Interstate Data Moving and the Last Block Problem: Lessons Learned in the CAPS Spring Experiment 2014

Keith A. Brewster, Ph.D.

Center for Analysis and Prediction of Storms

University of Oklahoma

SPC/NSSL Spring Program in the Hazardous Weather Testbed

- Testing and calibration of new forecasting methods in a simulated operational setting
- 6 weeks in spring season
- Collaboration among
  - NOAA research units
  - NOAA operational units
  - Universities
  - Private sector

- Testbed located between the NOAA Storm Prediction Center and Norman National Weather Service Forecast Office

CAPS Spring Experiment

- Part of NOAA/SPC Spring Experiment at the Hazardous Weather Testbed
- Run Large Ensemble of Convection-Allowing NWP Forecasts for 6-weeks in Spring
- New methods for severe weather prediction in 1-2 days time frame
- 25 NWP models run at XSEDE Centers
- 2013-2014 Darter at NICS (UTenn @ Oak Ridge)

Goal: "Real-time" 4D Data Visualization

Procedure Since 2007

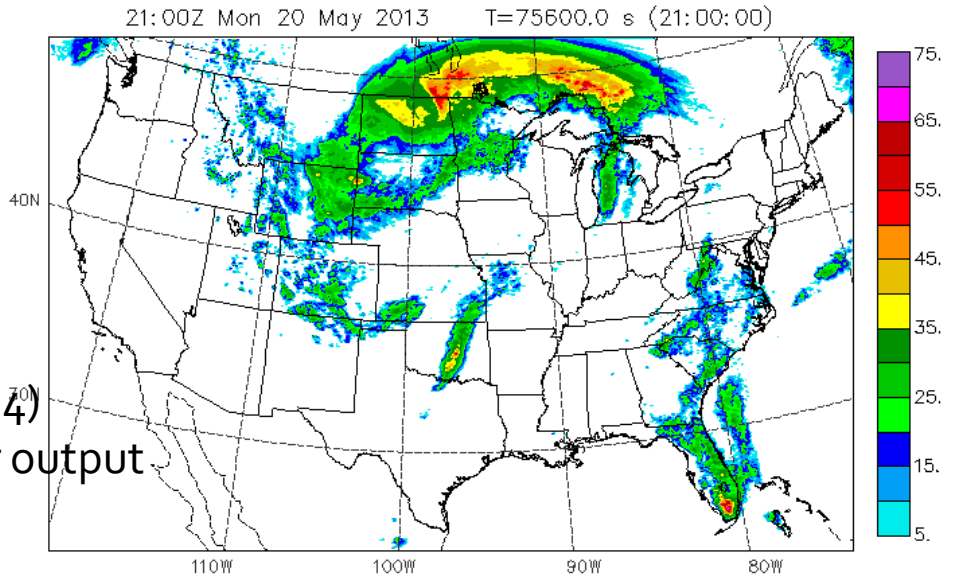
- Run models at PSC or NICS
- Bring 2D files and images back
- 2D fields and levels pre-selected

Issues

- Does not allow full 3D visualization
- May want to examine other fields and levels

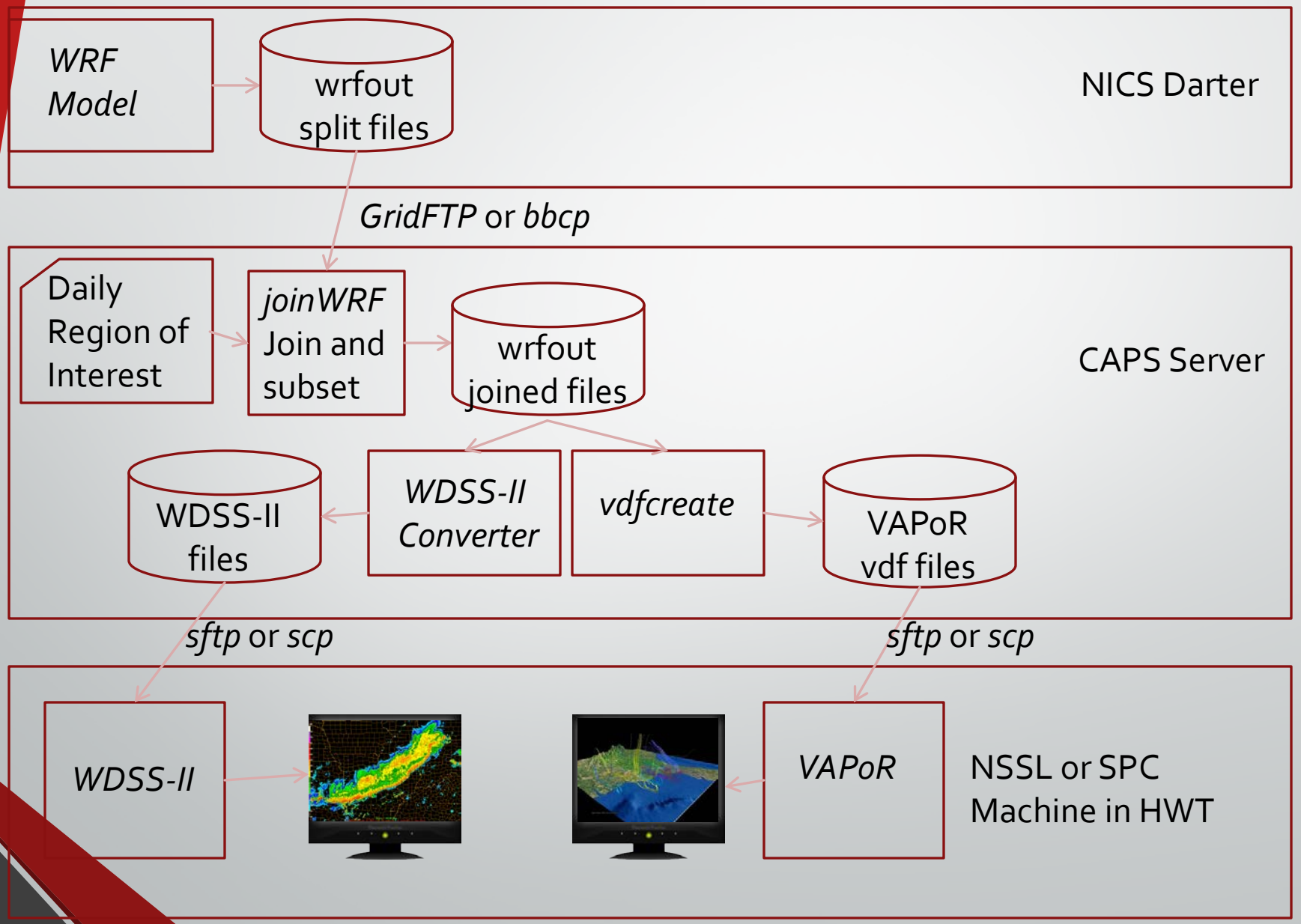
Need to move 3D Data from NICS at UTenn CAPS at OU

Scoping the Task



- CONUS Domain at 4-km Resolution
  1163 x 723 x 53
  Output for one time: **4.2 GB**
- Domain decomposition onto 384 (6x64) processors results in **384** split files per output time, each file = **11** MB
- For smooth animations, 10-minute output is generated for 5 members covering the afternoon and evening, forecast hours 18-30.
- Complete Forecast:
  - 60-h forecast, hourly output + 10 minute output 18h-30h:
    121 output times, **508 GB**
- Day-1 Afternoon and Evening for Animation
  - Forecast 18h-30h with 10-minute output:
    73 output times, **307 GB**

Plan "A" Workflow



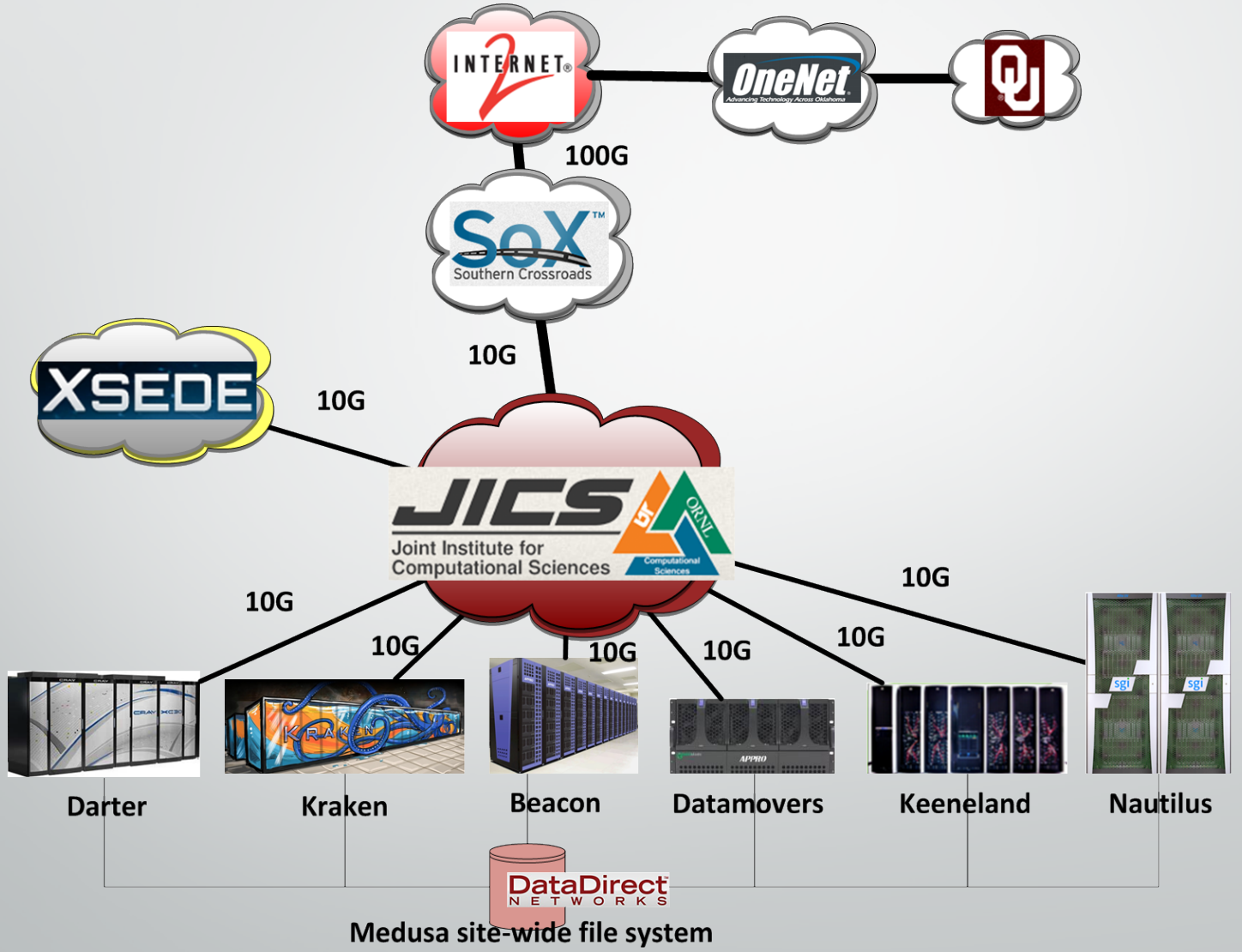
Data Route

NICS/U Tenn in Oak Ridge, Tennessee to CAPS/OU in Norman, Oklahoma

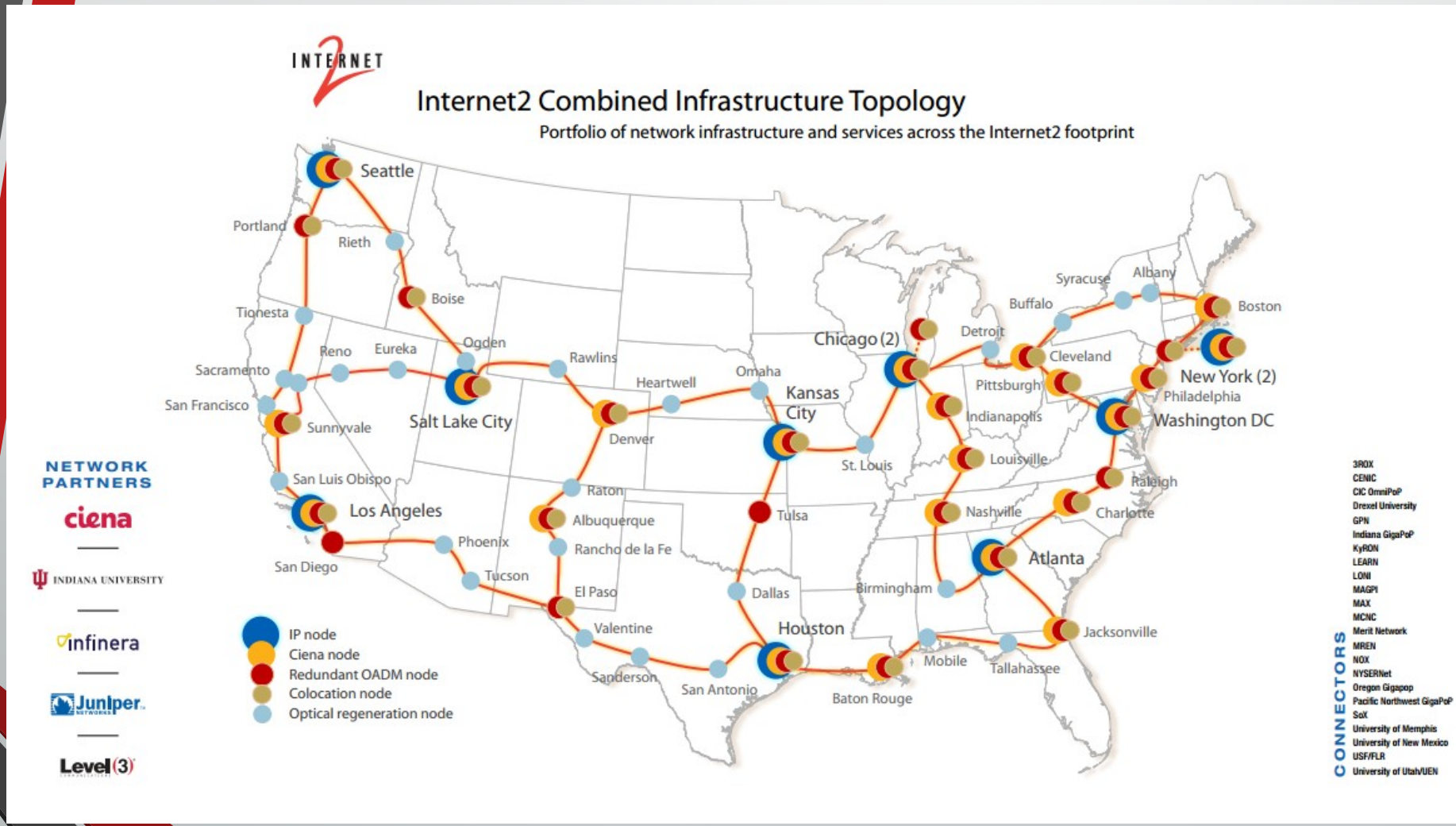
845 miles 13 hours via MichelinNet

Map from mapquest.com

Internet Route



Internet2


INTERNET 2
Internet2 Combined Infrastructure Topology
Portfolio of network infrastructure and services across the Internet2 footprint
Seattle
Portland
Rieth
Boise
Tionesta
Reno
Eureka
Ogden
Sacramento
Rawlins
Heartwell
Omaha
San Francisco
Sunnyvale
Salt Lake City
Denver
Kansas City
Chicago (2)
Detroit
Syracuse
Albany
Buffalo
Boston
Cleveland
New York (2)
Pittsburgh
Philadelphia
Indianapolis
Washington DC
St. Louis
Louisville
Raleigh
San Luis Obispo
Raton
Los Angeles
Albuquerque
Tulsa
Nashville
Charlotte
Phoenix
Rancho de la Fe
Atlanta
San Diego
Tucson
El Paso
Dallas
Birmingham
Valentine
Houston
Jacksonville
Sanderson
San Antonio
Mobile
Tallahassee
Baton Rouge
IP node
Ciena node
Redundant OADM node
Colocation node
Optical regeneration node
NETWORK PARTNERS
ciena
INDIANA UNIVERSITY
infinera
Juniper NETWORKS
Level(3)
CONNECTORS
3ROX
CENIC
CIC OmniPoP
Drexel University
GPN
Indiana GigaPoP
KyRON
LEARN
LONI
MAGPI
MAX
MCNC
Merit Network
MREN
NOX
NYSERNet
Oregon Gigapop
Pacific Northwest GigaPoP
SoX
University of Memphis
University of New Mexico
USF/FLR
University of Utah/UEN

Internet Route "Last Mile"

- OneNet Tulsa to OU – Norman (4PP)
- 4PP to National Weather Center
  - Across the parking lot
- National Weather Center to
  - CAPS Switch to File Server System
  - to CAPS Office Workstation via Firewall

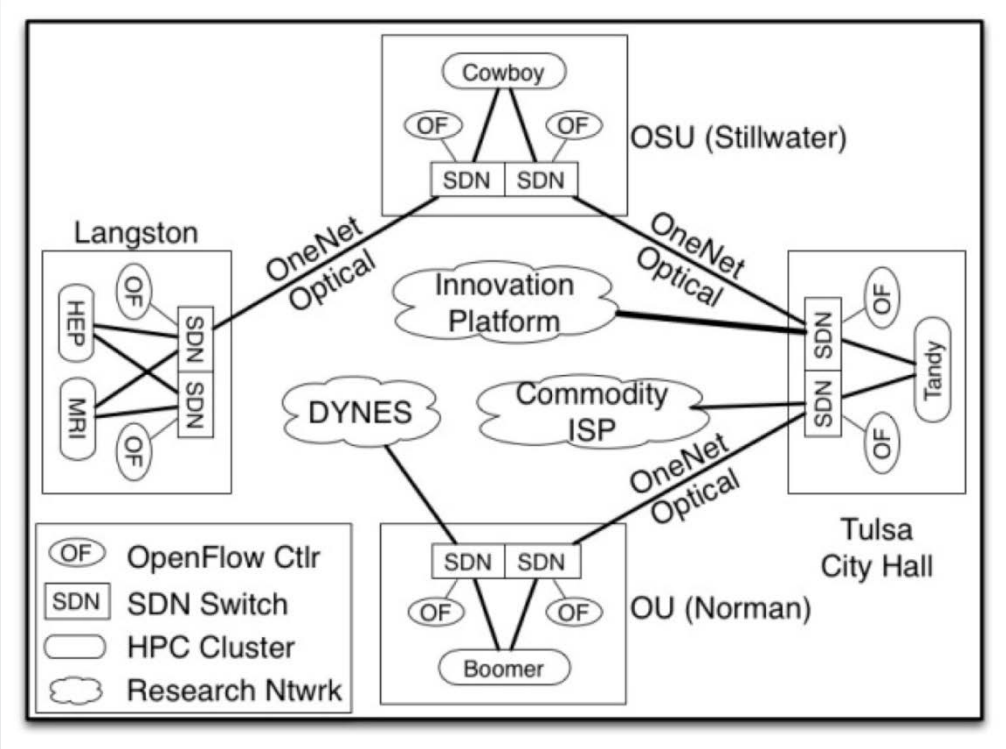
Recent Networking Initiatives

1. University of Tennessee BLAST
2. OneOklahoma Friction Free Network (OFFN)
3. National Weather Center Upgrade

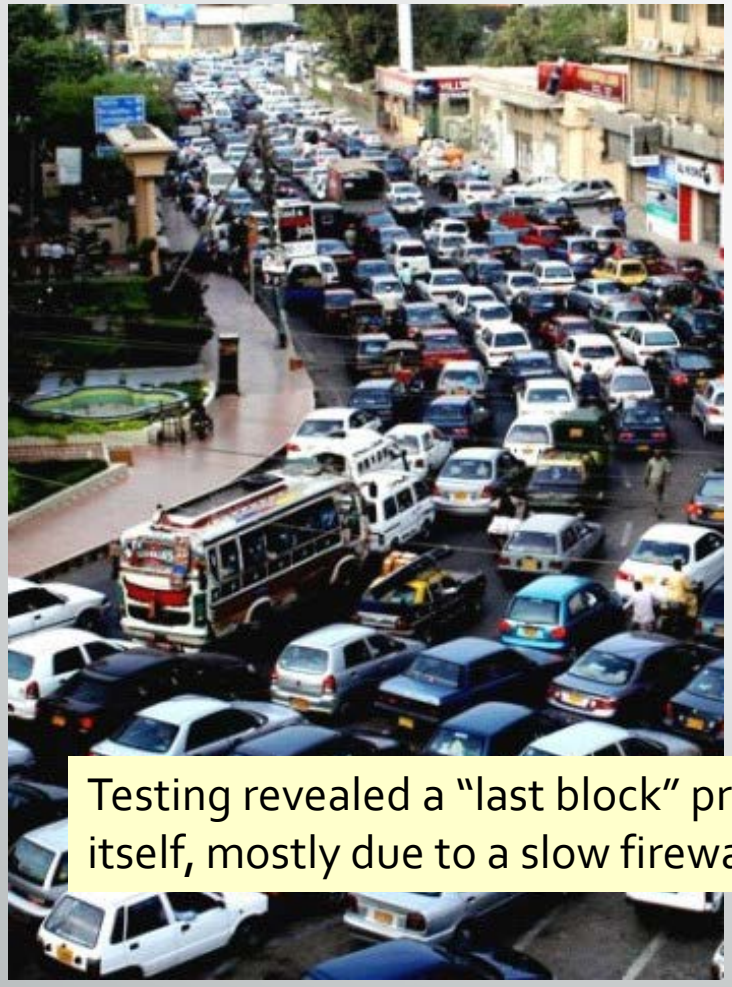
- University of Tennessee BLAST
  - 100 Gps Upgrade of Research Network
  - Includes connections to HPC at NICS

Recent Networking Initiatives

- OneOklahoma Friction Free Network (OFFN) NSF Campus Cyberinfrastructure-Network Infrastructure and Engineering Program (CC-NIE)
  - Establish 10 Gbps Network Ring
  - OU-OSU-Langston-Tandy Supercomputing Ctr



Traffic & Last Block Problem



Testing revealed a "last block" problem, actually within the building itself, mostly due to a slow firewall.

Packing and Compression?

Try creating compressed tar file before sending?
Sending large files is faster ~100 MB/s vs ~10 MB/s

BUT! Creating a compressed tar file takes time

Test for 1-hour of Full Domain Split Files

Operation	Time
bbcp individual split files	50 min

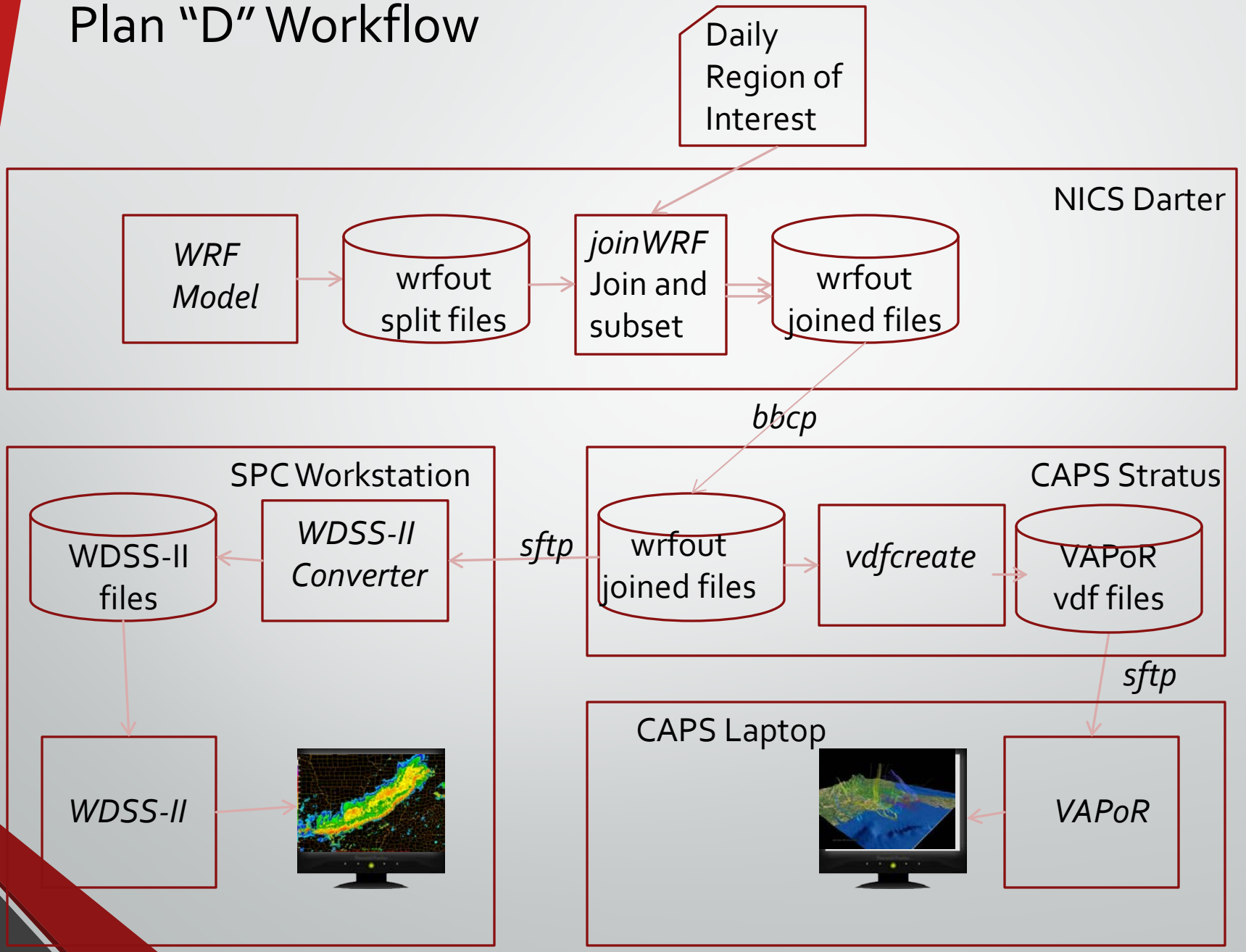
With Compression

Operation	Time
$ tar –zcvf	22 min
bbcp	2.5 min
Total	24.5 min

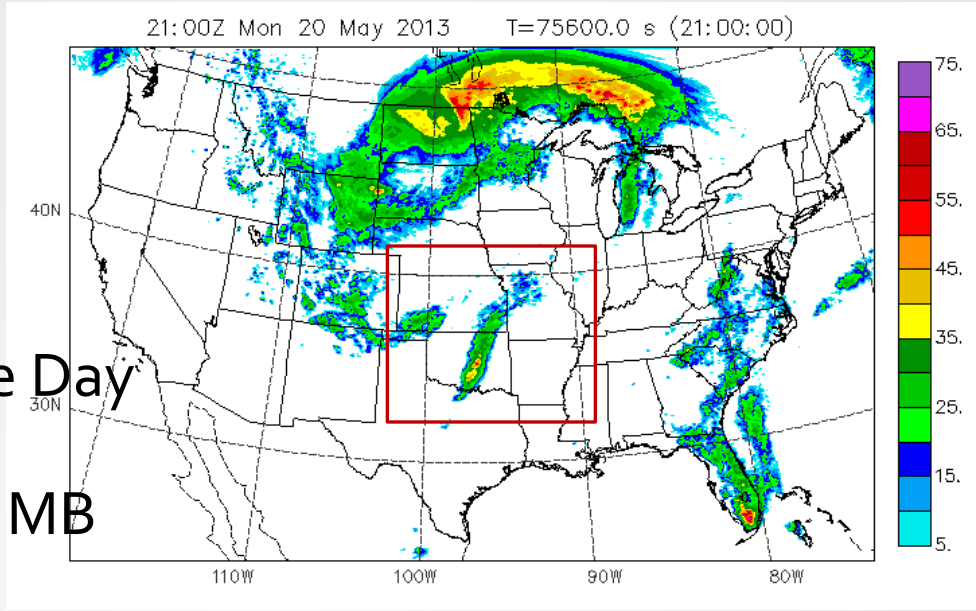
Without Compression

Operation	Time
$ tar -cvf	15 min
bbcp	5 min
Total	20 min

Plan "D" Workflow



Re-Scoping the Task

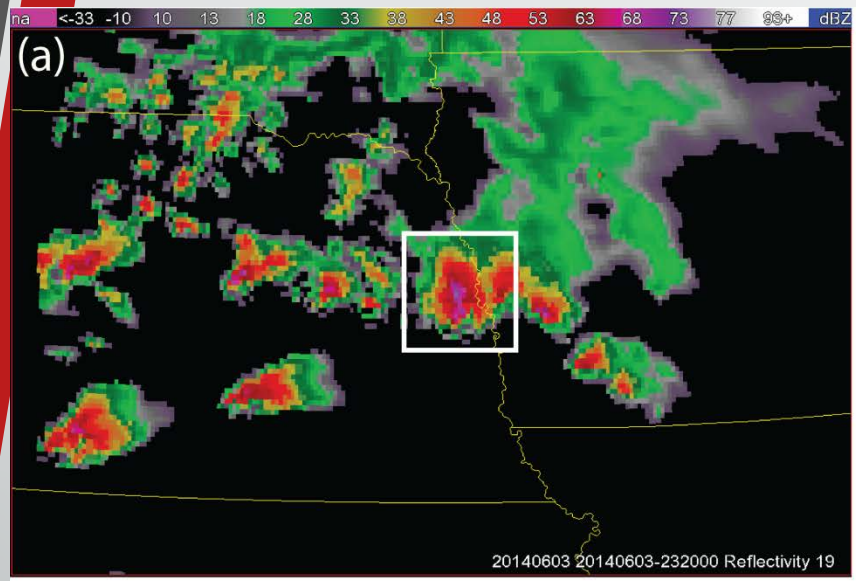


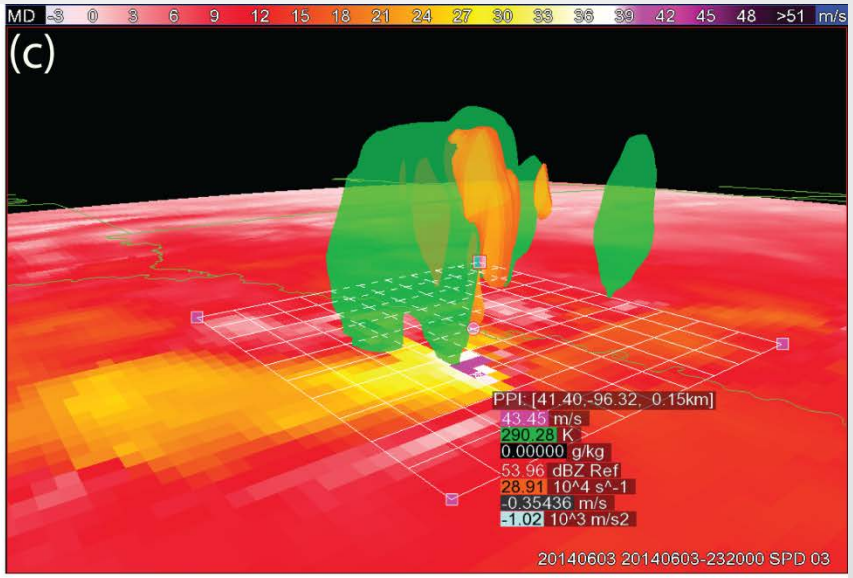
- Selected Subdomain of the Day
  203 x 203 x 53
  Output for one time: 111.5 MB
- Single file for each time =
  111.5 MB vs 4.2 GB

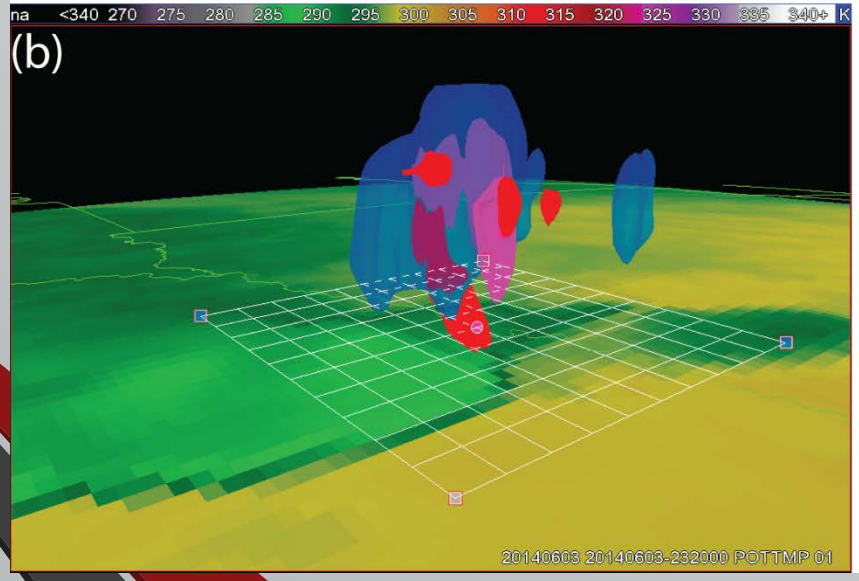
- Day-1 Afternoon and Evening for Animation
  - Forecast 18h-30h with 10-minute output:
    73 output times,
    8.2 GB vs **307 GB**
- Processing and Transfer: **~20 min**

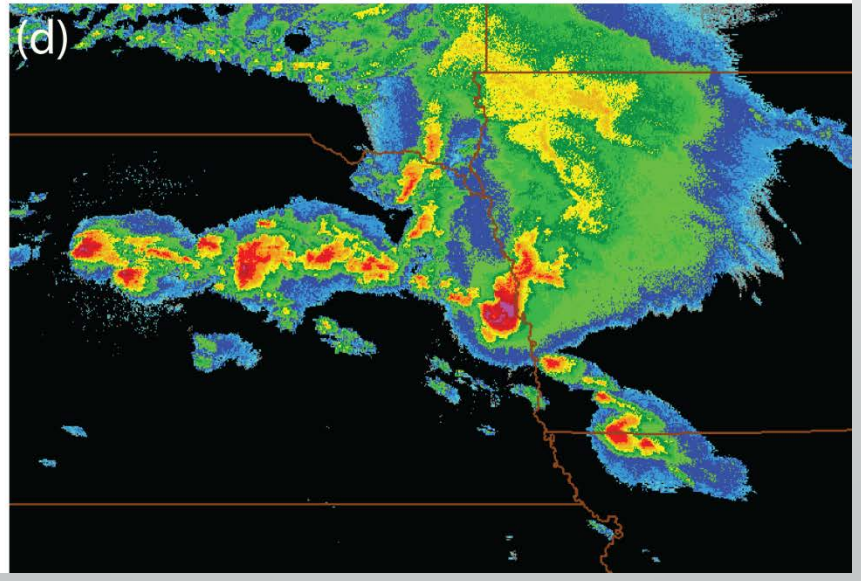
Fruits of Labor

WDSS-II



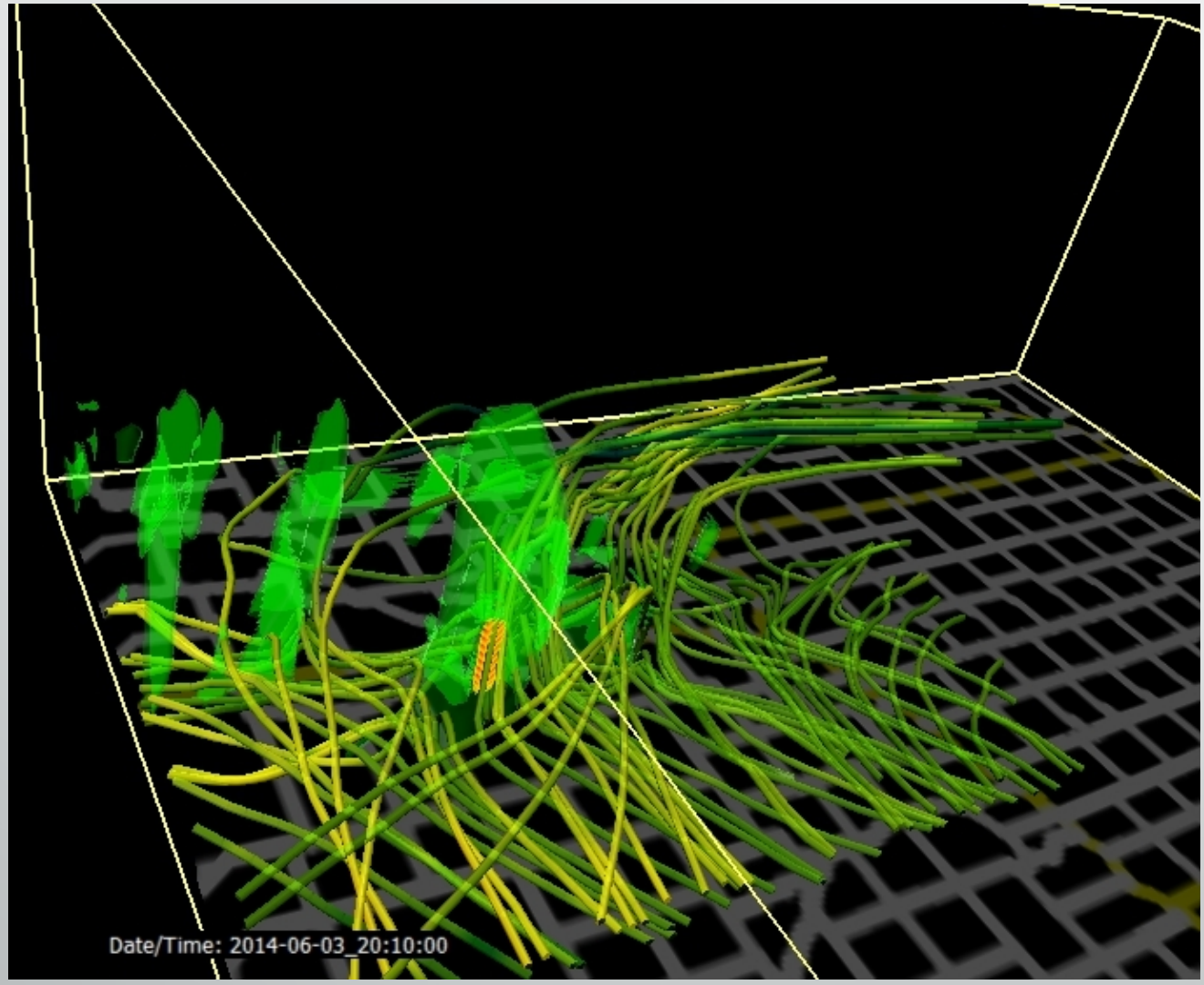






Fruits of Labor

VAPoR



Lessons Learned

- Involve networking pros early
- Your mileage (throughput) may vary
- Be flexible with workflows
- Evaluate overhead of all steps
- Find ways to fund equipment upgrades where needed – slowest link sets your rate
- Software Programmable Networks/Science DMZ's may be needed for the largest jobs

Future Plans

- National Weather Center (September, 2014)
  - Upgrading Network to 10 Gps Switches
  - Software Programmable Networking Enabled
  - Replacing Slow Firewall
- CAPS (September-October, 2014)
  - Upgrading to Two 10 Gps Gateway Servers with Virtual Router Redundancy Protocol (VRRP)
  - Software Programmable Networking Enabled
  - Upgrading servers to 10 Gps network interfaces
- For 2015
  - Set Jumbo Size Maximum Transmission Unit (MTU) for route
  - Explore use of Science DMZ within OneNet

Questions?

Thanks to:
Chris Cook, CAPS
Kevin W. Thomas, CAPS
Victor Hazlewood and his networking team, UTenn
Matt Runion & James Deaton, OneNet
Henry Neeman & OSCER Team
Mike Coniglio, NOAA/NSSL

Contact Info:
Keith Brewster
CAPS/University of Oklahoma
kbrewster@ou.edu