Multi-resolution Nested Dust Forecast System Feasibility Study

Karl Benedict (PI)
Earth Data Analysis Center, University of New Mexico
Chowei Yang (Co-I), Qunying Huang
Center for Intelligent Spatial Computing (CISC), George Mason University

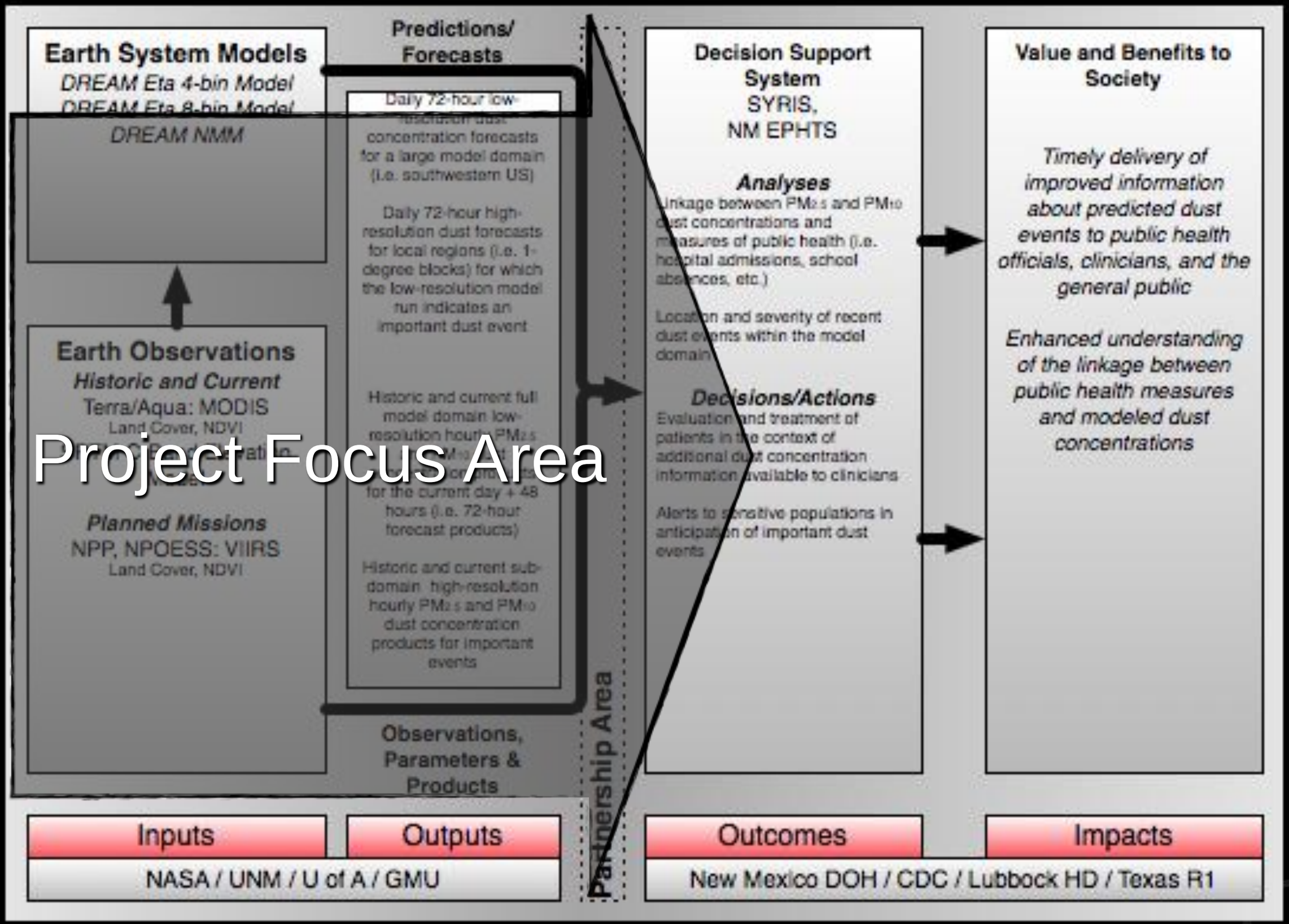
Project Background

- ✓ Public Health Applications
  - ✓ Public Health Applications in Remote Sensing (PHAiRS - NASA REASoN): 2003-2008
  - ✓ Adding NASA Earth Science Results to EPHTN via the NM/EPHT System (ENPHASYS - NASA DECISIONS): 2008-2011
- ✓ Interoperability Development & Testing
  - ✓ NASA GIO/PHAiRS Project Interoperability and High Performance Computing Test/Demonstration: 2007-2008

Goals

- ✓ Work with existing modeling cores (DREAM ETA-8, DREAM NMM)
- ✓ Modify model pre- and post-processors to support OGC and REST data transfer
- ✓ Develop algorithm for automated generation of dust forecast area(s) of interest
- ✓ Evaluate and report on performance characteristics of the nested model system

Integrated System Solution Diagram

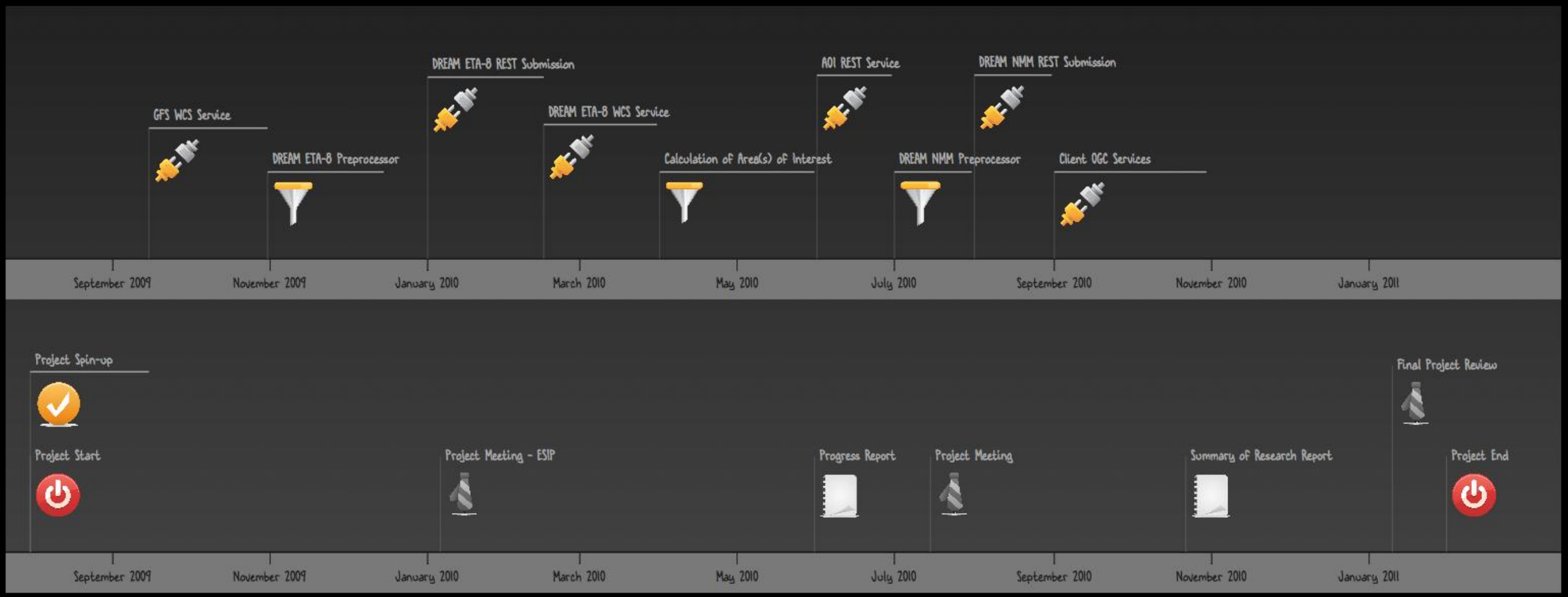


Project Focus Area

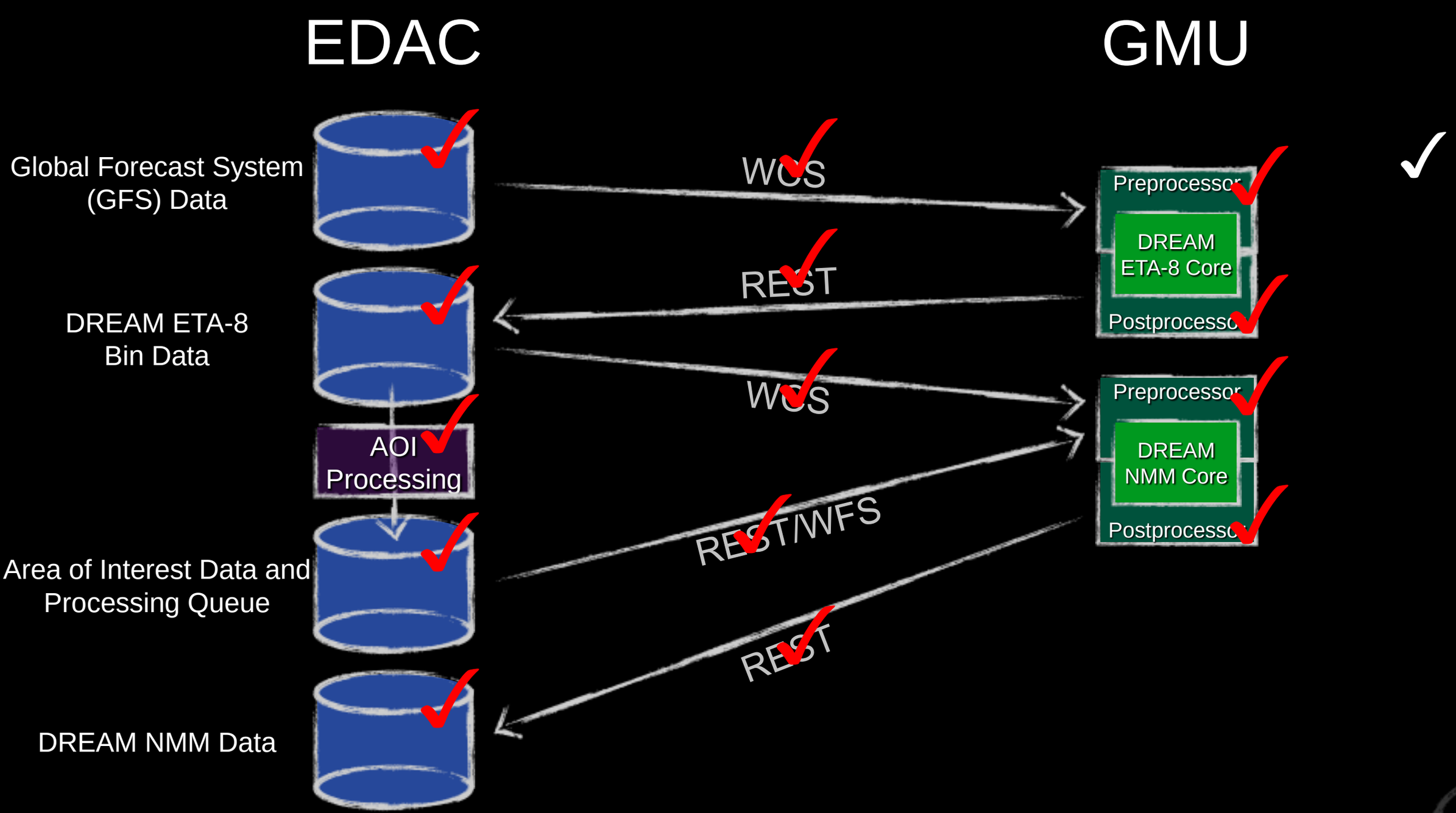
Feasibility Testing

- ✓ Systems Integration
  - ✓ Model pre- and post-processor implementation
  - ✓ Data management and storage
  - ✓ Appropriateness of implemented service standards
- ✓ Performance
  - ✓ Comparison of performance (time-to-delivery) of nested model vs. dedicated large domain/high-resolution model runs

Timeline


GFS WCS Service
DREAM ETA-8 Preprocessor
DREAM ETA-8 REST Submission
DREAM ETA-8 WCS Service
Calculation of Area(s) of Interest
AOI REST Service
DREAM NMM REST Submission
DREAM NMM Preprocessor
Client OGC Services
September 2009
November 2009
January 2010
March 2010
May 2010
July 2010
September 2010
November 2010
January 2011
Project Spin-up
Project Start
Project Meeting - ESIP
Progress Report
Project Meeting
Summary of Research Report
Final Project Review
Project End
September 2009
November 2009
January 2010
March 2010
May 2010
July 2010
September 2010
November 2010
January 2011

Systems Integration



Systems Integration

EDAC

GMU

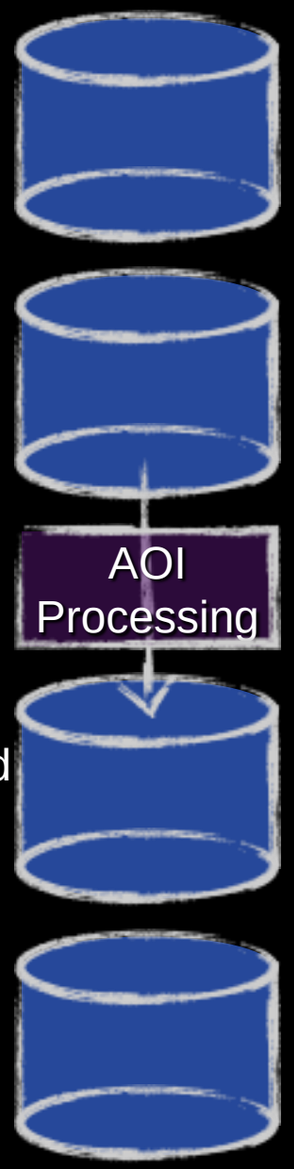
Global Forecast System (GFS) Data

DREAM ETA-8 Bin Data

AOI Processing

Area of Interest Data and Processing Queue

DREAM NMM Data



WCS/WMS

WCS/WMS

REST/HTTP/WFS

WCS/WMS

Preprocessor

DREAM ETA-8 Core

Postprocessor

Preprocessor

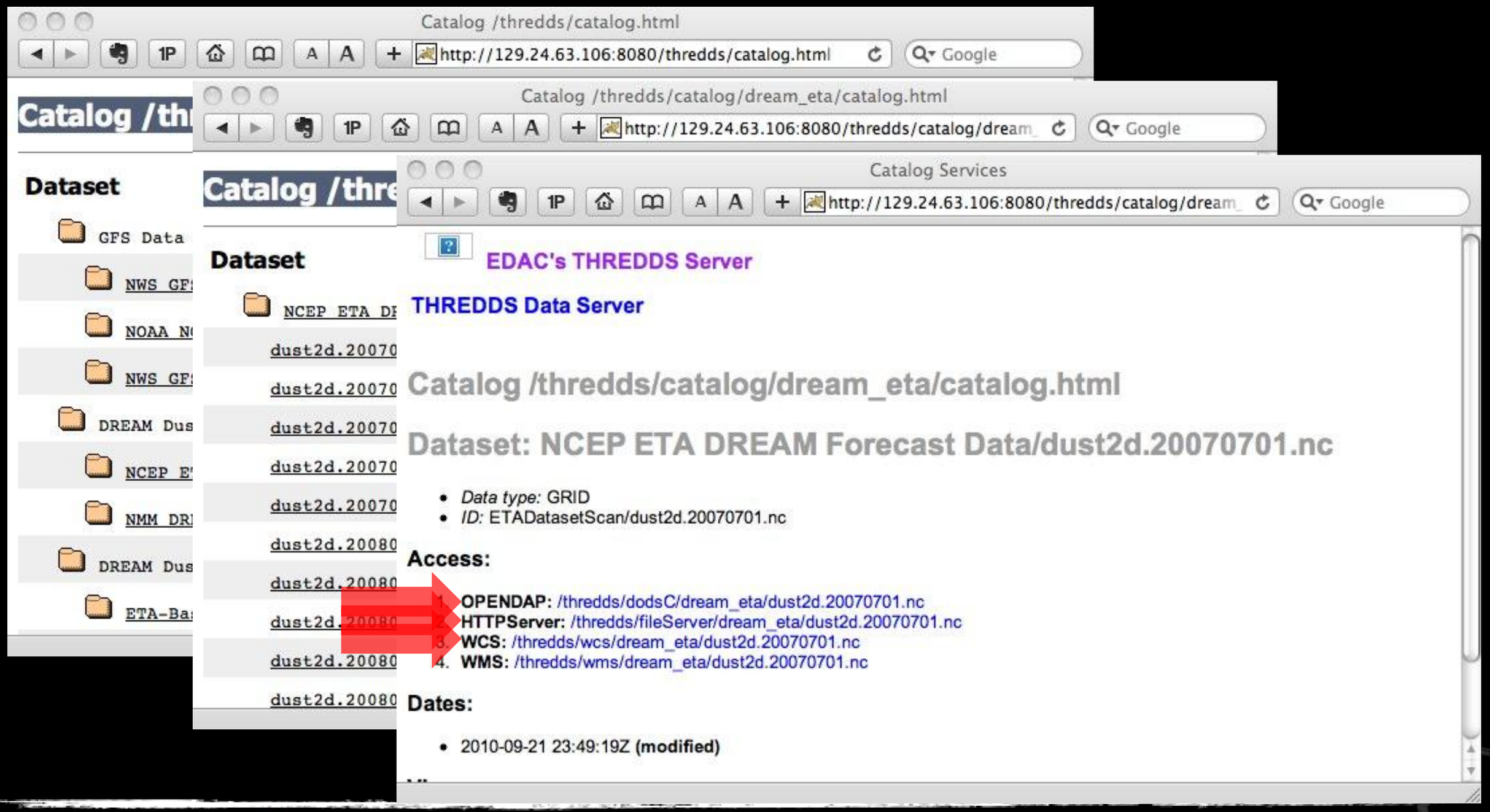
DREAM NMM Core

Postprocessor

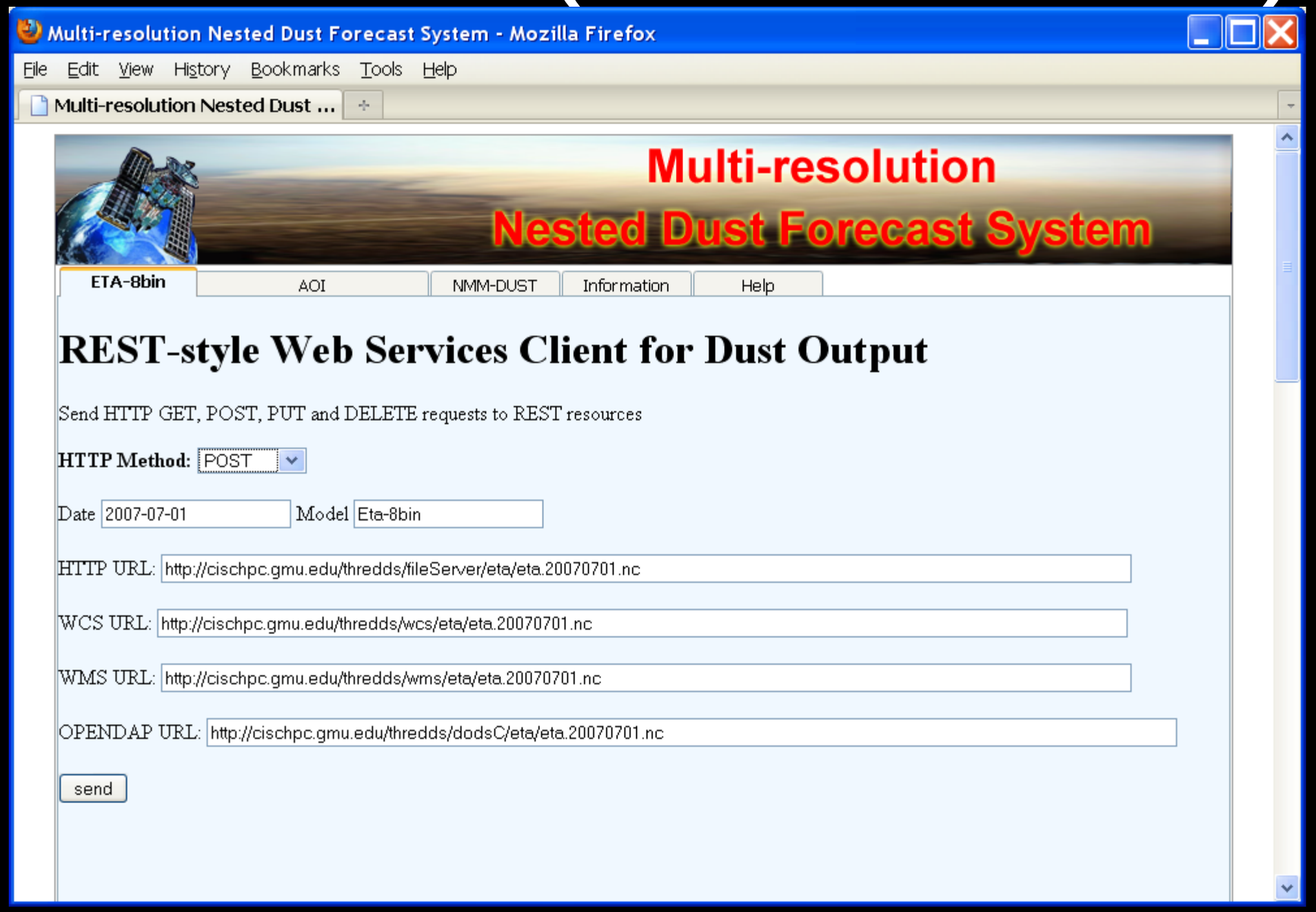
External Users/Systems

What Do These Components Look Like?

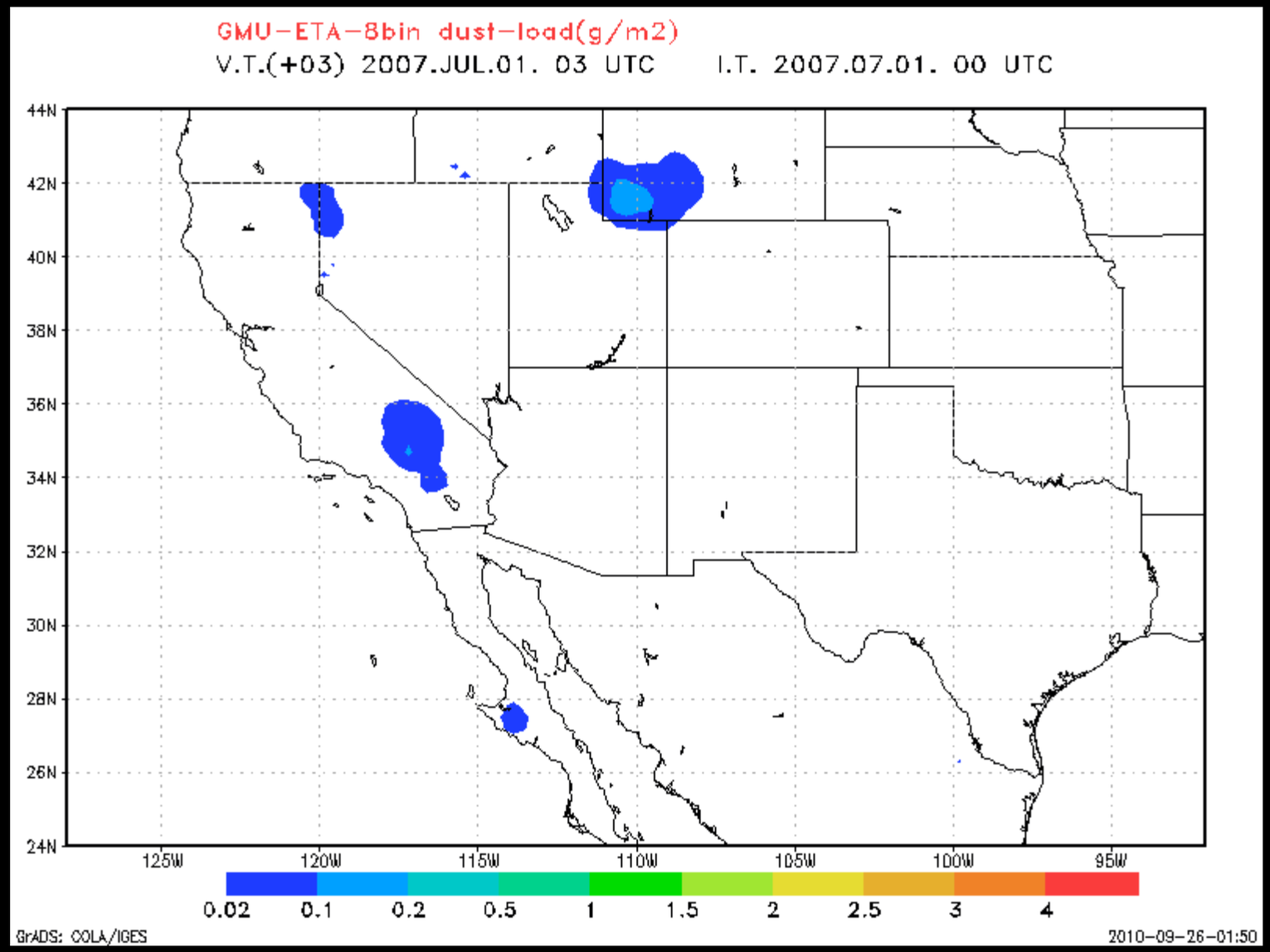
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)



REST Data Upload Services (EDAC/GMU)



DREAM ETA-8 Model



2007-07-01

Identified Areas of Interest

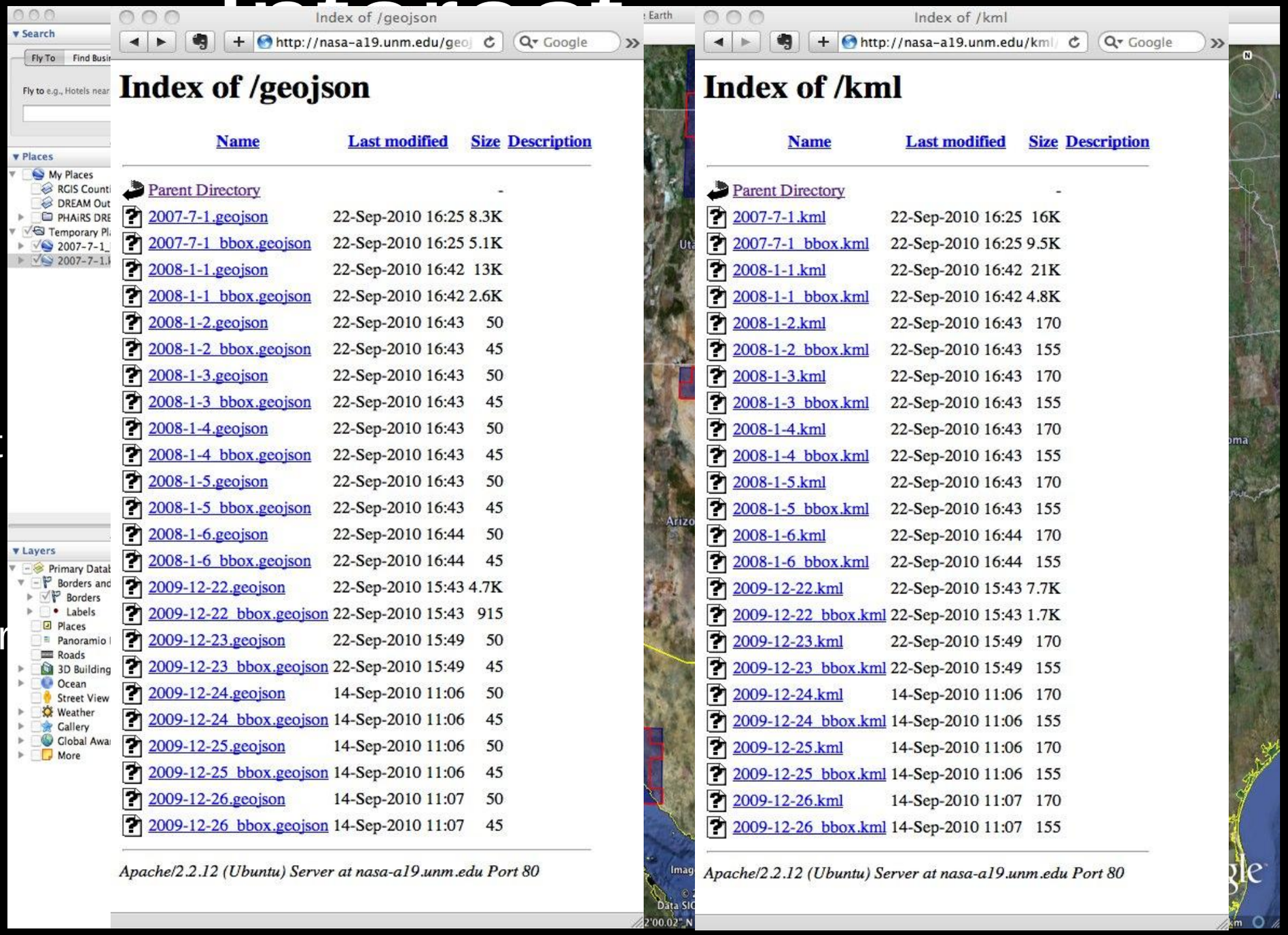
Matrix Processing of NetCDF Model Output

Generation of AOI Raster mask (GeoTiff)

Vectorization of discreet AOIs from raster

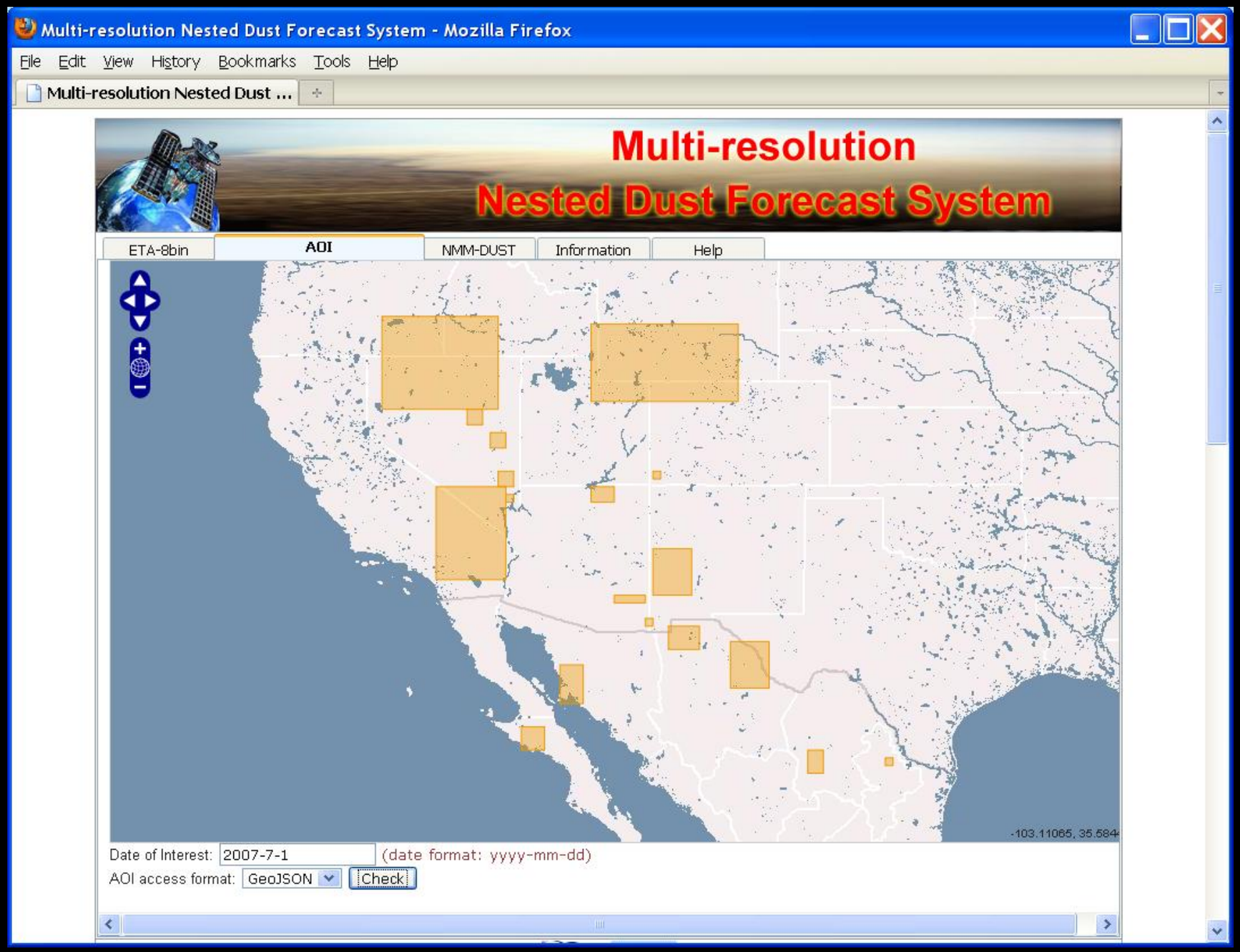
Generation of AOI vector files (KML, GeoJSON)

Publication of vector files via HTTP



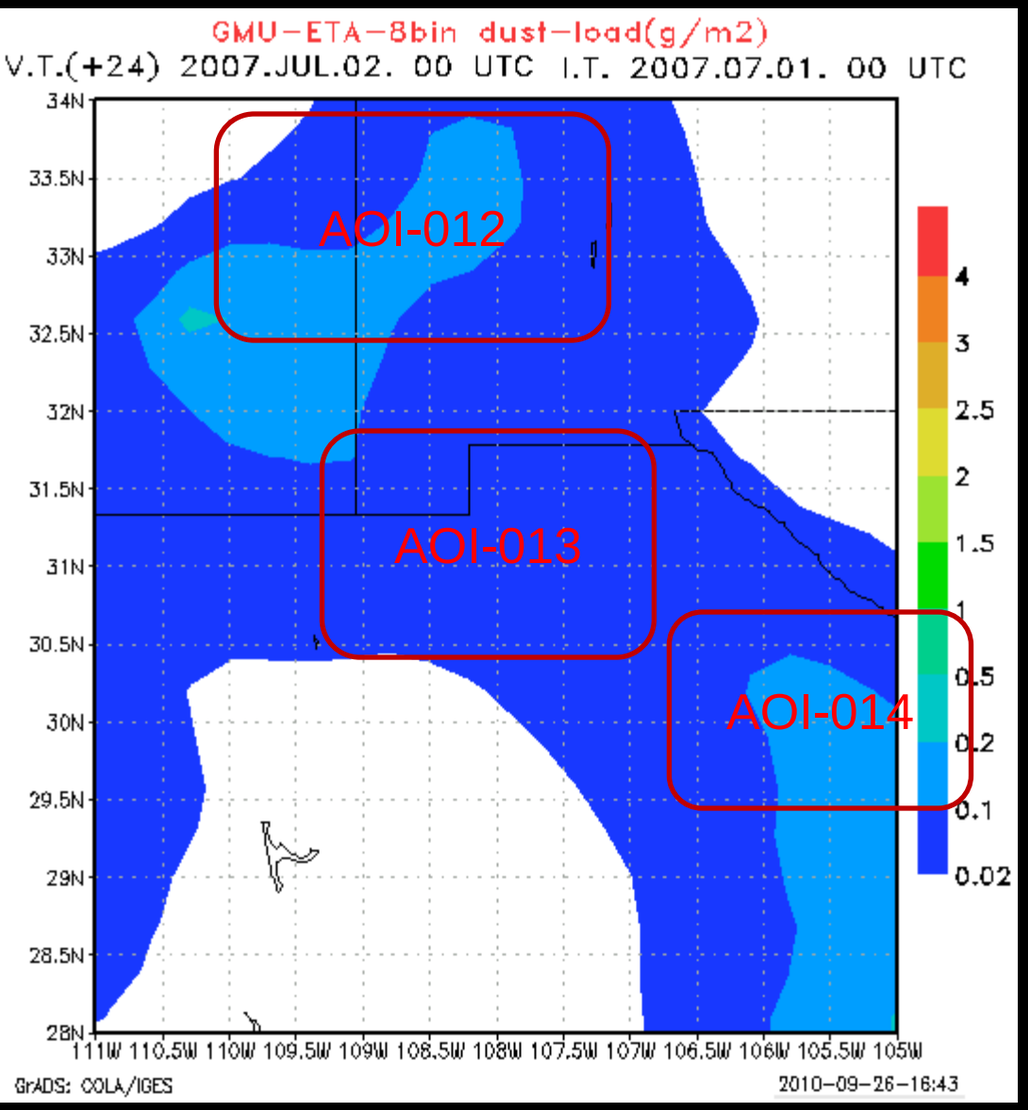
2007-07-01, 72-hour forecast, 3-hour time step, dl (dust loading, gm/m^2),1*10^{-7} base threshold, 8 aggregate thres

AOI Access Client(GeoJson/KML)

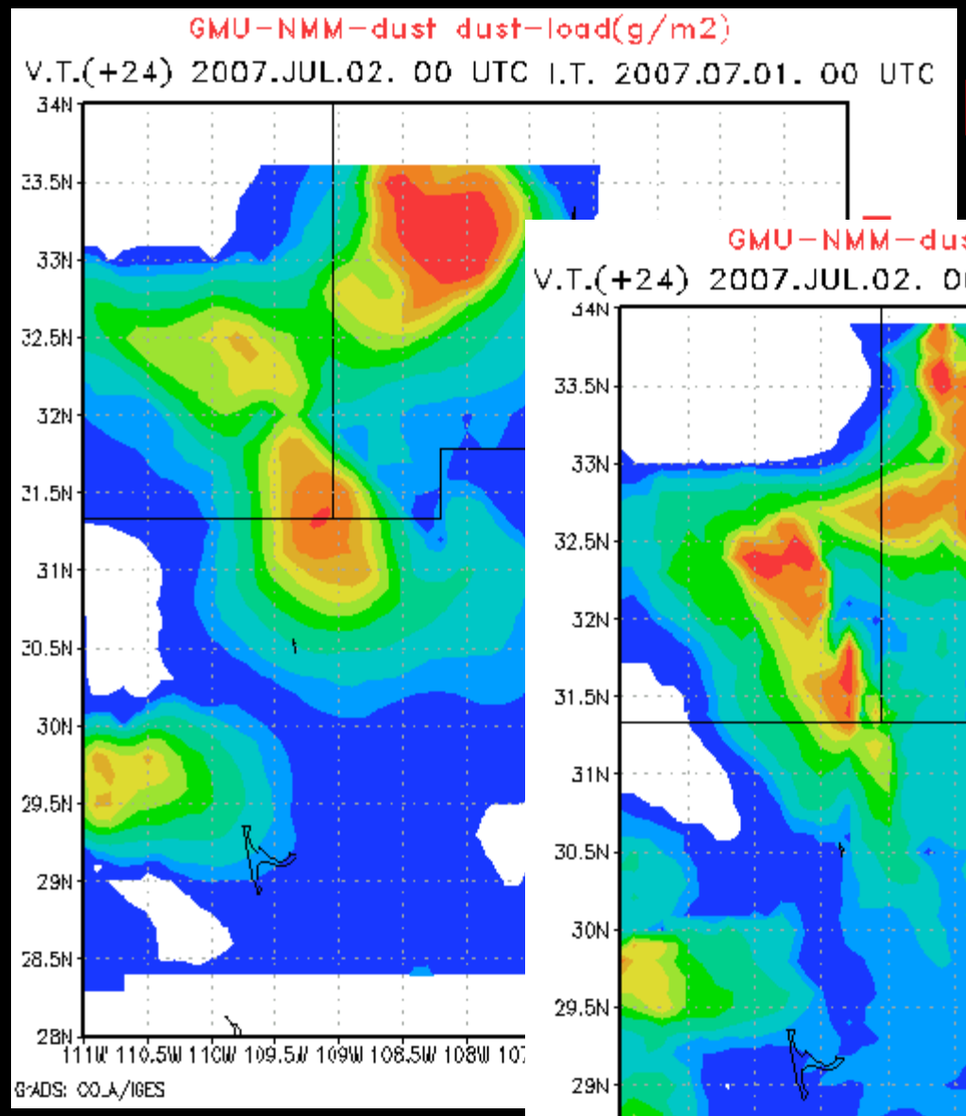


NMM Model

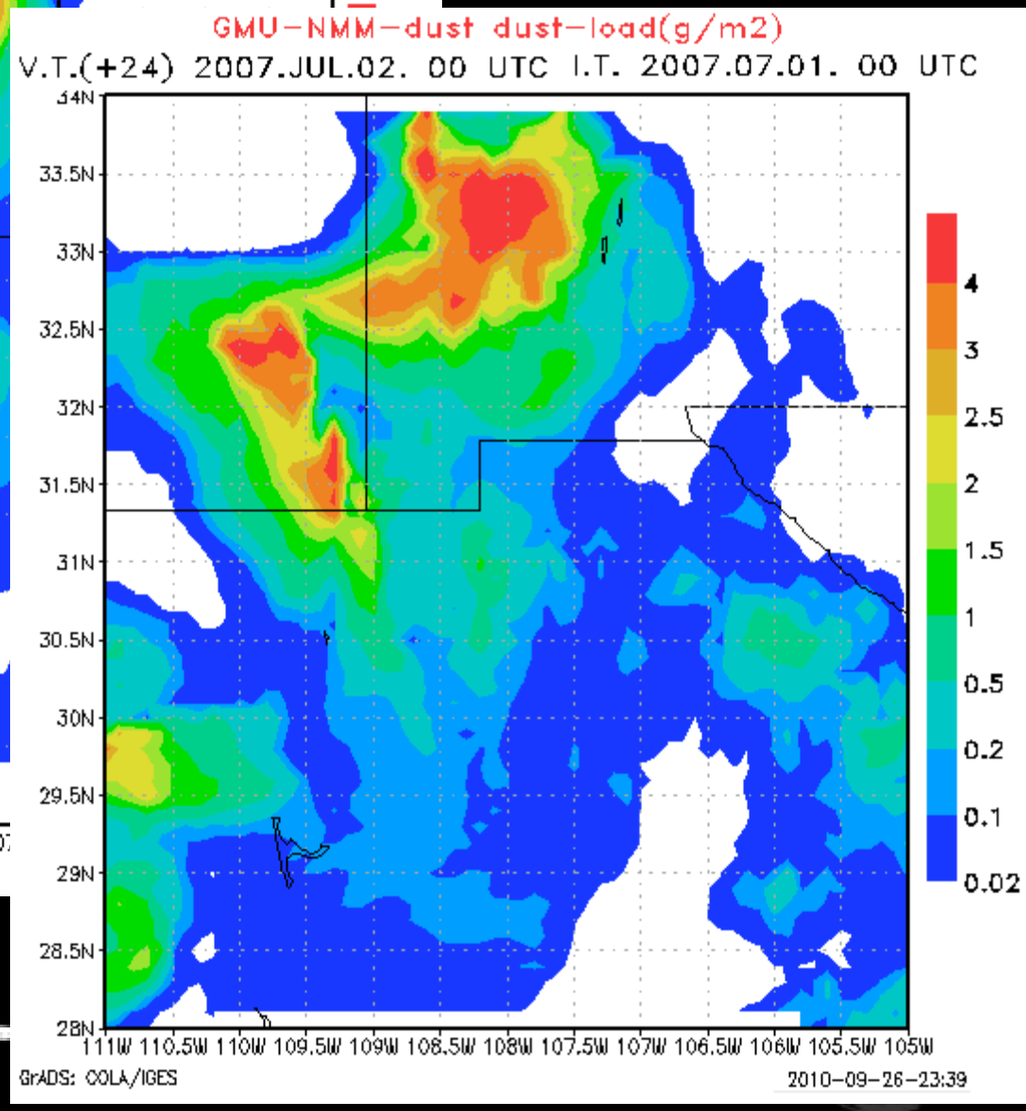
ETA-8bin: 50km



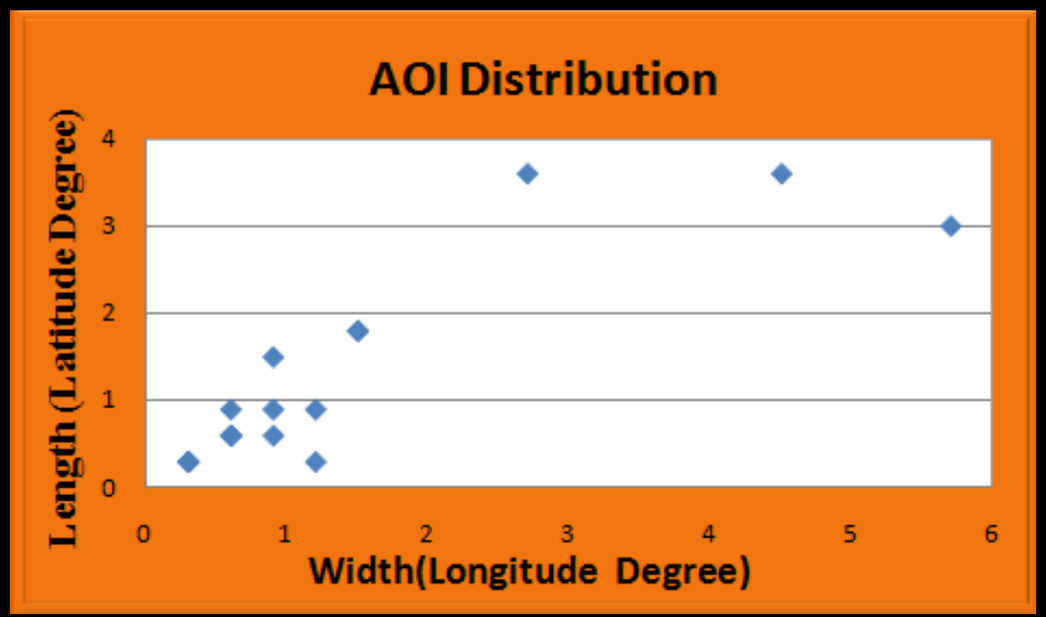
NMM-dust: 22km

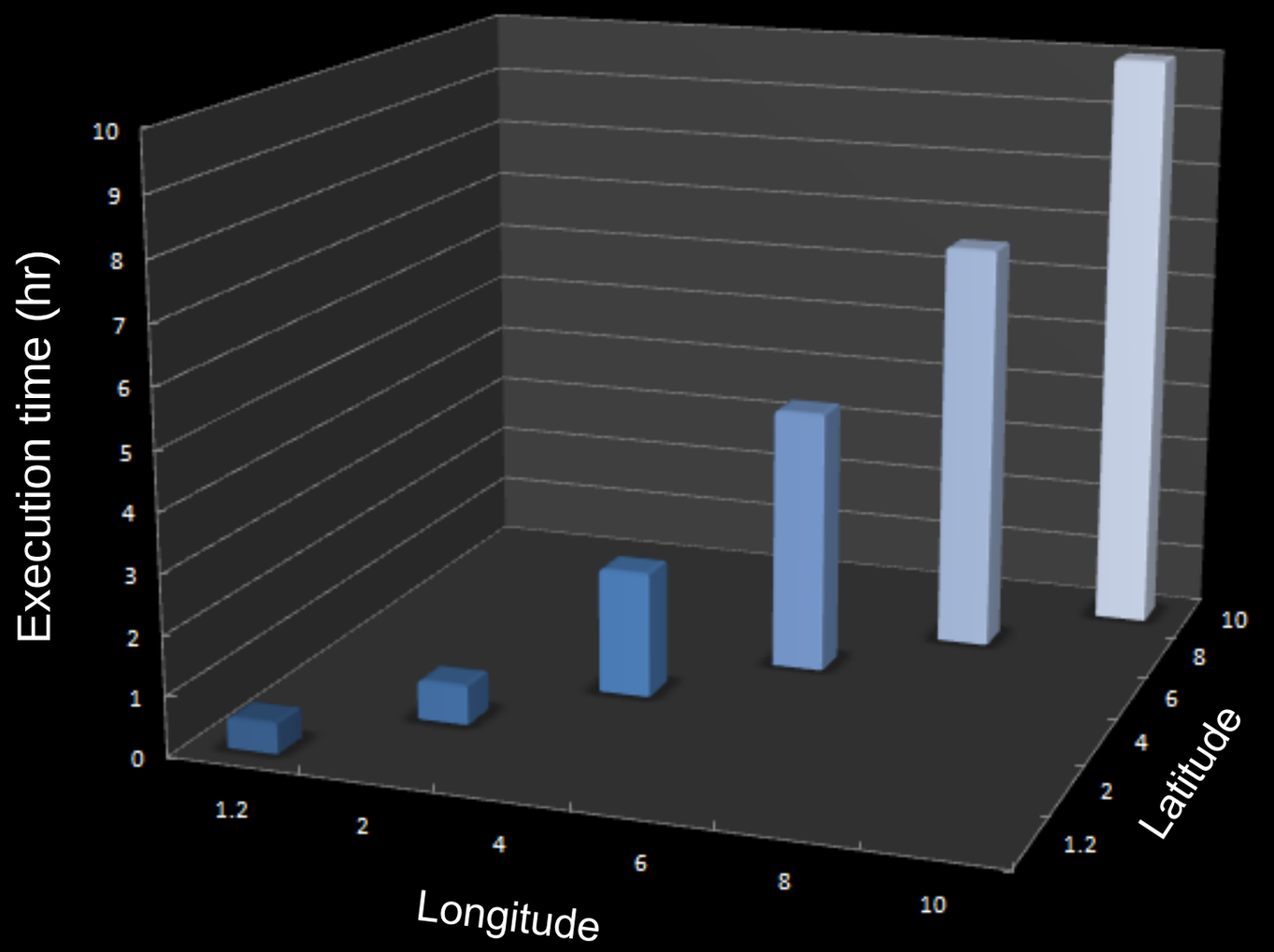


NMM-dust: 3km



Performance





Performance Comparison

Nested Running

- Retrieve initialization data from EDAC:1.5 mins
- Run ETA model: 20 mins
- Push ETA model output to EDAC: 4.5 mins
- AOI analysis: ~4.4 seconds
- Retrieve initialization data from EDAC for NMM : 3.9 mins
- Retrieve AOI data from EDAC for NMM: 30 s
- Execute NMM for each AOI
  - ✓ Depends on the AOI domain size
  - ✓ AOI 012: around 48 mins
- Push NMM outputs to EDAC: 1.5 mins per AOI output

NMM-Dust Only

- Retrieve initialization data from EDAC: 2.45 min
- Execute NMM
  - Depends on the size of domain
  - Cannot execute full domain
  - 10*10 degree : 12.7 hours with 8 CPU

Performance Comparison

		serial	parallel
retrieve ETA initialization data	0:01:32	0:01:32	0:01:32
run ETA model	0:20:00	0:20:00	0:20:00
push ETA output to data server	0:04:37	0:04:37	0:04:37
AOI Analysis	0:00:04	0:00:04	0:00:04
retrieve NMM initialization data	0:03:50	1:09:00	0:03:50
retrieve AOI data for NMM model run	0:00:30	0:09:00	0:00:30
execute NMM model for all AOIs	5:55:00	5:55:00	2:08:00
push NMM output to data server	0:01:30	0:27:00	0:01:30
		8:06:13	**2:40:03**

Feasibility Analysis

- ✓ Systems Integration
  - ✓ Model pre- and post-processor implementation
    - ✓ Relatively straightforward process
    - ✓ Challenge posed by models that require re-compilation to change model domain or other execution parameters
  - ✓ Data management and storage
    - ✓ Simple file system approach works well
    - ✓ Separation of model execution from file storage allows for optimization for modeling independent of storage capacity
    - ✓ Need to develop more structured data management system (i.e. data registry & management utilities) in move towards operationalization
- ✓ Appropriateness of implemented service standards
  - ✓ WCS is very effective in supporting data subsetting prior to network transfer (i.e. parameter extraction from large model products)
  - ✓ WMS remains useful for quick visualization of products in a variety of platforms, but for the modeling activity is not key
  - ✓ For this application, WFS seems like overkill, simple HTTP access to GeoJSON data files is sufficient for delivery of AOIs in a compact data model to remote systems
  - ✓ REST exchange remains useful for flexible transfer of data products between systems where the OGC services don't have a standard request-response model

Feasibility Analysis

- ✓ Performance
  - ✓ With all components now in place, performance testing is under way
  - ✓ Different geographic resolution will require different time to simulate.
    - ✓ One time of higher resolution (e.g., 4x4->2x2) will need 8-10 times more computing time
  - ✓ NMM model execution at high resolution (3km) remains computationally and time intensive, but smaller domains are more feasible
    - ✓ Different domain size will require different operation time.
    - ✓ Most sub-domain size is within 2x2 degrees, which can be processed within one hour for 3X3 km^2
- ✓ Given long execution times for NMM model, network latency for transfer of initialization parameters and outputs is a small fraction of total execution time.
  - ✓ About 1-2 minutes for transfer data between two sites (UNM & GMU)
- ✓ More dynamic parameters, such as soil moisture, should be assimilated into the model and implementation of this additional modeling capacity would require increasing computing power (potentially provided by cloud computing)

Follow-on/Related Opportunities

- ✓ Operationalization
  - ✓ Develop AOI generation service
  - ✓ Develop processing queue that tracks which AOIs have been processed for publication of un-executed AOIs as a feed (e.g. ATOM or RSS)
  - ✓ Enable time-enabled WMS for more efficient access and use
  - ✓ Auto-mosaic/overlay of time-enabled WMS for low- and high-resolution model outputs within a single service
- ✓ Integration with other modeling systems.
  - ✓ *Soil Moisture from Hydrologic Models for Model Initialization (NASA EPSCoR Proof of Concept Project)*
  - ✓ *Community Multi-Scale Air Quality (CMAQ) model (ENPHASyS Project)*
- ✓ Extension of on-demand high-resolution model execution into public/private cloud
- ✓ Automated air-quality alerts based upon AOI system

Contact Information

Karl Benedict

kbene@edac.unm.edu

(505) 277-3622 x 234

MSC01 1110, 1 University of New Mexico

Albuquerque, NM 87131

http://edac.unm.edu

Chowei Yang

cyang3@gmu.edu

GMU/CISC, MS 6A2

4400 Univ. Dr., Fairfax, VA, 22030-4444

http://www.cisc.gmu.edu/