Welcome to the Year-2 NSF Site Visit

University Corporation for Atmospheric Research

Colorado State University

Howard University

Indiana University
Quality Education. Lifetime Opportunities.®

LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY

UNC Chapel Hill

UAH The University of Alabama in Huntsville

University of Illinois at Urbana-Champaign

Millersville University 'Ville

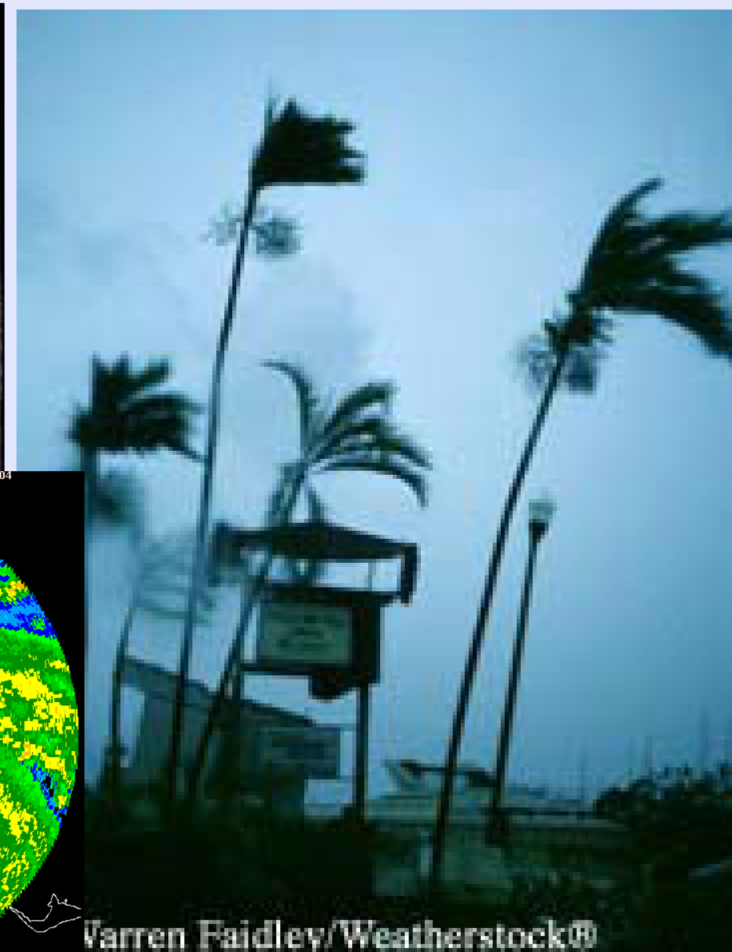
the university of Oklahoma

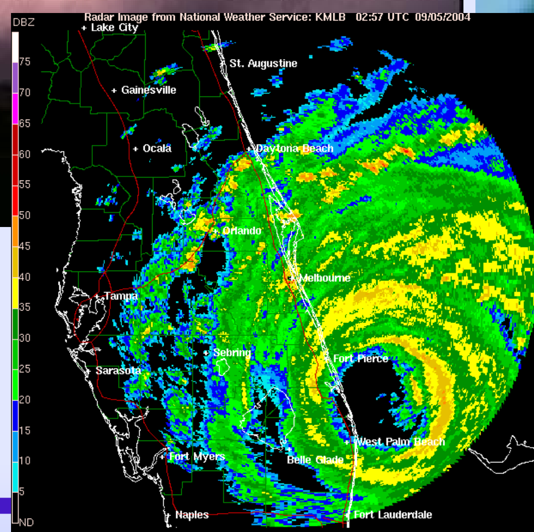
Overview of LEAD

Kelvin Droegemeier

Year-2 Site Visit
21-22 July 2005

What Would <u>YOU</u> Do if These Were About to Occur?

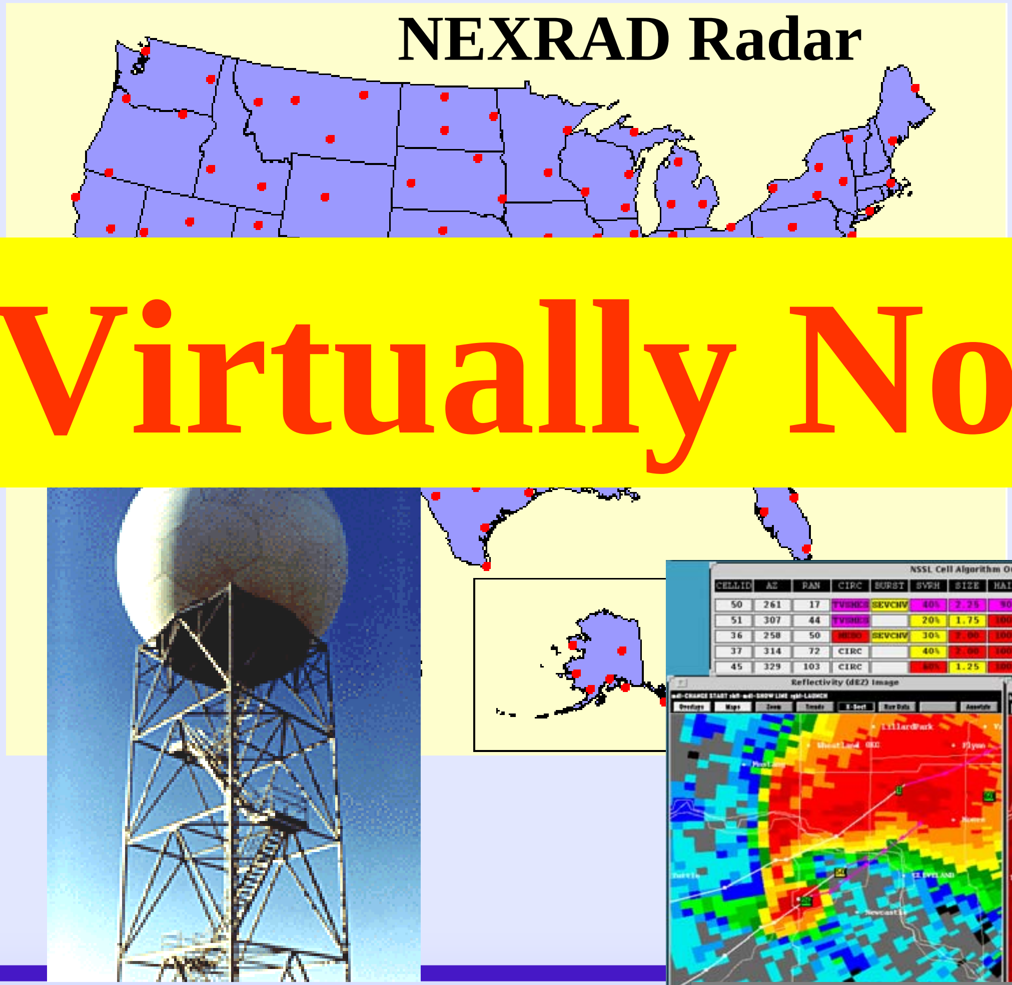




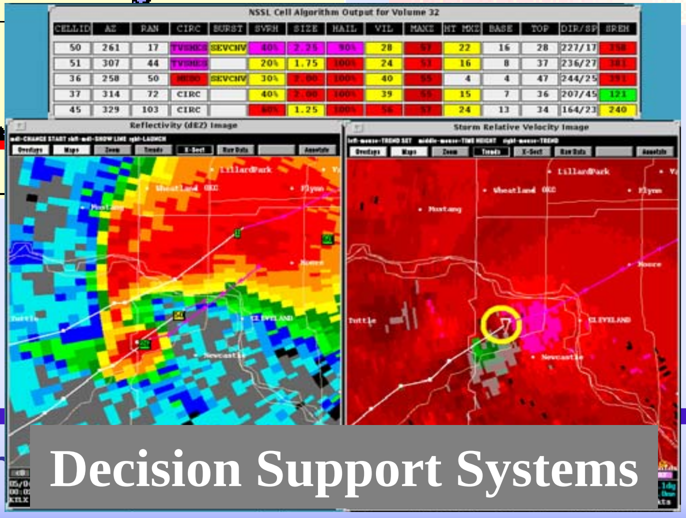
What THEY Do to Us!!!

- Each year, **<u>mesoscale weather</u>** – local floods, tornadoes, hail, strong winds, lightning, and winter storms – causes hundreds of deaths, routinely disrupts transportation and commerce, and results in annual economic losses > $13B.

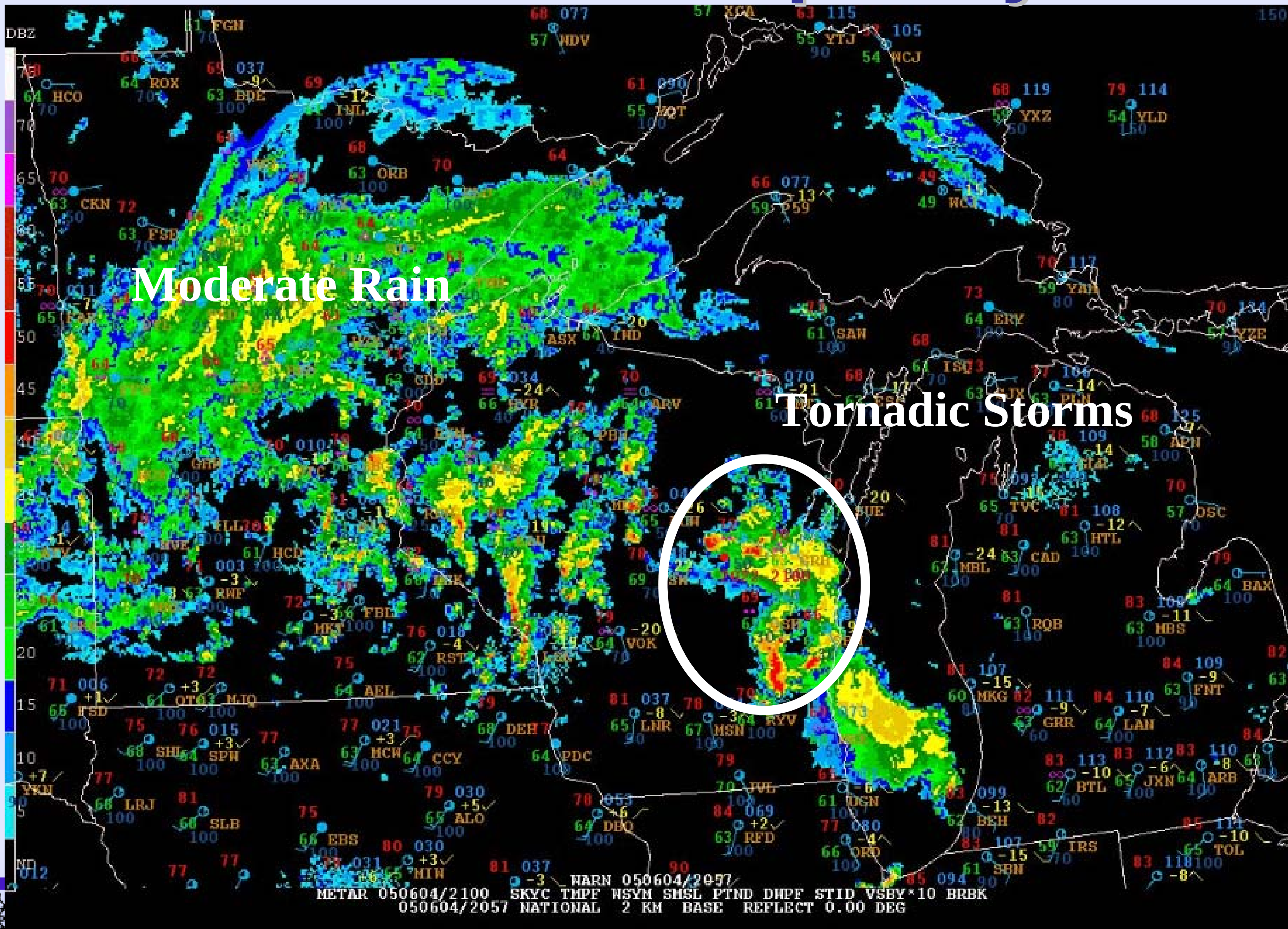
What Weather & Associated Information Technologies Do…



Virtually Nothing!!!

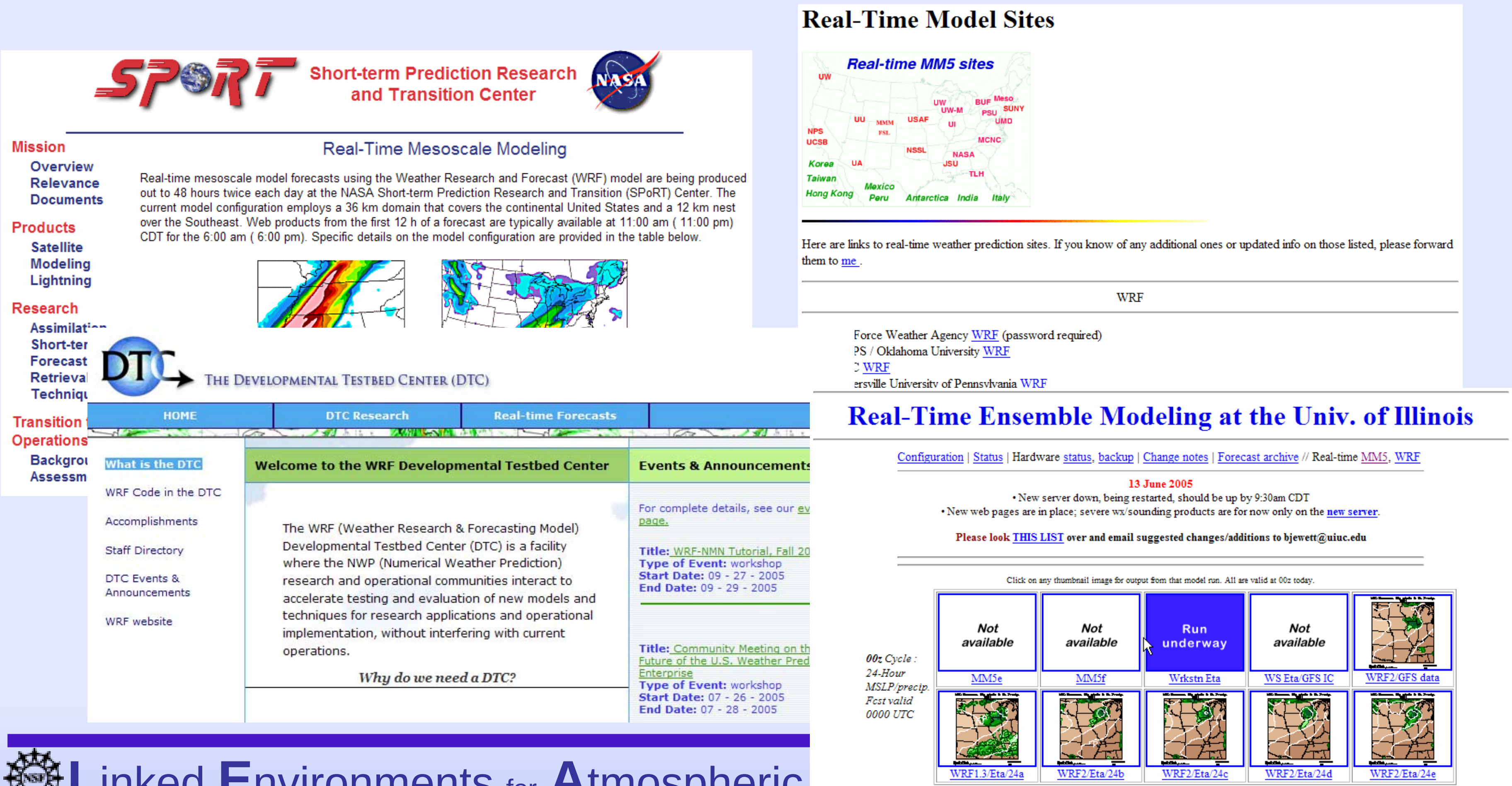


Radars Do Not Adaptively Scan

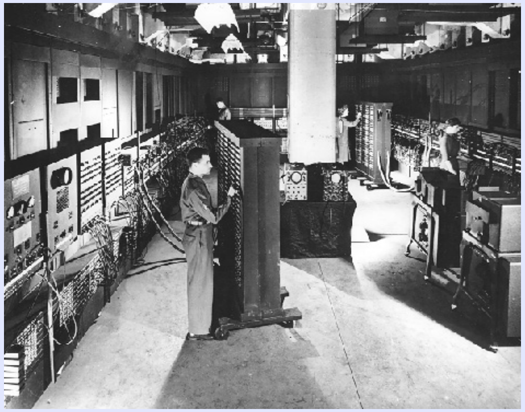


(lead.ou.edu)

Real Time Research Models Use Very Simple Static Configurations and Most Don't Even Assimilate Observations

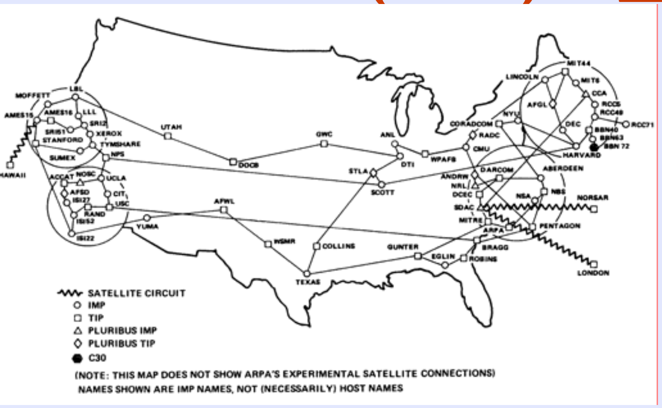


Cyberinfrastructure is Virtually Static



ENIAC (1948)

Earth Simulator (2005)



ARPANET (1980)

National Lambda Rail (2005)

Abilene Backbone (2005)

The Bottom Line

- Mesoscale weather is VERY DYNAMIC and impactive but our tools, cyber environments, research methodologies and learning modalities are VERY STATIC
- Getting even static capability is an enormous challenge due to the complexity of the tools and the primitive inf… ctures used to link

The LEAD Vision

Conduct the research and development necessary to allow

People (scientists, students, operational practitioners)

and

Technologies (models, sensors, data mining)

TO INTERACT WITH MESOSCALE WEATHER TO ADVANCE KNOWLEDGE & UNDERSTANDING

Two Principal Goals of LEAD

- **Goal #1**: Dynamic Adaptation to Weather
  - Models and hazardous weather detection systems responding to observations and their own output
  - Models and hazardous weather detection systems driving the collection of observations
  - IT infrastructures providing on-demand, fault tolerant services
- **Goal #2**: Lowering the barrier for using complex end-to-end weather technologies
  - Democratize the availability of advanced weather technologies for research and education
  - Empower application in a grid context
  - Facilitate rapid understanding, experiment design and execution

The Potential Payoff

Posters #2, #4

- **An improved understanding** of mesoscale weather by studying it in ways that are consistent with its behavior → improved forecasting

Posters #15-18, 22-23

- Advanced meteorological capabilities available to a much **broader community** of users

Poster #24

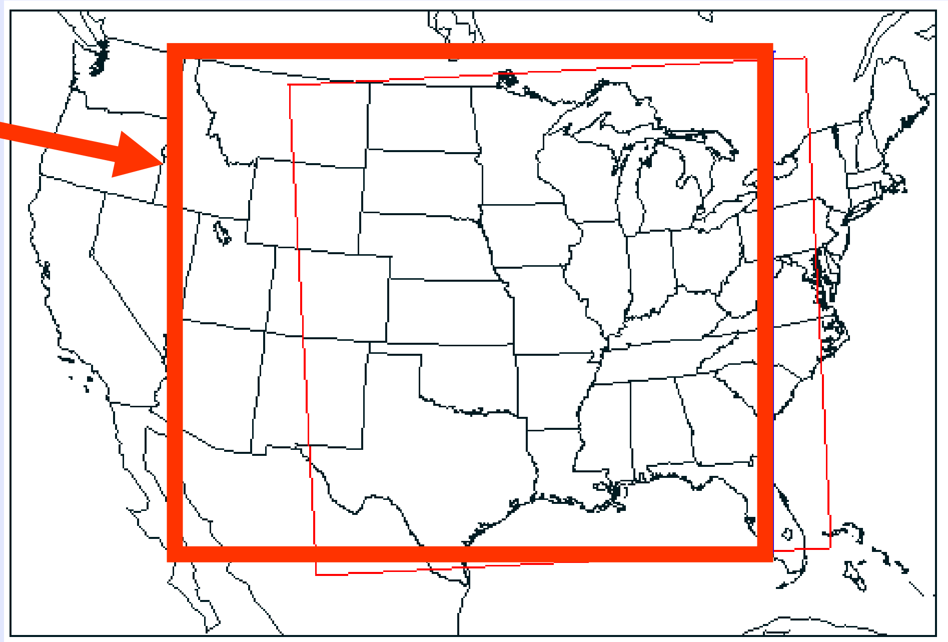
- Much **shallower learning curve** -- design and execute experiments in minutes rather than months.
- Application of weather technologies in a **grid** context
- An environment for **ongoing basic research** in computer science (e.g., data, workflow, monitoring, QoS)
- An environment where meteorologists can learn about new computing concepts

Setting the Context: Today's Capabilities

- Sophisticated mesoscale numerical models, systems and other tools abound (ARPS, MM5, WRF, COAMPS)
- All components are very complex even if used individually
- Process control infrastructures are not widely available (e.g., CAPS) and are unwieldy (50K line Perl scripts)
- The NET RESULT
  - Huge learning curve, especially for students (70-30 rule)
  - Limited sophistication of experiments
  - Cannot run in grid environments
  - Disincentive for use – requires substantial human capital and physical infrastructure

Real Time Forecasting: Simple Configurations + Limited to a Few Groups

- CAPS made one 30-hour WRF forecast per day at 2 km horizontal grid spacing
- Fixed model configuration
- Initialized with NCEP analysis – no radar or other data added
- Dedicated time on Lemieux at PSC
  - 1228/3000 processors from midnight – 7 am each day



Preparation took 3 top scientists and 2 technical support staff 7 months + help from NCAR & PSC Staff

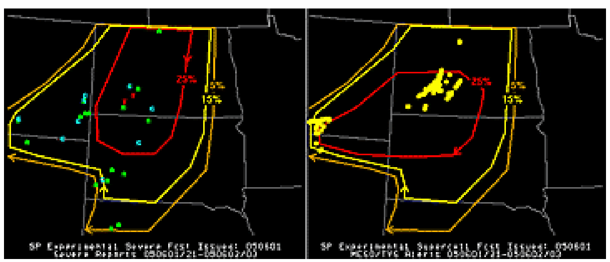
Sample Result

Poster #14

Improved Storm Forecast Capability Demonstrated

In a multi-partner spring program led by NOAA, the Center for Analysis and Prediction of Storms and the Pittsburgh Supercomputing Center generated the highest-resolution numerical weather forecasts yet attempted, with results suggesting that storms may be more predictable than previously thought.

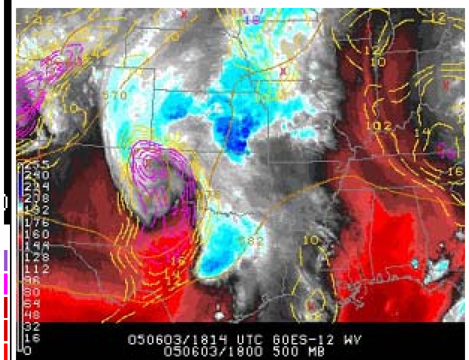
PITTSBURGH, July 5, 2005 — As it has during many storm seasons over the past decade, the Pittsburgh Supercomputing Center (PSC) this spring collaborated with the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma, Norman, to produce real-time numerical forecasts of storms.



This year, however, rather than forecasting for a small region over the Great Plains, CAPS harnessed PSC resources to generate forecasts over two-thirds of the continental United States. "This was an unprecedented experiment," said CAPS director Kelvin Droegemeier, "that meteorologists could only dream of several years ago."

Relying on LeMieux, a leading computing resource of the National Science Foundation's TeraGrid, the CAPS team produced the highest resolution storm forecasts that have yet been attempted. CAPS used the Weather Research and Forecasting Model (WRF), an advanced model designed for research as well as operational use, and with LeMieux - running on 307 nodes (1,228 processors) - successfully produced an on-time, daily forecast from mid-April through early June.

NOAA SCIENTISTS WORK TO IMPROVE SEVERE WEATHER FORECASTS

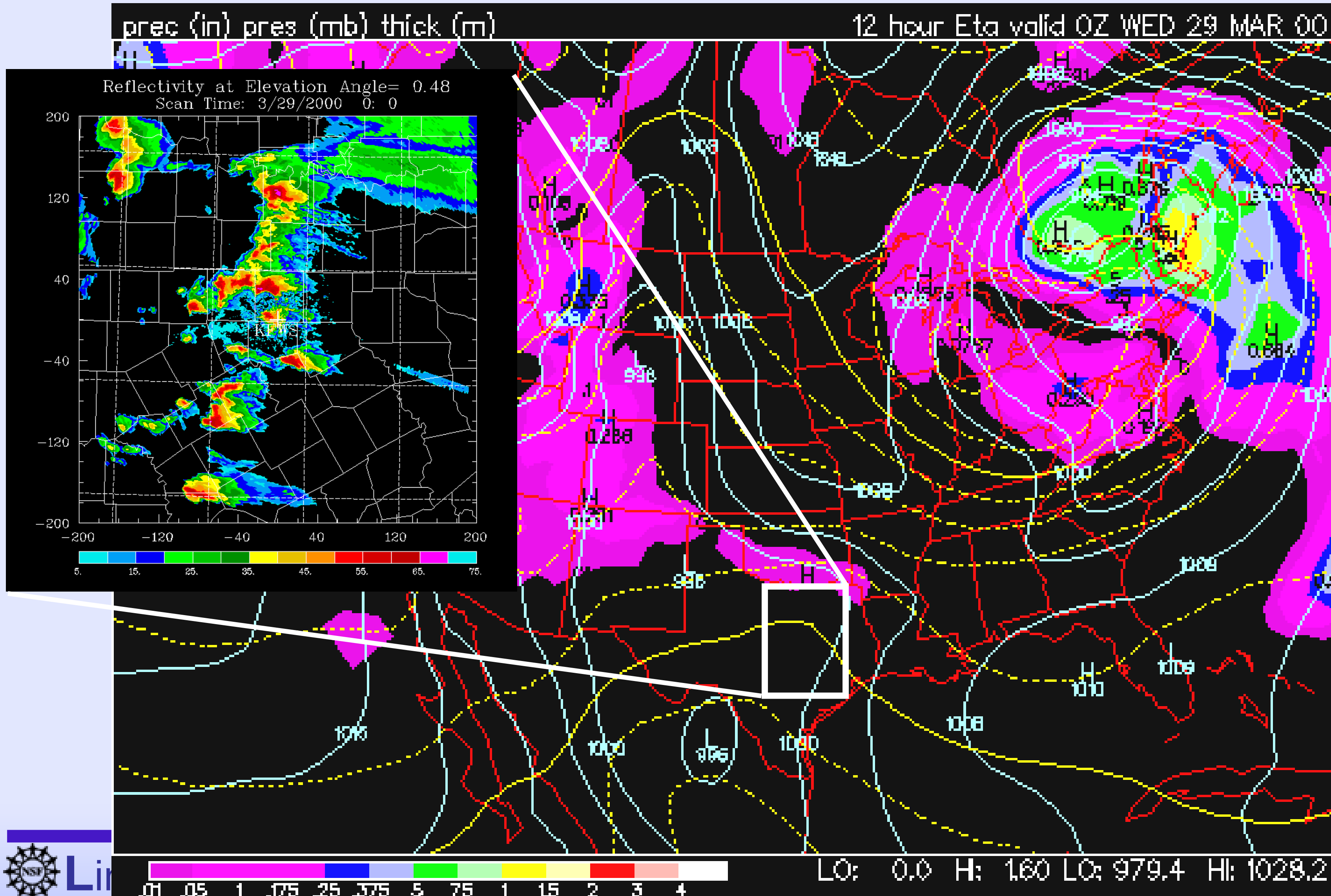


June 20, 2005 — Thunderstorms with lightning, hail, strong winds and tornadoes can be devastating, resulting in hundreds of deaths and millions of dollars in damage each year. Researchers and forecasters with NOAA in Norman, Okla., are working together to improve the tools forecasters use to predict such storms, ultimately providing the public more time to prepare for severe thunderstorm events and more specific information about what type of severe weather to expect. **(Click NOAA image for larger view of radar and satellite composite of storm crossing the middle of the United States on June 3, 2005. Please credit "NOAA.")**

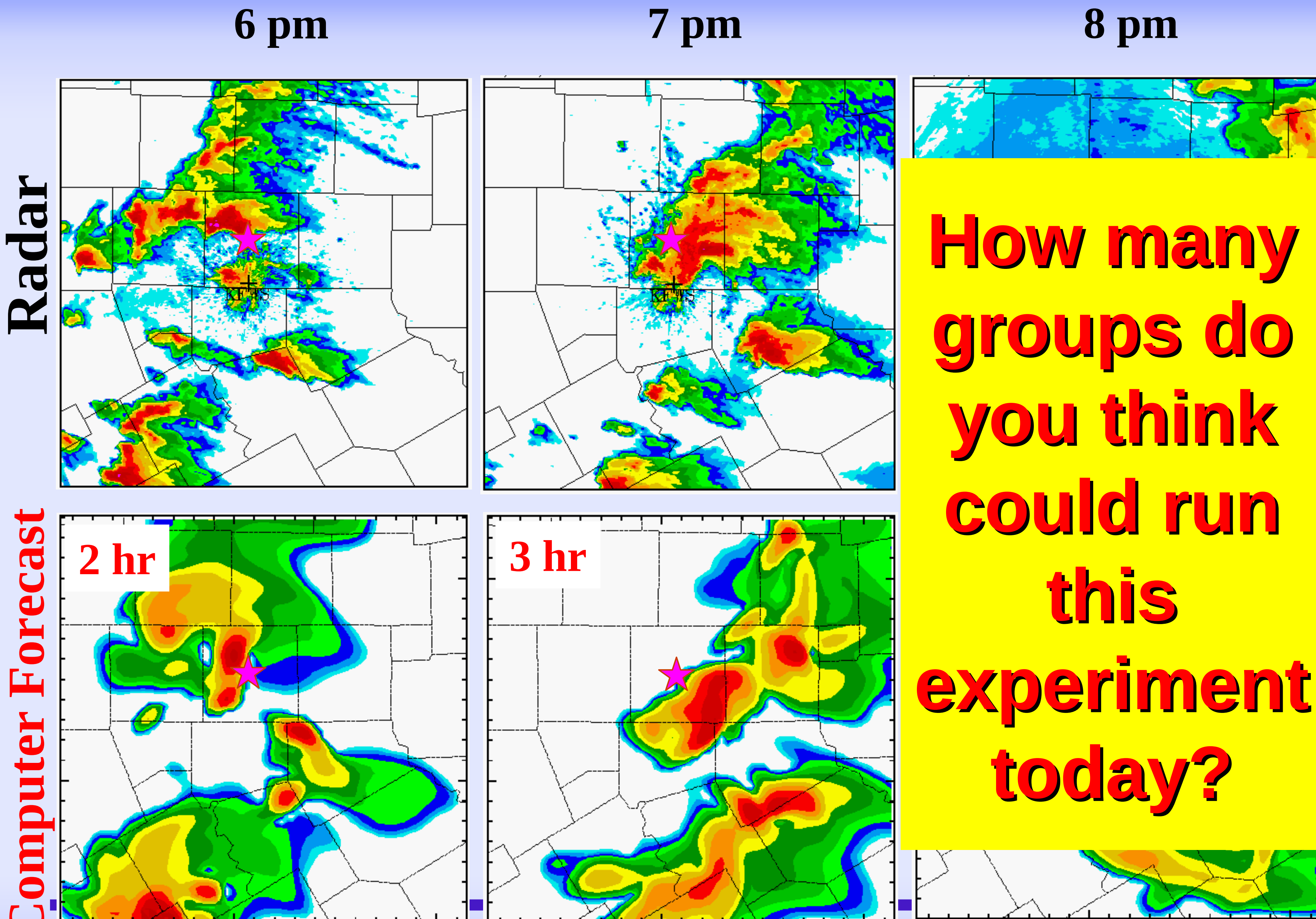
This one-of-a-kind collaboration between the research and forecast communities occurs each year at NOAA in what is known as the Spring Program, and is the cornerstone of the NOAA Hazardous Weather Testbed, operated jointly by the NOAA's Storm Prediction Center and the National Severe Storms Laboratory. It provides a unique environment where research meteorologists can interact directly with the end users of their products—operational forecasters—with obvious benefits for all involved.

The NOAA Storm Prediction Center and the NOAA National Severe Storms Laboratory worked closely with the NOAA National Weather Service office in Norman and partnered with three external organizations to generate a unique collection of high resolution numerical weather prediction models. These experimental models were generated three times a day. The predictions were made from several different versions of the Weather Research and Forecasting (WRF) model, an advanced weather prediction system being designed for use by research scientists and forecasters in the United States.

A Simpler Problem: A Non-Real Time Real Data Forecast

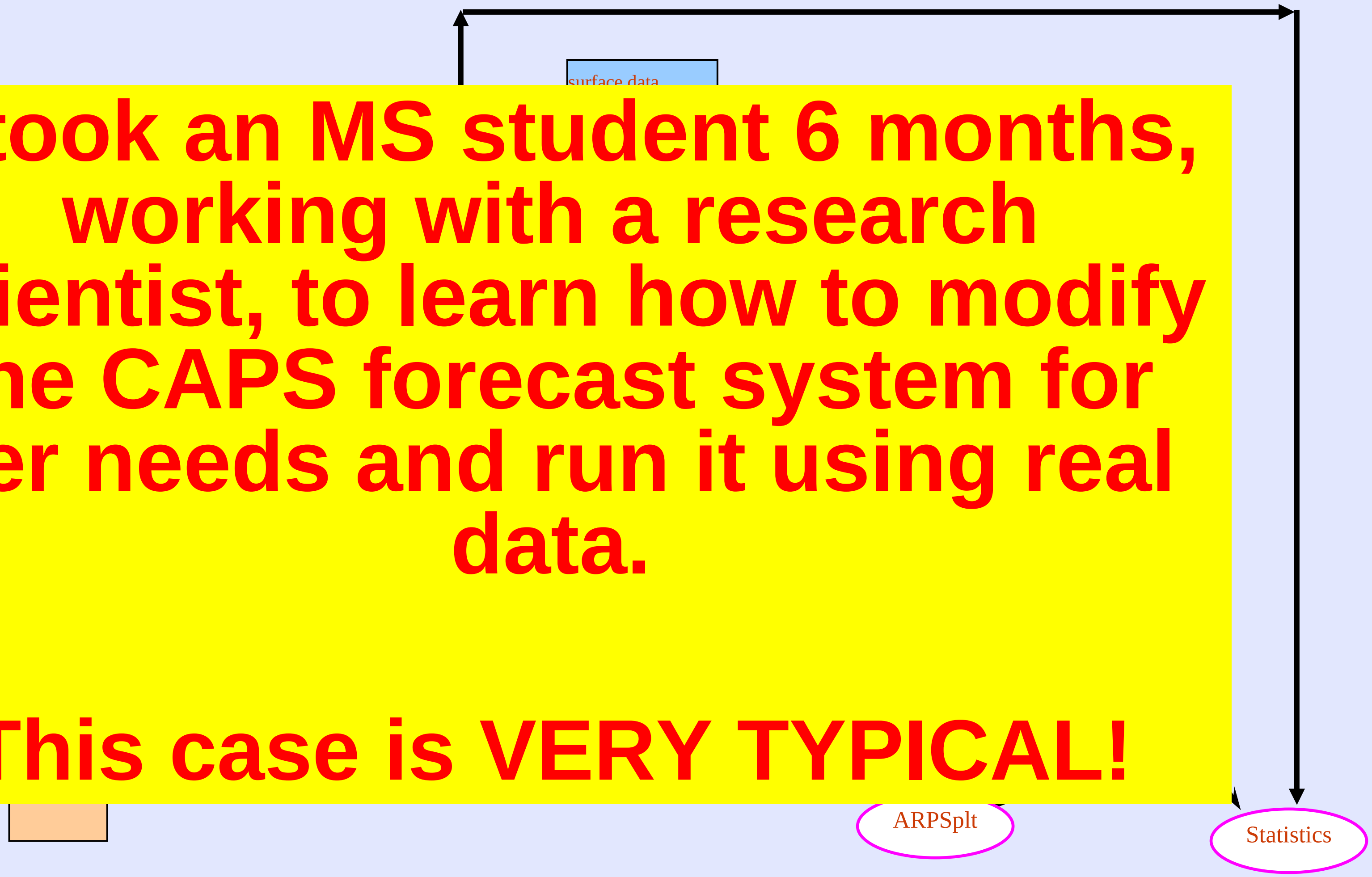


(lead.ou.edu)



How many groups do you think could run this experiment today?

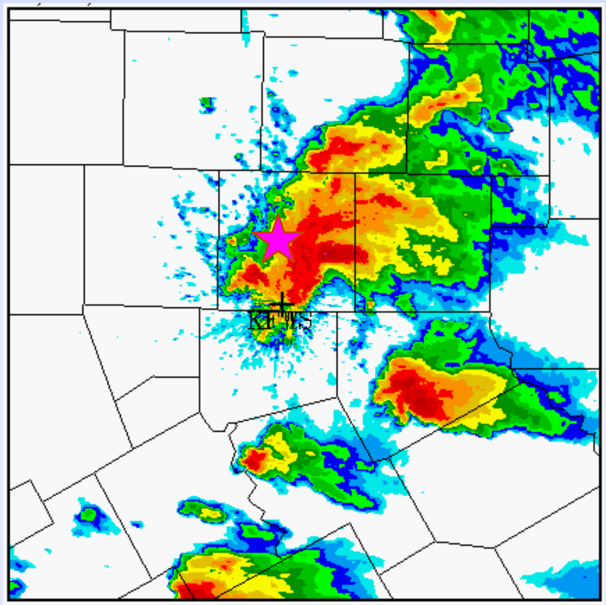
Experiment Components & Flow

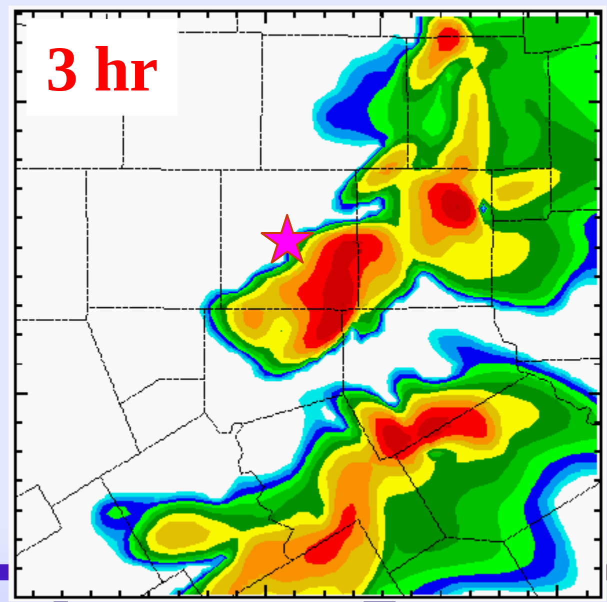


It took an MS student 6 months, working with a research scientist, to learn how to modify the CAPS forecast system for her needs and run it using real data.

This case is VERY TYPICAL!

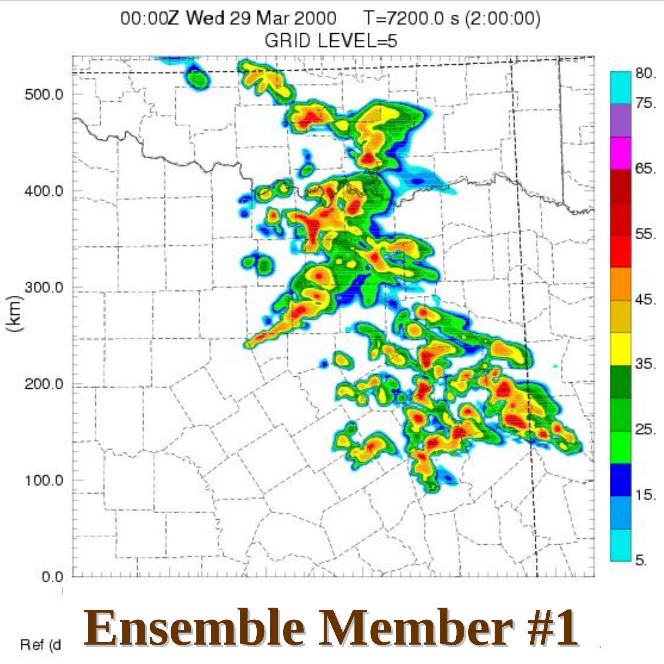
7 pm

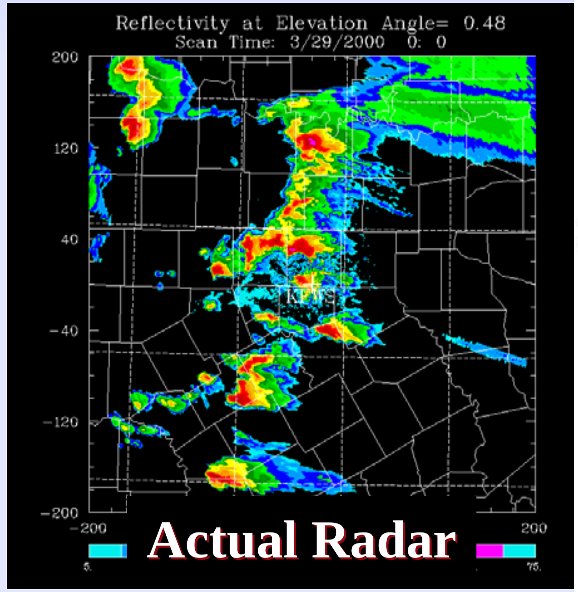


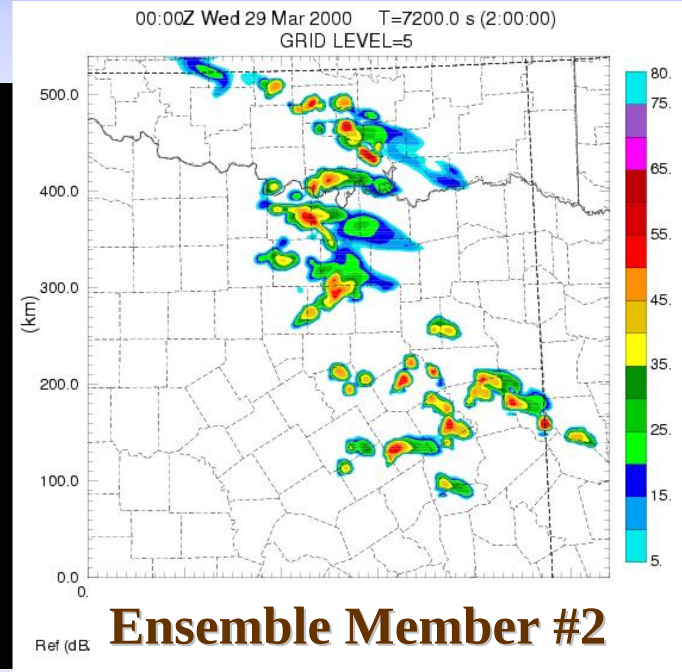


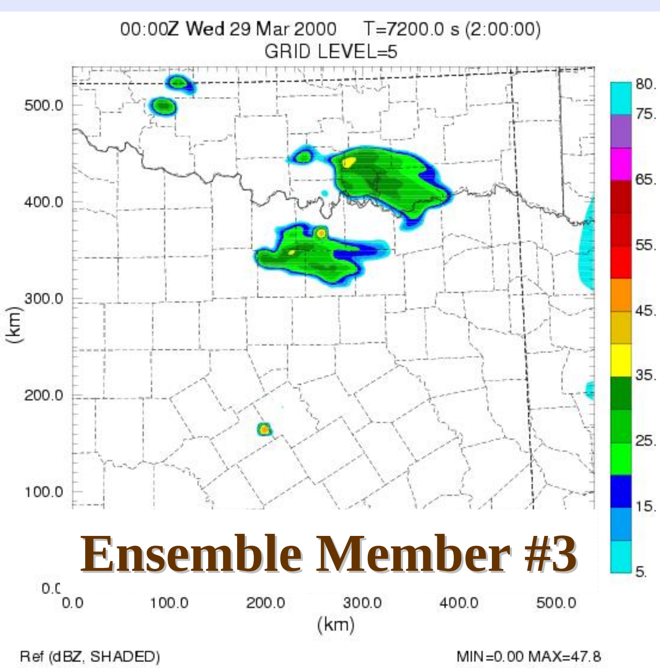
How much do you trust this forecast?

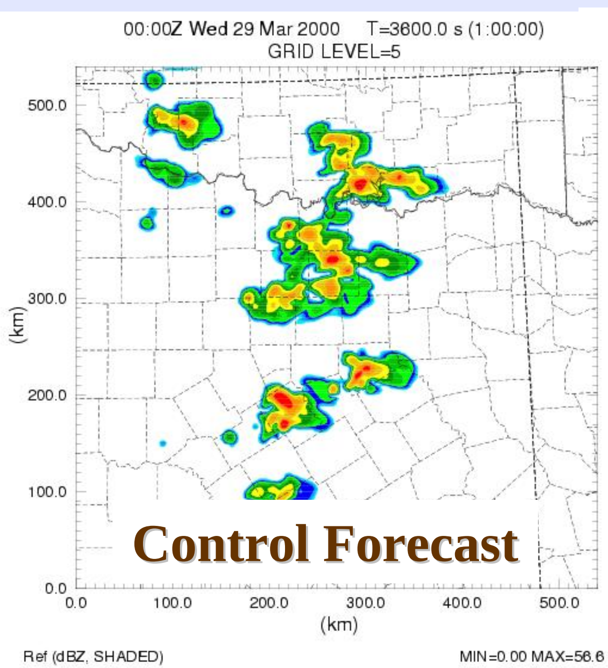
5 ENSEMBLES

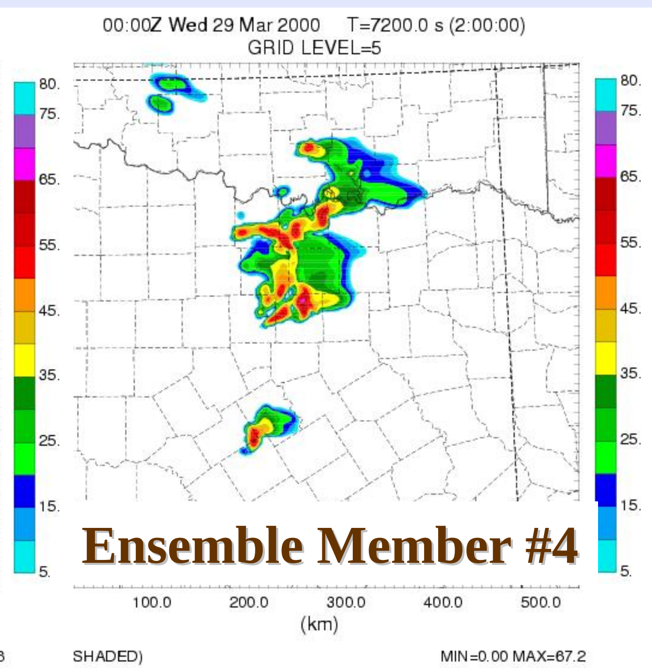

00:00Z Wed 29 Mar 2000 T=7200.0 s (2:00:00)
GRID LEVEL=5
(km)
Ensemble Member #1


Reflectivity at Elevation Angle= 0.48
Scan Time: 3/29/2000 0: 0
KFWS
Actual Radar


00:00Z Wed 29 Mar 2000 T=7200.0 s (2:00:00)
GRID LEVEL=5
(km)
Ensemble Member #2

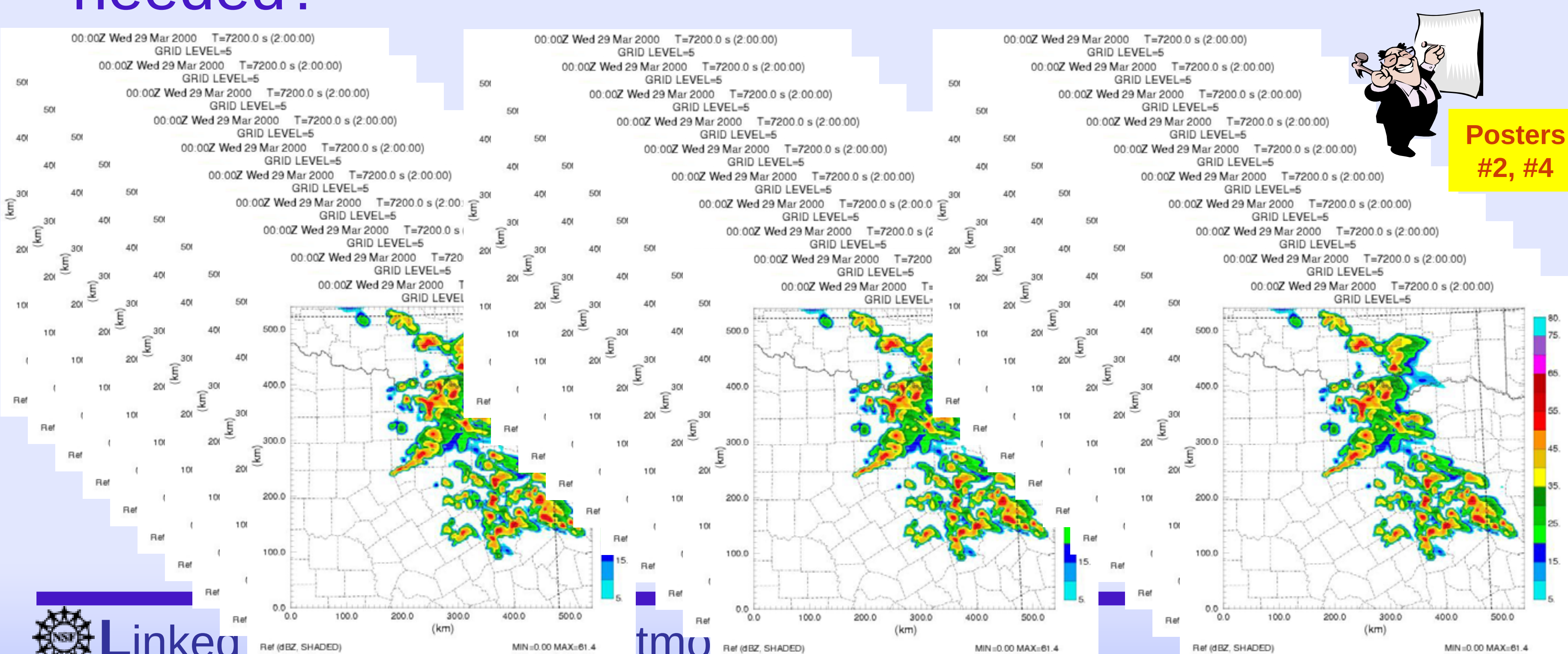

00:00Z Wed 29 Mar 2000 T=7200.0 s (2:00:00)
GRID LEVEL=5
(km)
Ensemble Member #3
Ref (dBZ, SHADED)
MIN=0.00 MAX=47.8


00:00Z Wed 29 Mar 2000 T=3600.0 s (1:00:00)
GRID LEVEL=5
(km)
Control Forecast
Ref (dBZ, SHADED)
MIN=0.00 MAX=56.6


00:00Z Wed 29 Mar 2000 T=7200.0 s (2:00:00)
GRID LEVEL=5
(km)
Ensemble Member #4
SHADED)
MIN=0.00 MAX=67.2

Resources and Decisions

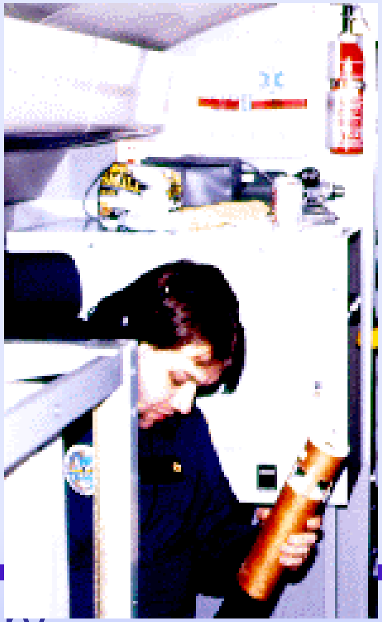
- How many ensembles are needed?
- Can the model or observations provide the answer?
- Will sufficient cyberinfrastructure be available when needed?



Posters #2, #4

Can Observations Also be Dynamic?

- Targeted observations have been explored at large scales for quite some time
- How does one do this at the mesoscale, where radars are the foremost observing tool?

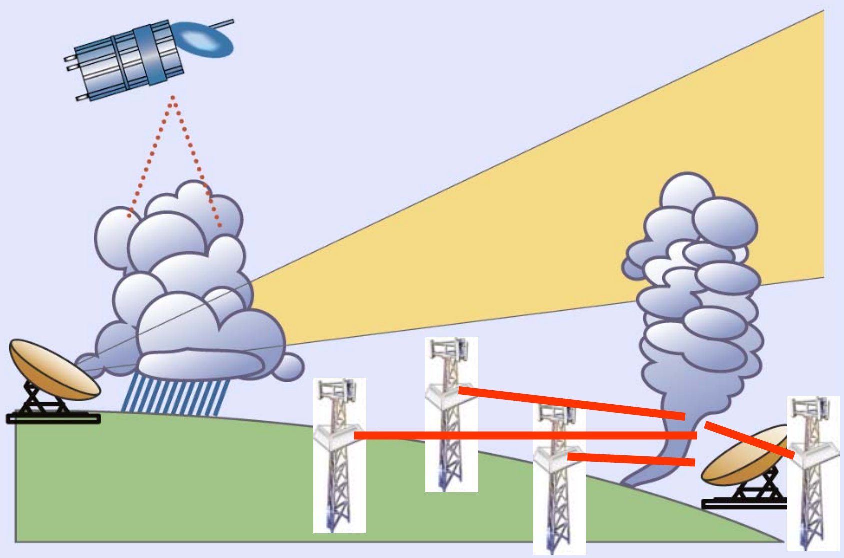


Courtesy US Navy and FASTEX Program

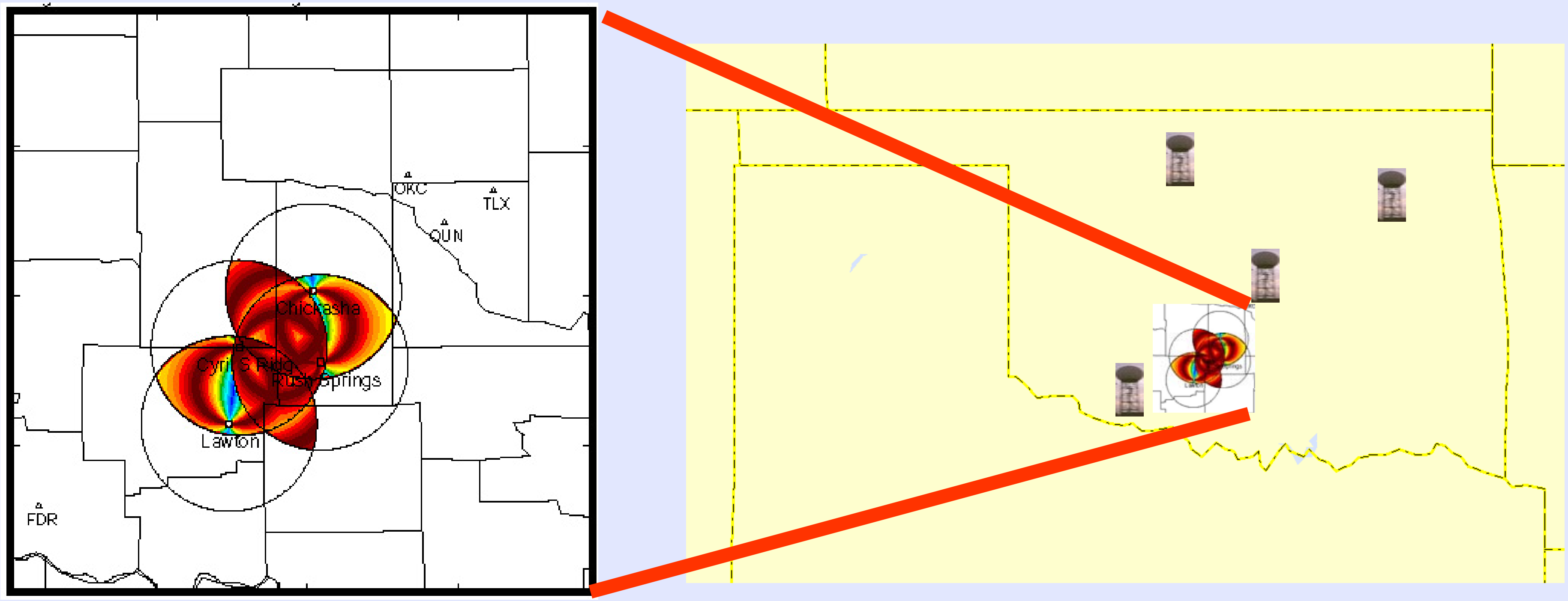
NSF Engineering Research Center for Adaptive Sensing of the Atmosphere (CASA)

Poster #5

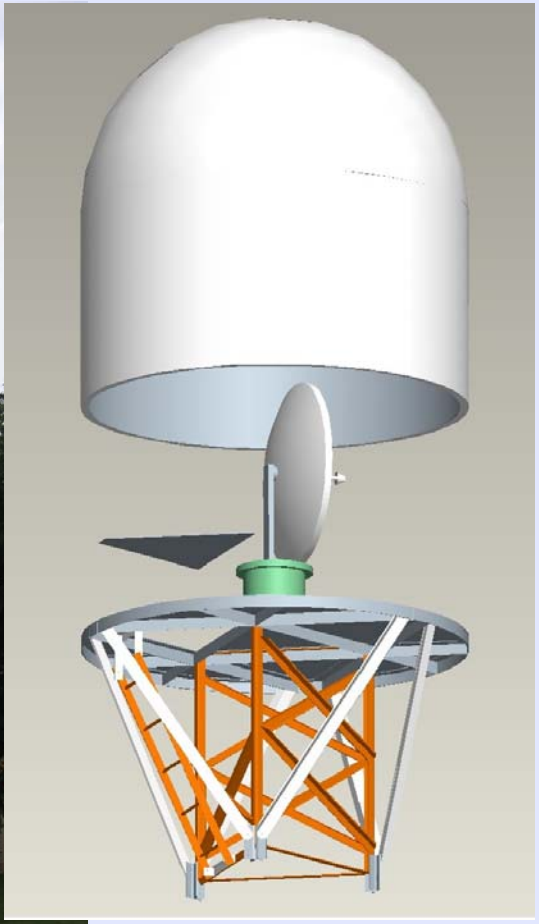
- Concept: inexpensive, dual-polarization phased array Doppler radars on cell towers and buildings
- Adaptive dynamic sensing of multiple targets simultaneously
- UMass/Amherst is lead institution, OU & CSU are core partners



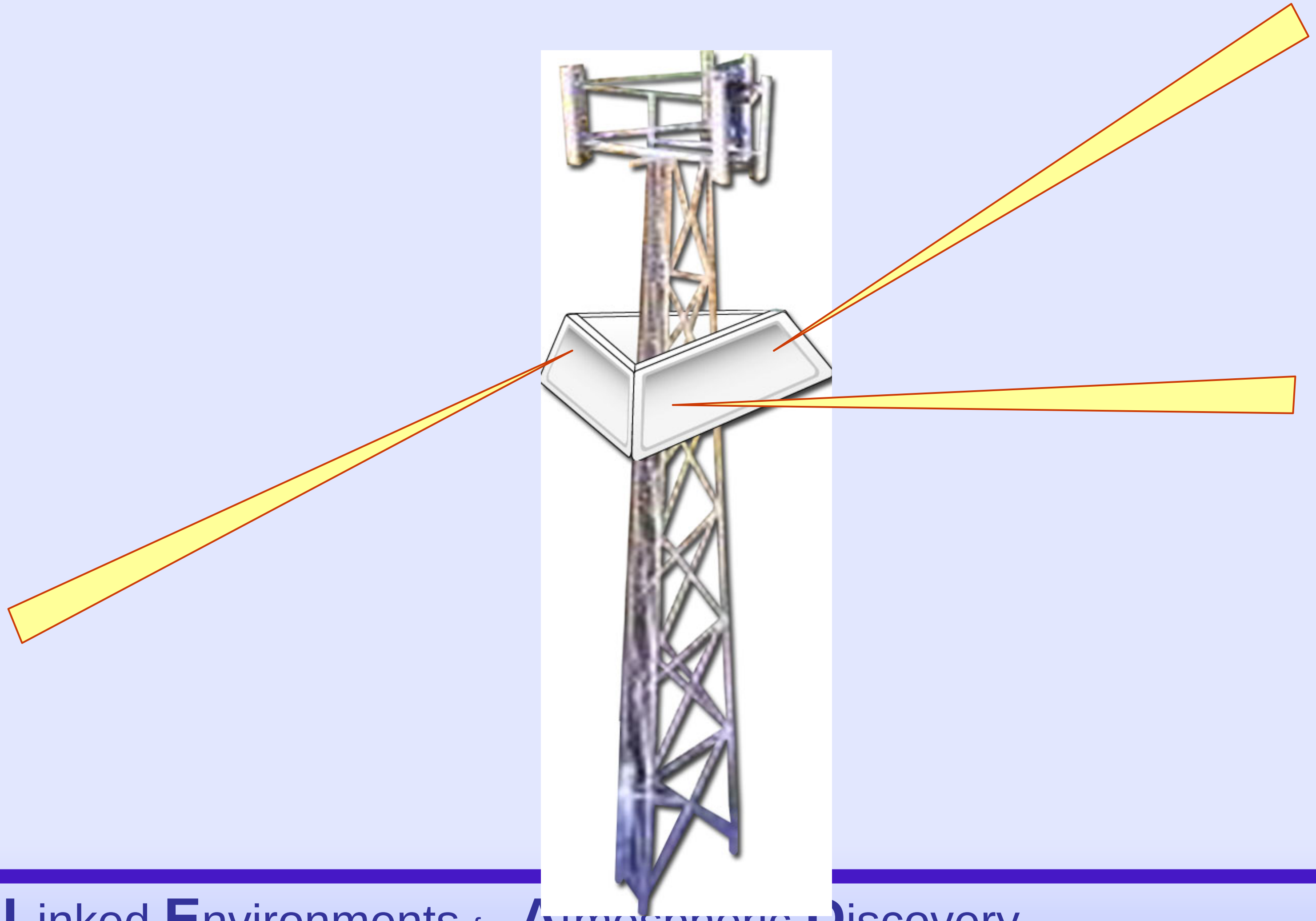
Oklahoma Test Bed: Spring 2006



Spring 2006: Mechanical Scanning



By 2008: Fully Solid State Electronic Scanning + More Radars



LEAD-CASA Interaction

Posters #1, #5

- CASA provides
  - dynamically adaptive radars
- LEAD provides
  - portal and distribution of data
  - meta data schema
  - dynamic weather applications
- Numerous experiments now underway using simulated CASA data
- Joint CASA-LEAD science meeting to be held during AMS Radar and Mesoscale Conferenced in October, 2005

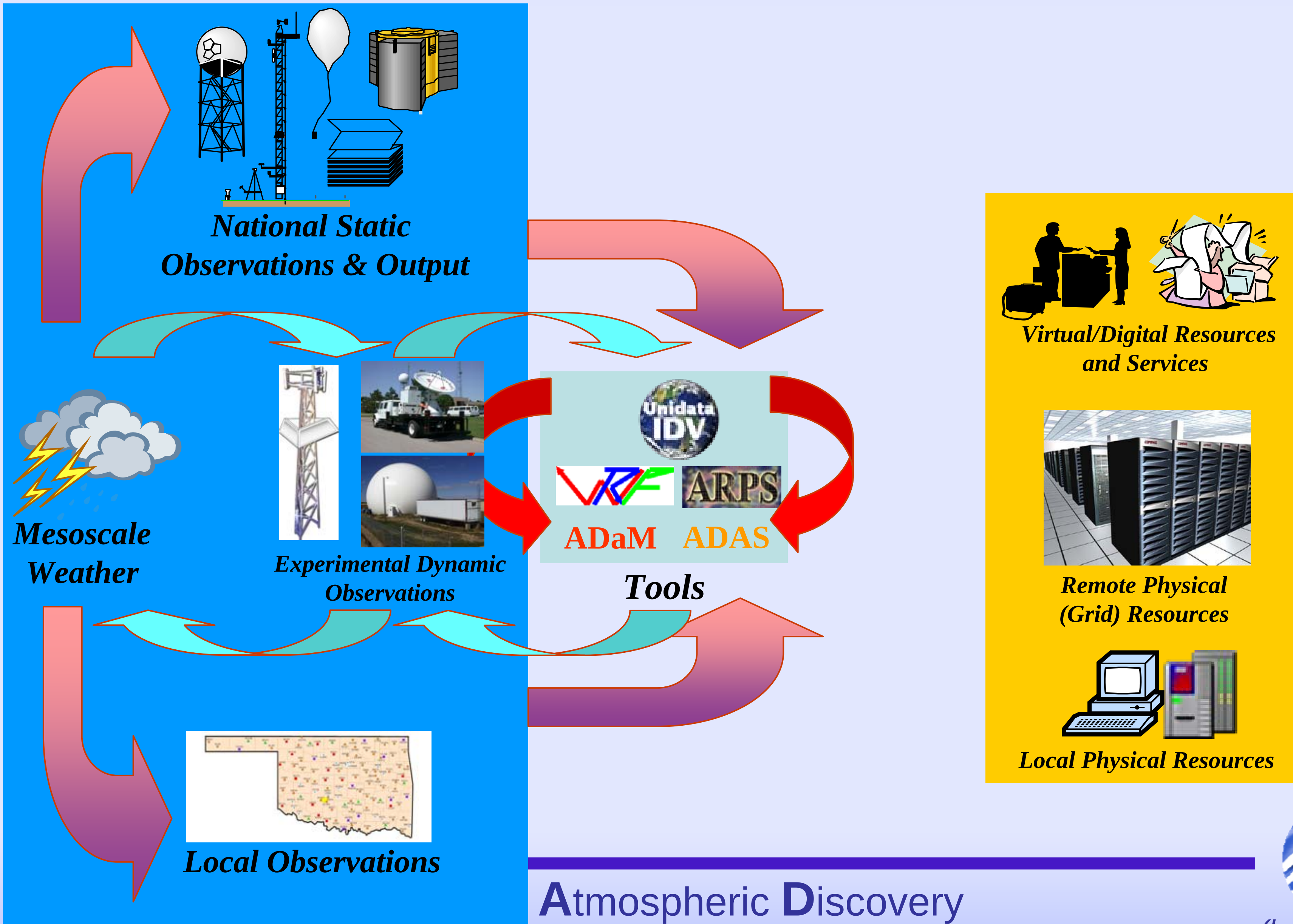
Another Component of Adaptation

- **Cyberinfrastructure including application and environment monitoring, performance estimation and scheduling**

Poster #7

- **More on this ... nya**

The Three Components of Adaptation



Key CS Research Challenges

- Creating sequences (workflows) of complex interacting services that can change dynamically with time
- Synchronizing applications, data, computing and networking for automated on-demand, fault-tolerant execution
- Monitoring performance, detecting vulnerabilities, and estimating resource needs in a dynamic environment
- Coupling remote sensing systems and streaming data with real time weather applications
- Distributing work across heterogeneous systems in a grid-based architecture

Key Meteorology Research Challenges

- Packaging complex applications (e.g,. prediction models) as easy-to-use, fault-tolerant services
- Accommodating truly real time streaming observations in data assimilation and prediction
- Developing ensemble strategies for deep convective storms
- Automatically detecting weather features in assimilated data sets versus raw sensor data
- Developing strategies for dynamically adaptive forecasting and comparing them to traditional static methodologies

The LEAD Framework/Environments

Fundamental Capabilities

Query & Acquire | Simulate | Assimilate | Analyze/Mine | Predict | Vi e

Foundational User Tools

Portal | ADAS | MyLEAD | WRF | ADaM

Enabling Functions

Define | Query | Store | Publish | Allocate | Monitor | Edit | Authorize | Build | ge

Acquire | Execute | Compile | Steer | Catalog | Configure | Move | Trigger | ode

Middleware Services

- Authorization
- Authentication
- Notification
- Monitoring
- Workflow
- Security
- ESML

- VO Catalog
- THREDDS Catalog
- MyLEAD
- Control
- Query
- Stream
- Transcoder

- Ontology
- Host Environment
- GPIR
- Application Host
- Execution Descripti
- Application Descript

Outcomes

Data Sets | Model Output | Gridded Analyses | Animations | Static Images | Relatio

Meteorology

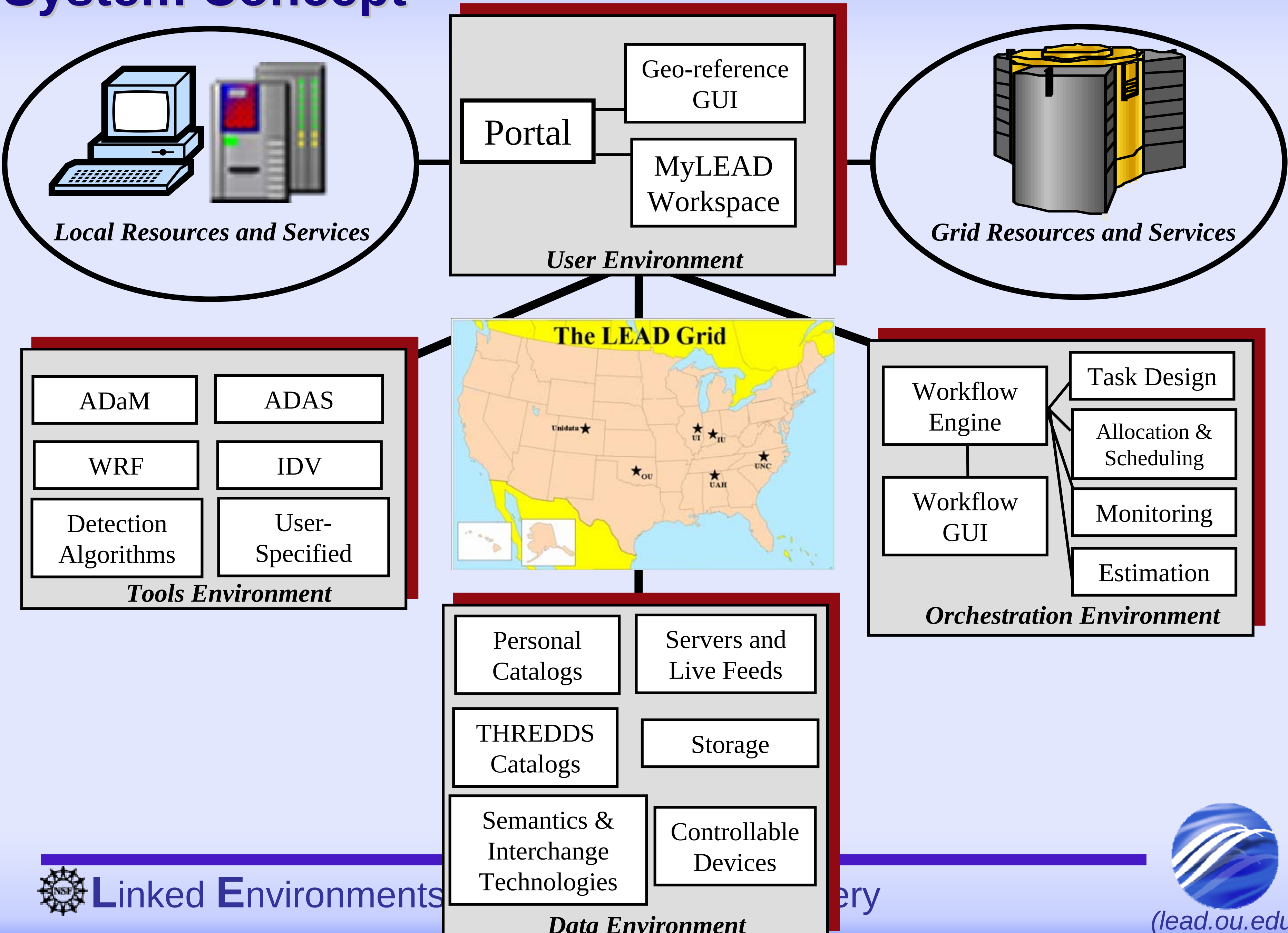
CS

CS

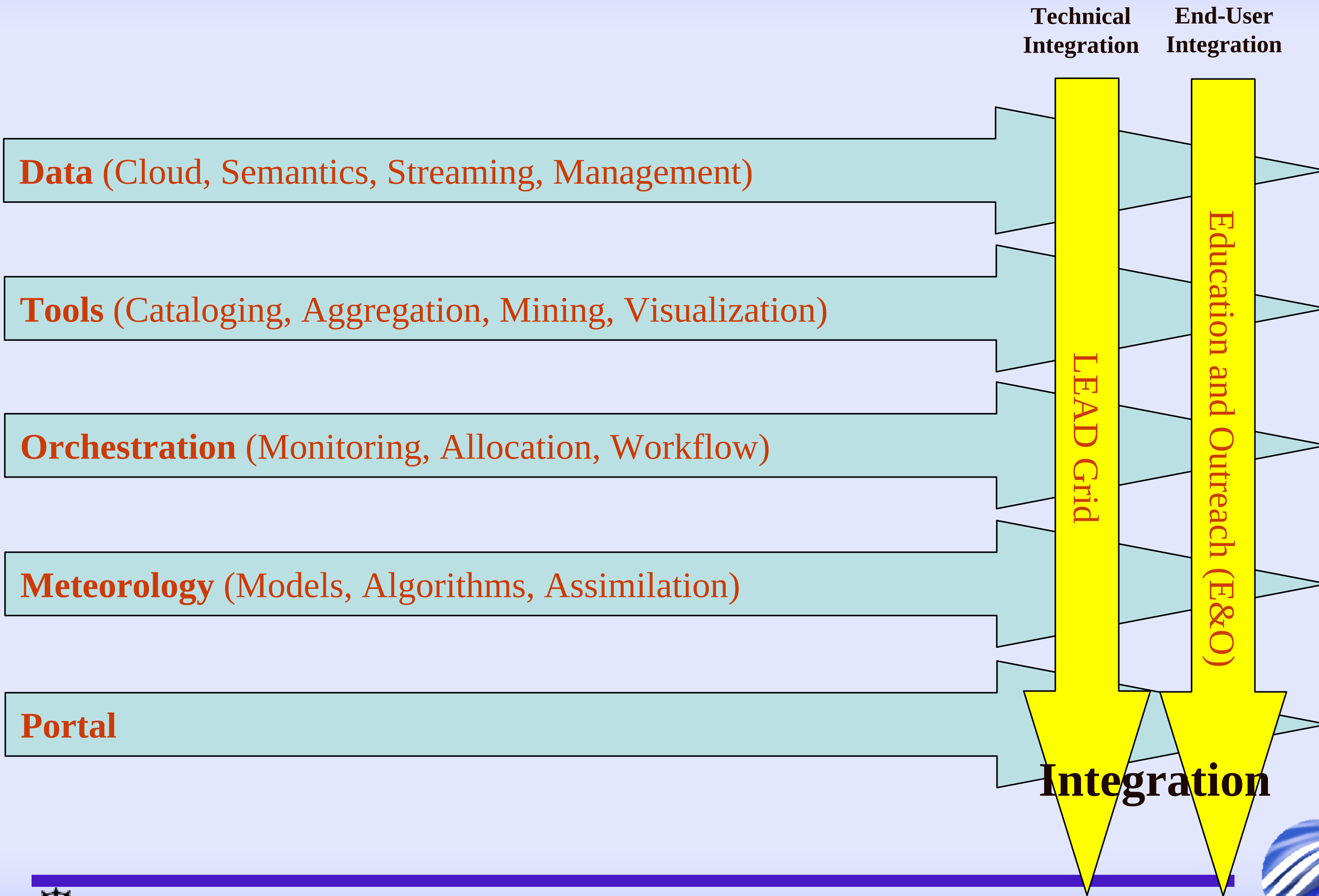
Meteorology

New Knowledge, Understanding, Ideas

System Concept



Research Organization



LEAD Service-Oriented Architecture

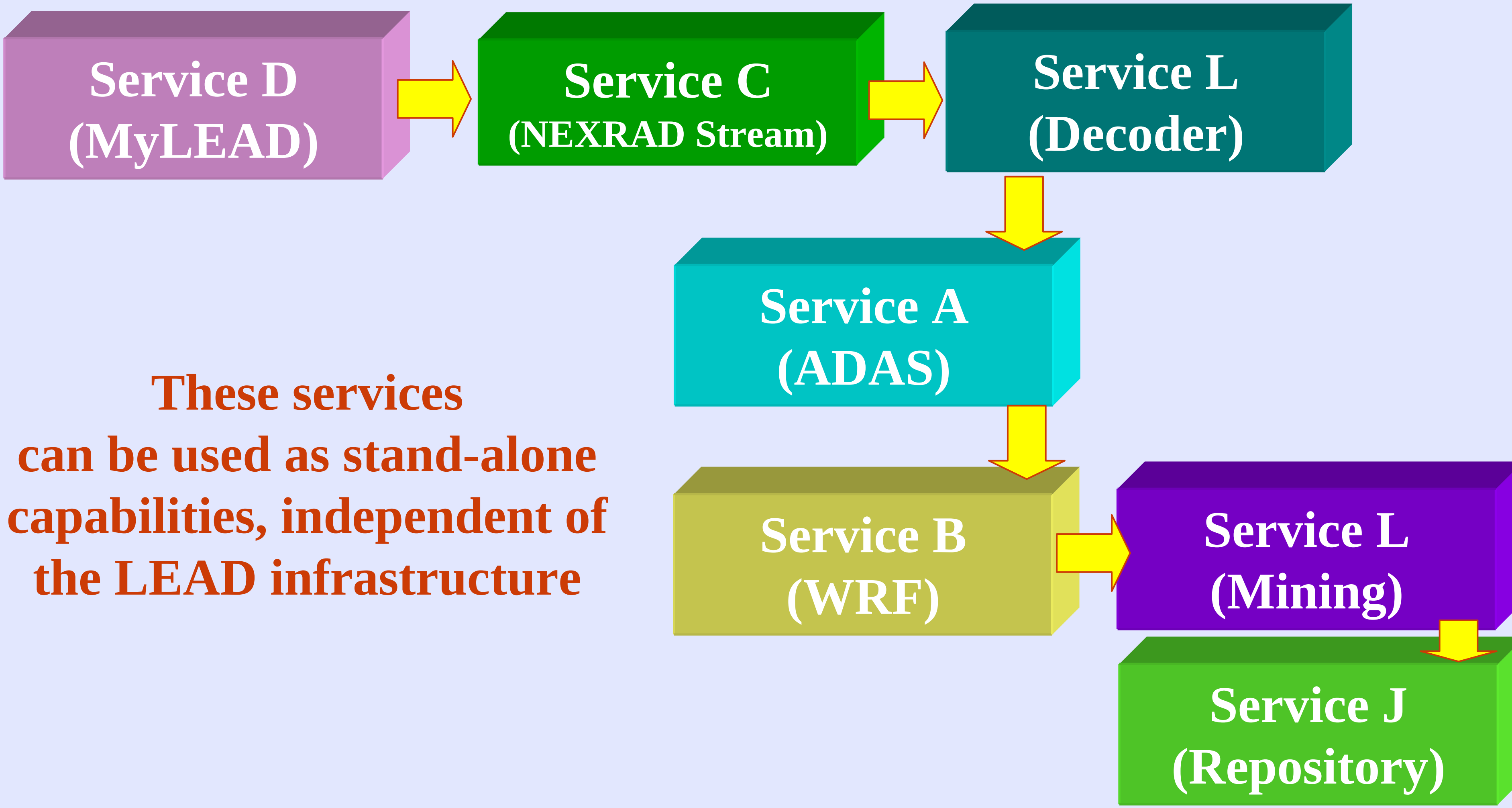
Why A Service-Oriented Architecture?

- Flexible and malleable
- Platform independence (emphasis on protocols, not platforms)
- Loose integration via modularity
- Evolvable and re-usable (e.g. Java)
- Interoperable by use of standards → robustness

Establish the Services

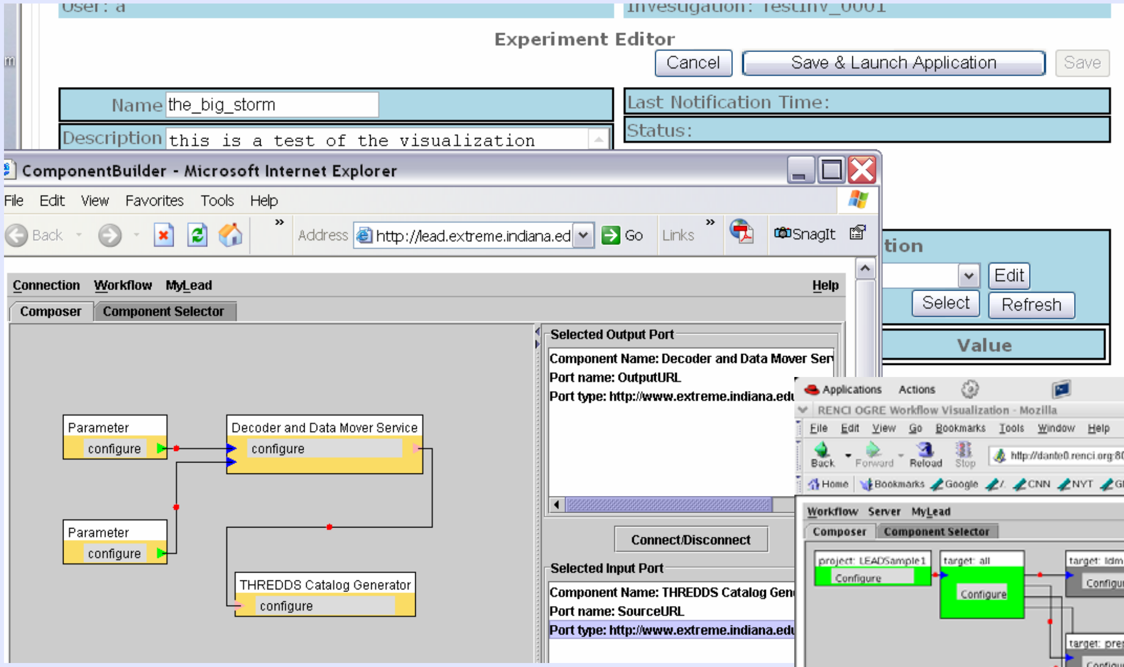
- Service A (ADAS)
- Service B (WRF)
- Service C (NEXRAD Stream)
- Service D (MyLEAD)
- Service E (VO Catalog)
- Service F (IDV)
- Service G (Monitoring)
- Service H (Scheduling)
- Service I (ESML)
- Service J (Repository)
- Service K (Ontology)
- Service L (Decoder)
- Many others…

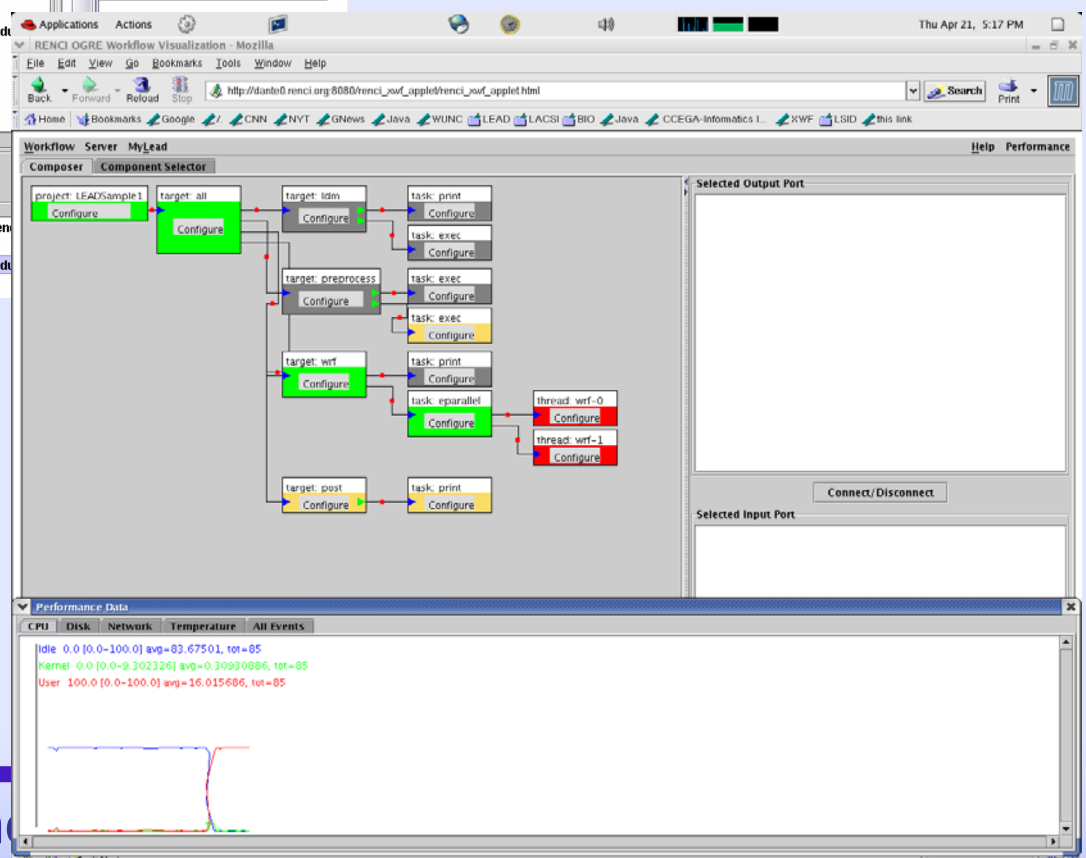
…and then Solve General Problems by Linking them Together in Workflows



These services can be used as stand-alone capabilities, independent of the LEAD infrastructure

raphical Workflow Composer & Monitori





Posters #2, #3, #20

Leveraging + Basic Research

Capability/Resource	Principal Technologies
Atmospheric, Oceanographic, Land-Surface Observations	CONDUIT, CRAFT, MADIS, IDD, NOAAPort, GCMD, SSEC, ESDIS, NVODS, NCDC
Operational Model Grids	CONDUIT, NOMADS
Data Assimilation Systems	ADAS, WRF 3DVAR
Atmospheric Prediction Systems	WRF, ARPS
Visualization	IDV
Data Mining	ADaM
NSF NMI Project	Globus Tool Kit
Semantic Interchange and Formatting	ESML, NetCDF, HDF5
Adaptive Observing Systems (Radars)	CASA OK Test Bed, V-CHILL
LEAD Portal	NSF NMI Project (OGCE)
Workflow Orchestration	BPEL4WS
Monitoring	Autopilot
Data Cataloging/Management	THREDDS, MCS, SRB

System Functional Requirements, Service Descriptions, and Architecture Design & Implementation Plan

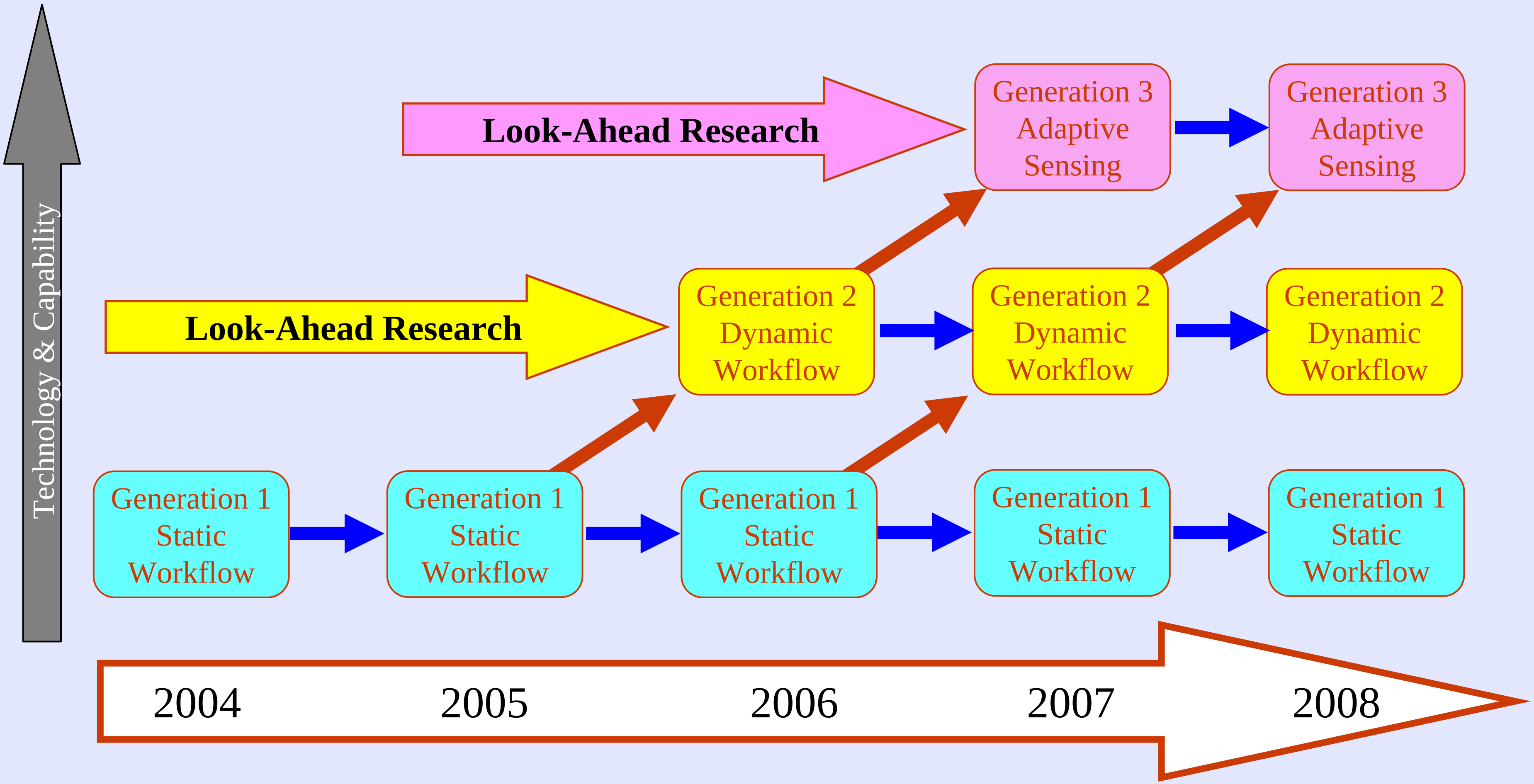
Version 2.0

LINKED
ENVIRONMENTS
FOR ATMOSPHERIC
DISCOVERY

10 July 2005

- Core Values
- System Capabilities
- Tools Description
- Functionality
- Design Principles
- Assumptions, Policies, Constraints
- Functional Requirements
  - User
  - Technical
- Architecture
- Implementation

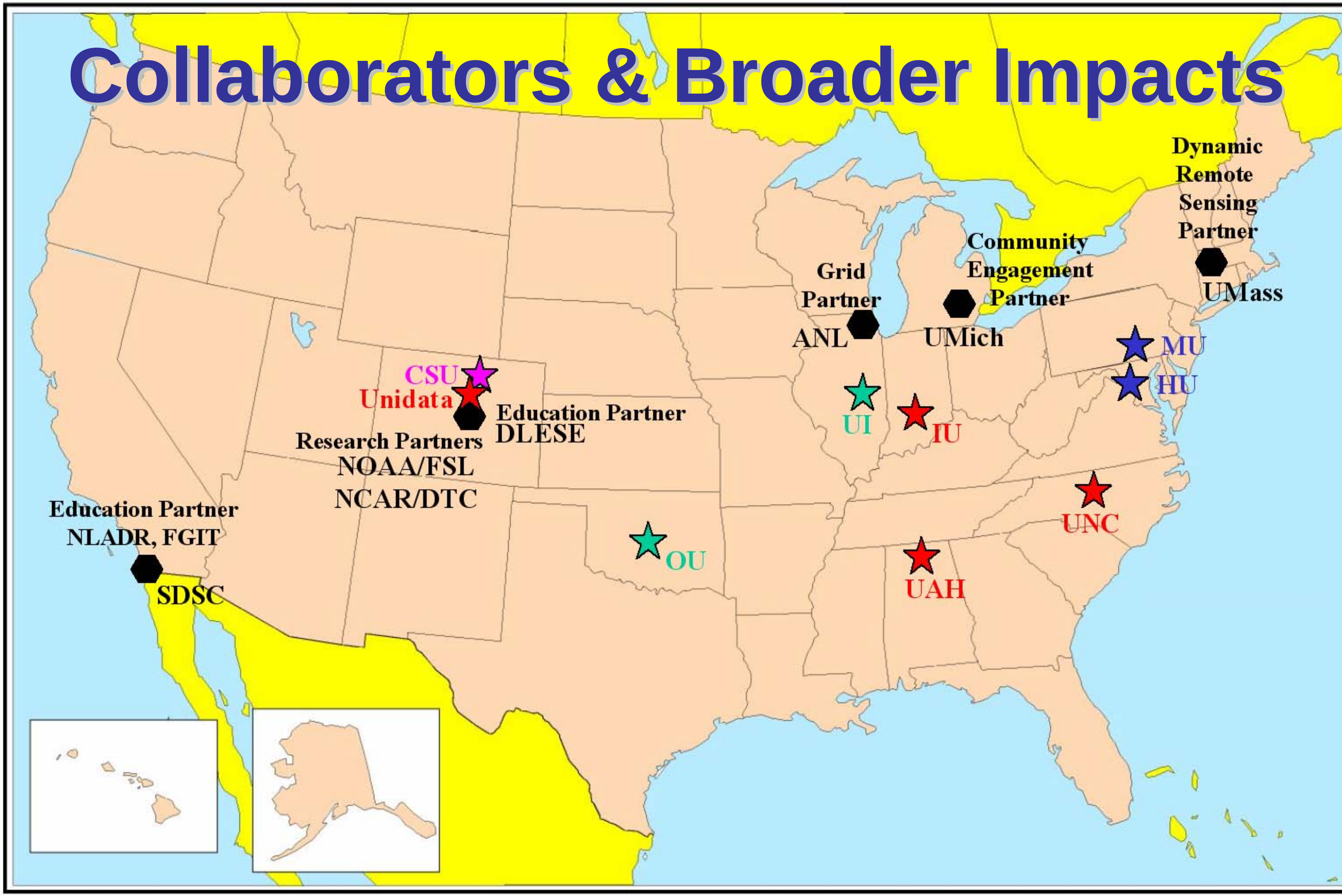
LEAD Technology Generations



Bringing it Together and Making it Work: Technical Integration

- Rapid prototyping
- Systems and software engineering
- Integrating
- Hardening
- Extending beyond the LEAD Grid
- Measuring impact: Has LEAD made a difference?

Collaborators & Broader Impacts



Core Academic Partner

Core Academic Partner + Grid Test Bed

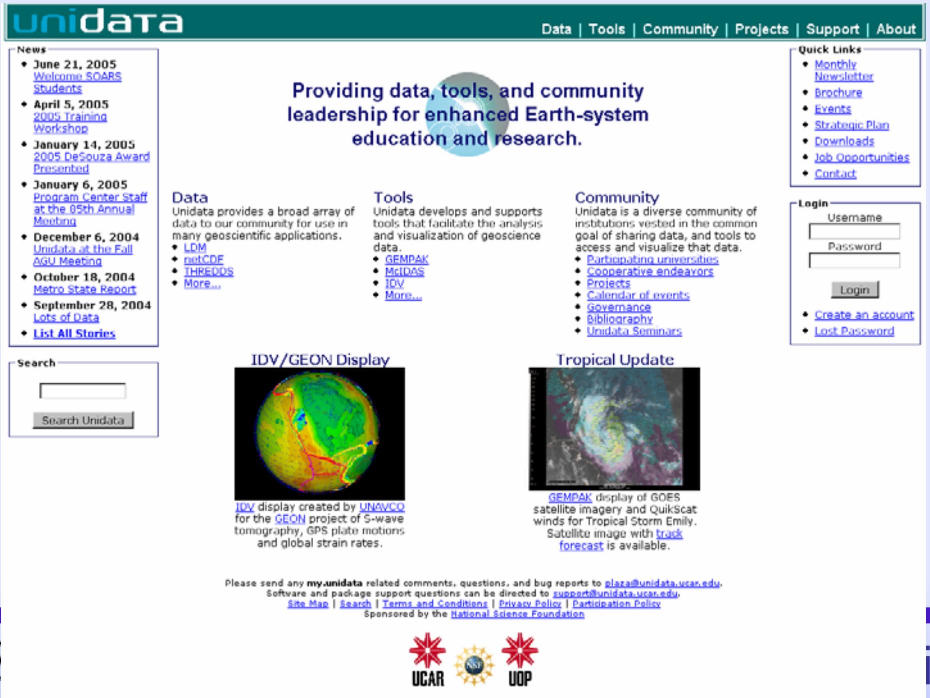
Core Academic Partner + Education Test Bed

Core Academic Partner + Grid Test Bed + Education Test Bed

Special Partners

Deployment & Community Engagement

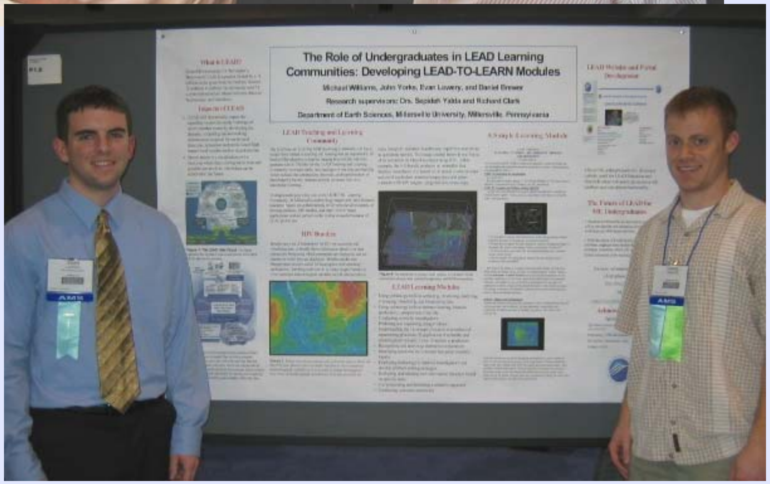
- **Audience:** Higher education, operations research, students and teachers in grades 6-12
- Scalability and extensibility are key
- LEAD seeks to create a maintainable system that will have sustained broad impact



Education

Posters
#15-18, 22-23

- Led by Howard University and Millersville University
- 3 Phases
- 6 Education Test Beds
- National Network of Teacher-Partners
- LEAD Learning Communities
- LEAD-to-LEARN Education Modules
- EarlyLEAD
- Dynamically Adaptive Learning
- Links to the CASA WeatherRATS Program
- CI-TEAM Proposal



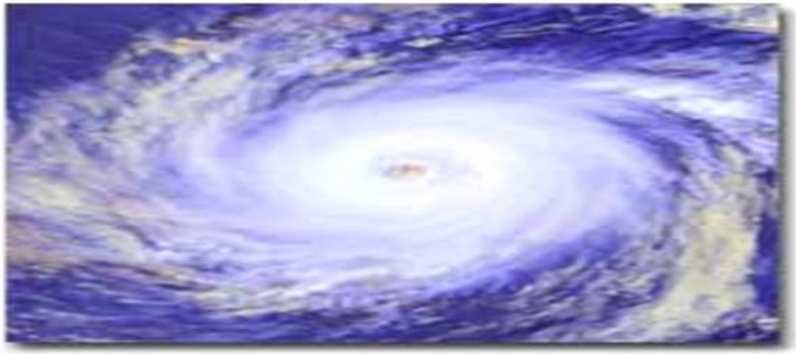
Diversity

- Led by Howard University
- Emphasizes end-to-end engagement of women and minorities and capacity building
  - Jackson State, UTEP, UPRM, SOARS
- NCAS Weather Camp (2005)
- WRF Workshop (2006)

Weather Camp 2005

Hosted by the Howard University

NOAA Center for Atmospheric Sciences



WHEN: July 10 – July 29, 2005

WHERE: Howard University – Washington , DC

FEE: **THERE IS NO COST TO THE PARTICIPANT**

REQUIREMENTS: Scholars must have an interest in weather, meteorology, atmospheric sciences, environmental sciences, or applied physical sciences

WHO SHOULD APPLY: All rising high school juniors and seniors (MUST BE U.S. CITIZENS OR PERMANENT RESIDENTS)

DEADLINE: Completed applications must be received no later than Friday May 27, 2005 by 5:00pm EST

Research Accomplishment Highlights

- Built the LEAD Grid and deployed related services
- Converted all of the major application components to services
- Developed and deployed the LEAD Portal
- Developed the myLEAD personal work space
- Completed 85% of supporting services including prototype monitoring system
- Developed the experiment builder
- Developed graphical composer and compiler for static workflows
- Developed geographical reference GUI
- Developed ontology, meta data schema, catalogs
- Developed data query system
- Demonstrated functionality for a complete WRF forecast scenario
- Demonstrated value of weather detection in assimilated data sets in comparison to raw observations

Posters #9, #11, #13

Posters #8, #12, #19

Poster #6

Quantitative Output to Date

- 49 conference and journal articles
- Several Internet disseminations
- Degrees awarded (all in computer science)
  - 3 MS at IU
  - 1 MS at OU
  - 4 MS at UAH
  - 1 Ph.D. at IU
  - 2 other MS degrees in CS at IU for work supporting LEAD
- 56 oral presentations

Research Challenges Ahead of Us: Managing Risk

Poster #21

- Dynamic workflow – the most difficult
- Moving beyond the LEAD Grid to TeraGrid
- → The shifting sands of our underpinnings…
- Steering of the CASA radars
- Software extensibility and maintainability
- Scalability
- Granularity (number and complexity) of services
- Component integration
- User-friendliness in interfaces (HCI)

Year-1 (9/04)	Year-2 (9/05)	Year-3 (9/06)	Year-4 (9/07)	Year-5 (9/08)
Establish LEAD Grid and Data Streams	1st Generation SOA, Portal, Orchestration, Metadata Catalog	Usability testing	Community Outreach (DTC and others)	Full control of CASA radars
Define User and Tech Requirements	Deployment of Services on Testbeds	Data & Tools subsystem integration	On-Demand with TeraGrid	Fully functional on TeraGrid
Engage Education Partners	Link Educational Materials with LEAD Tools	Initial control of CASA radars	Dynamically adaptive learning	Deployment to DTC and Unidata
Define Service Architecture	Real Data WRF Forecasts Using Static Orchestration	Adaptive workflows	Fault tolerance	Continuation of selected components
Enable WRF Real Data Forecasts		TeraGrid gateway		
		SPC Experiments		

Administrative Overview

LEADership

National Science Foundation

PI and Project Director (K. Droegemeier)

External Advisory Panel

Executive Committee (PIs)

Forum on Geosciences Information Technology (FGIT)

Project Coordinator (T. Leyton)

Technical/Integration Coordinator (R. Wilhelmson)

Technical Integration Assistant (T. Baltzer)

Facilities and Technical Integration Committee (Technical Liaisons)

Partner Organizations

University of Oklahoma
(K. Droegemeier, PI; M. Xue, K. Brewster, D. Weber, Co-PIs)
Meteorological Research, Education

University of Alabama in Huntsville
(S. Graves, PI; R. Ramachandran, J. Rushing, Co-PIs) *Data Mining, Interchange Technologies, Semantics*

UCAR/Unidata
(M. Ramamurthy, PI; B. Domenico, D. Murray, A. Wilson, Co-PIs)
Data Streaming and Distributed Storage

Indiana University
(D. Gannon, PI; Beth Plale, Co-PI)
Data, Workflow, Orchestration, Services

University of Illinois/NCSA + UNC
(R. Wilhelmson, PI + D. Reed, PI)
Monitoring and Data Management

Millersville University
(R. Clark, PI; S. Yalda, Co-PI)
Education and Outreach

Howard University
(E. Joseph, PI; V. Morris, Co-PI)
Meteorological Research Education and Outreach

Colorado State University
(Chandra, PI)
Instrument Steering, Dynamic Updating

u)

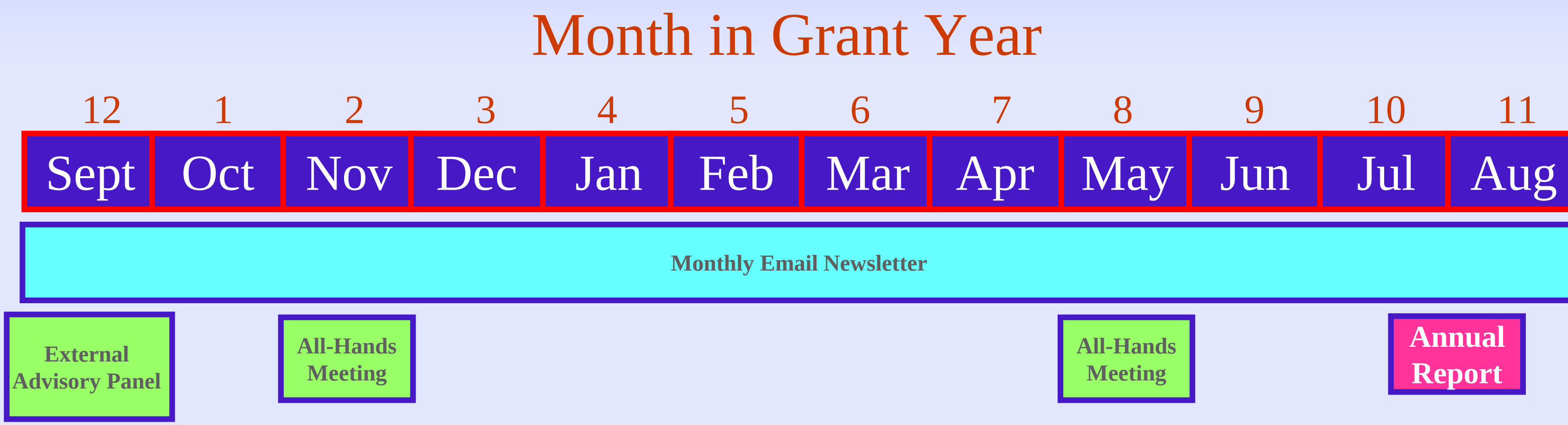
Participants

Institution	FTE on Project	Leveraged Activities (FTE)
University of Oklahoma (OU)	3.65	0.87
University of Illinois at Urbana-Champaign (UIUC)	2.00	0.95
University of Alabama in Huntsville (UAH)	3.21	1.58
UCAR Unidata Program	3.65	1.05
Indiana University (IU)	4.99	3.18
Howard University (HU)	2.44	1.21
Millersville University (MU)	1.19	0.30
Colorado State University (CSU)	0.00	0.00
University of North Carolina (UNC)	1.66	0.00
Totals	**22.79**	**8.14**

Total Number of LEAD Personnel

(Funded Directly by LEAD + Leveraged From Other Projects)

	OU	UIUC	UAH	Unidata	IU	HU	MU	CSU*	UNC		Total
Faculty	2	2	1	0	2	3	2	0	1		13
Non-Faculty Researcher	5	8	7	2	3	1	0	0	2		28
Technical Support Staff	2	2	1	9	0	0	1	0	4		19
Graduate Student	2	0	5	0	9	5	0	0	3		24
Undergraduate Student	0	0	0	0	0	0	12	0	0		12
Other	0	0	0	0	0	2	0	0	0		5
Total	11	12	14	11	14	11	16	0	10		101



Purpose	Location	Date
Special Metadata Planning	UAH	14-15 April 2004
Regular External Advisory Panel Meeting	OU	8 September 2004
Special Portal Planning Meeting	IU	27-28 September 2004
Special Meeting Between LEAD and DLESE/DTC	Unidata	14 October 2004
Special Metadata Planning Meeting	UAH	20 October 2004
Special Data Thrust Meeting	UAH	21 October 2004
Regular Fall All-Hands Meeting	Pittsburgh, PA (SC 2004)	12-13 November 2004
Special Architecture Planning Meeting	NCSA	9-10 December 2004
Special CI-TEAM E&O Proposal Planning Meeting	Unidata	6-7 March 2005
Special Metadata Planning Meeting	OU	12 May 2005
Regular Spring All-Hands Meeting	OU	12-13 May 2005
Year-2 NSF Site Visit	UIUC	21-22 July 2005

Weekly Meetings (16 per month)

	Mon	Tue	Wed	Thu	Fri
Week 1	Data & Tools AG		Portal AG Integration/Prototype AG		
Week 2		Education	Integration/Prototype AG Meteorology AG	Orchestration	GWSTB AG
Week 3	Data & Tools AG		Portal AG Integration/Prototype AG		
Week 4		Education	Integration/Prototype AG Meteorology AG	Orchestration	GWSTB AG

Weekly Meetings (8 per month)

	Mon	Tue	Wed	Thu	Fri
Week 1				2-hour All-Hands AG Session	1-hour PI Conference Call
Week 2				2-hour All-Hands AG Session	1-hour PI Conference Call
Week 3				2-hour All-Hands AG Session	1-hour PI Conference Call
Week 4				2-hour All-Hands AG Session	1-hour PI Conference Call

Summary

- LEAD is
  - Evolving a new paradigm for studying the atmosphere: Interacting dynamically with weather
  - Greatly simplifying the application of sophisticated capabilities
  - Bringing all of this to the masses for use on local as well as remote (grid) resources
  - A driver for research in the grid and web services communities – broader impacts to other disciplines
  - Working to have a sustained impact in the community beyond its own lifetime and beyond the involvement of its developers (via Unidata, DTC)