Data Users Meeting
The first Open Data Users Meeting was held on Sep 10, 2015. The primary message ...

Dates for Year 1 of Open Data Initiative
The Open Data Initiative first live on Feb 9, 2015 and is currently

Browse Projects

This All My Tags

Parent projects (1)
HIWPP

Peer projects (0)

Child projects (0)

Enter Tag

[Save Tag](#) [Reset](#) [Go](#)

Start typing, or use the 'Delete' key to show all available tags.

HIWPP_OpenData Tags: None

