A Quick Tour of LEAD via the LEAD Portal & Application Orchestration

Dennis Gannon, Marcus Christie, Suresh Marru Indiana University

Year-2 Site Visit 21-22 July 2005



Outline

- The Big Picture
- The Portal overview
- Looking at your MyLEAD space
- How to build a simple workflow
 - Use remote ADaM services to do data mining.
- Creating and running an experiment

 Launching a large simulation
- Monitoring the performance of WRF



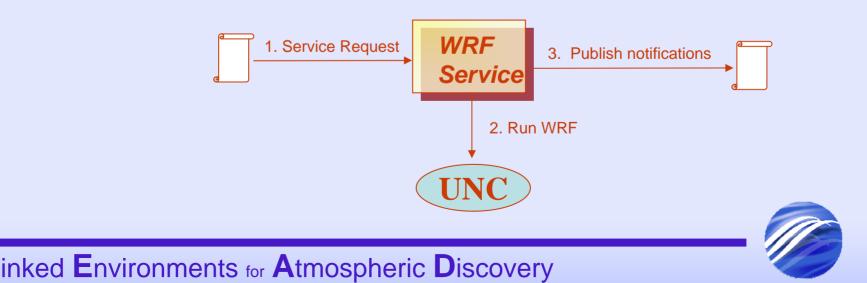
The User Perspective

- The Lead System Portal is designed to allow users to
 - -Explore data
 - -Compose, run and catalog experiments
 - -Learn by interacting with the weather.
- The Foundation of the software is based on the concept of "services" and "service oriented architectures."



What's a service anyway?

- A "web server" that runs an application for you.
 - You send it requests (XML documents) and it processes the information and send replies (notifications) when it is done.
- More on this concept and service architectures later (Beth's talk)



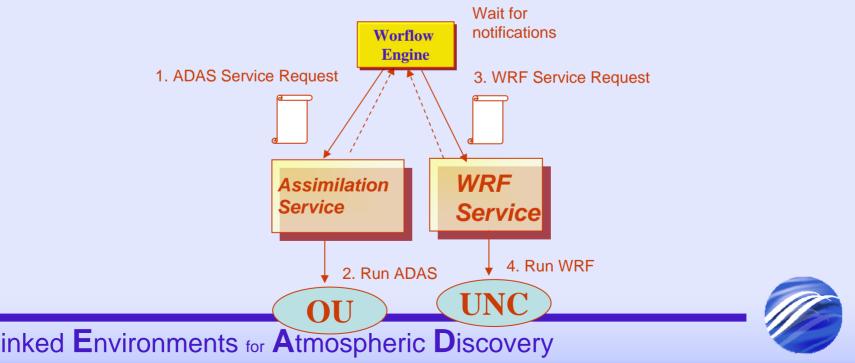
System Architecture Components

- Data Services
 - MyLEAD : a personal metadata catalog
 - Resource Catalog : finding data and services
 (much more: next talk!)
- Experimental & Workflow Services
 - A way to compose a set of "remote services" into an application.
 - A way to run the application
 - A way to share it with others.



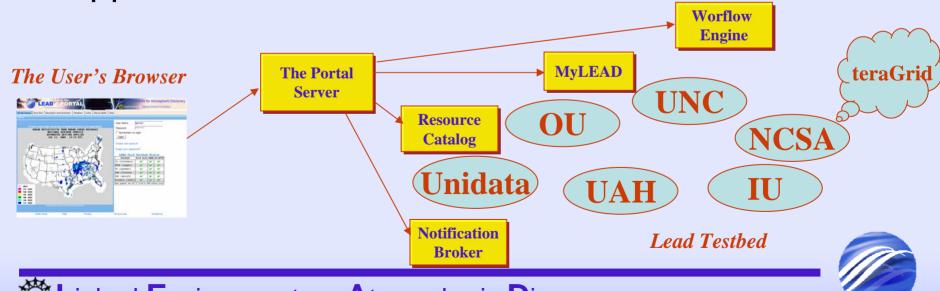
What's a Workflow?

- A sequence of service interactions that are needed to accomplish a bigger task.
 - We use two different workflow systems
 - BPEL & OGRE



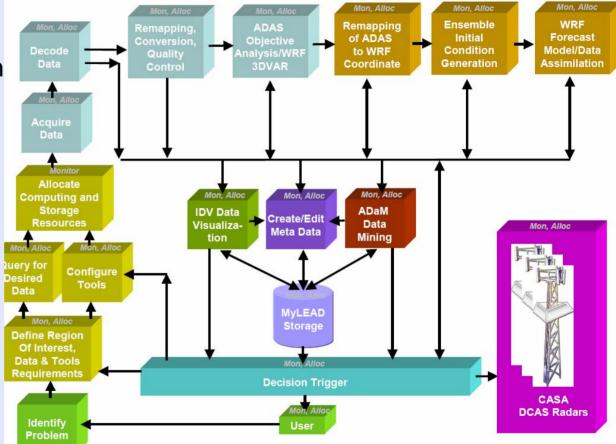
The Portal as a Gateway

- User interacts with a web portal
- The web portal talks to services running on the LEAD testbed "Grid" and, next year, teraGrid!
- These Services manage your data and run the applications



Building an application from Services

- Users create experiments by composing application services into "workflows"
- A workflow can be viewed as a pathway through this picture.
- When combined with data sources it is executed on the computational resources.



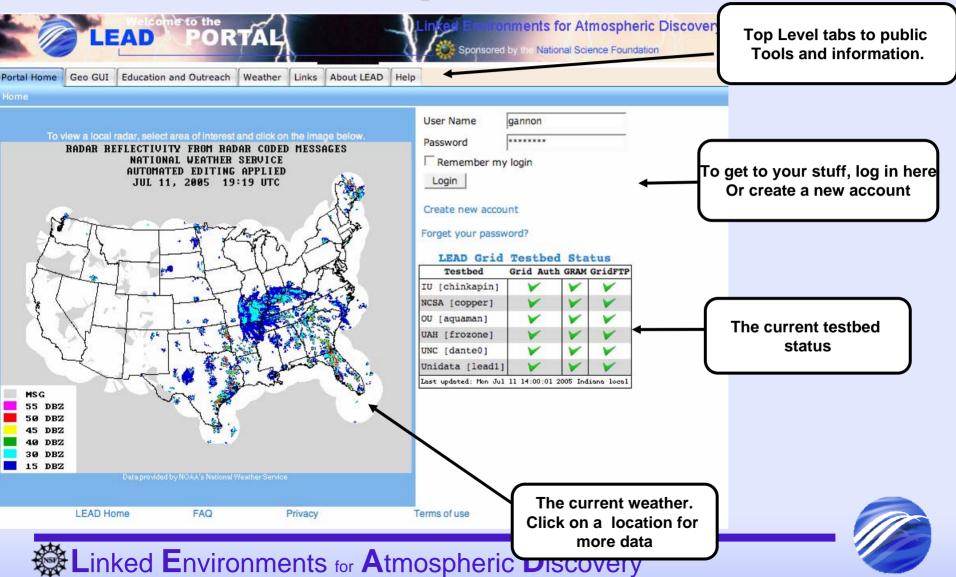


A Quick Look at the portal

- The top level
 - Access to public resources
 - A look at the Geographic region selector under development
- Logging in
 - Your private view of LEAD.
 - MyLEAD
 - Doing a search over old experiments.
 - A few words about security.



The top level view



GEO Reference GUI Prototype

Portal Home	Geo GUI	Education and Outreach	Weather	Links	About LEAD	Help
Geo Referenc	e GUI					
initialize						
Spatial Exte	ent					
En 🗨 🖲	Zoom to	State(s) 🗾 🖓 🖓 🤇	Ø			Selection Layers Legend
	.,					refresh
Rain	water -				· · · · /	Select by Rectangle Left: -180.0 Top: 90.00
MN	R	The St. Clair			Studente	Right: 180.00 Bottom: 0.00
K WI	E VA		vrence	₩ T <i>></i>	ME	C Select by Polygon
Wascon	sin of {	MI ST Still	ty and	NH NH	+ John Mar	(lon, lat) Pairs:
IA J	-1	Nill Start	Muenanas Sustuenanas	MA OCT B	n Se	
Calloo	9 IL 1	V ОН 2 РА		-		
Missour	har	WV POW	MD DE			1
	JA-	KY VA	- Al			
the of		V Alberta Contraction	Z			
1	15 AL	GA SC S				
LA	السوح ا	SOUTH V			_	
Lon, Lat: -88	.2528, 34.40	175				
Temporal R						
From Date:		Time: 12:00				
To Date:	07-11-05	Time: 9:00				
Data Produ	cts					
Hourly S						
🗌 12-hour	ly UpperAir I	Balloons (Rawinsondes)				

- Use mouse to drag a region of interest.
- Fill in the data requirements
- The tool, when finished will gather the data for you.



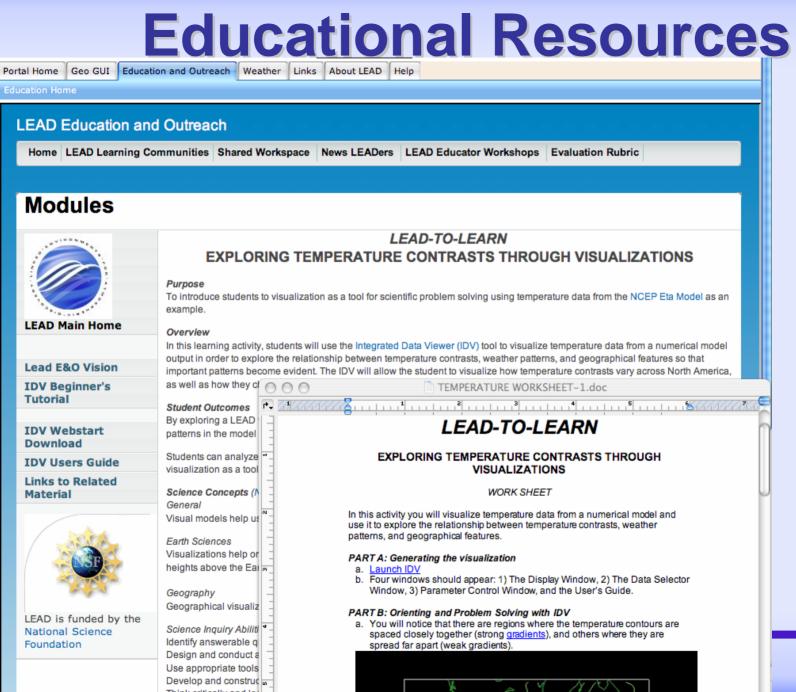
very

Half-hourly GOES Vis/IR

NEXRAD Level II (5-10 min)

NEXRAD Level III (5-10 min)

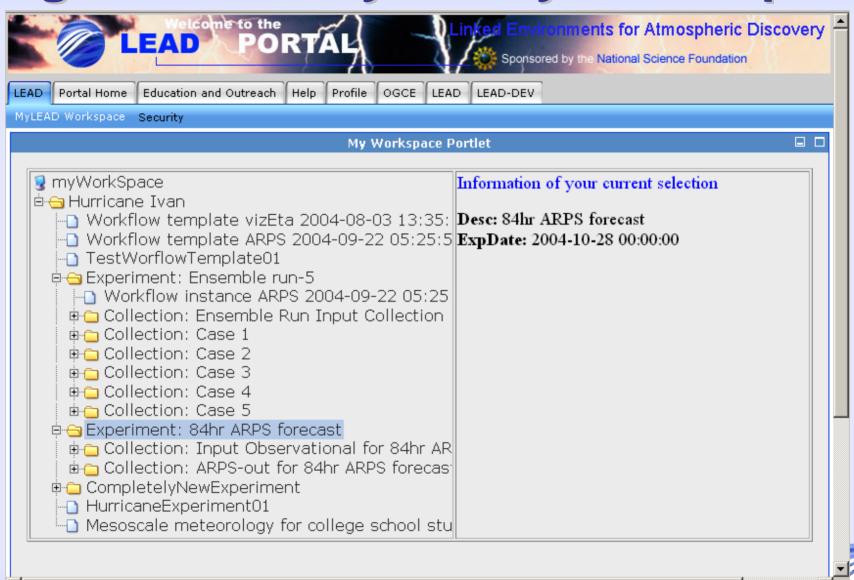
5-minute ACARS Commercial Aircraft T. V



The local standard and the standard law

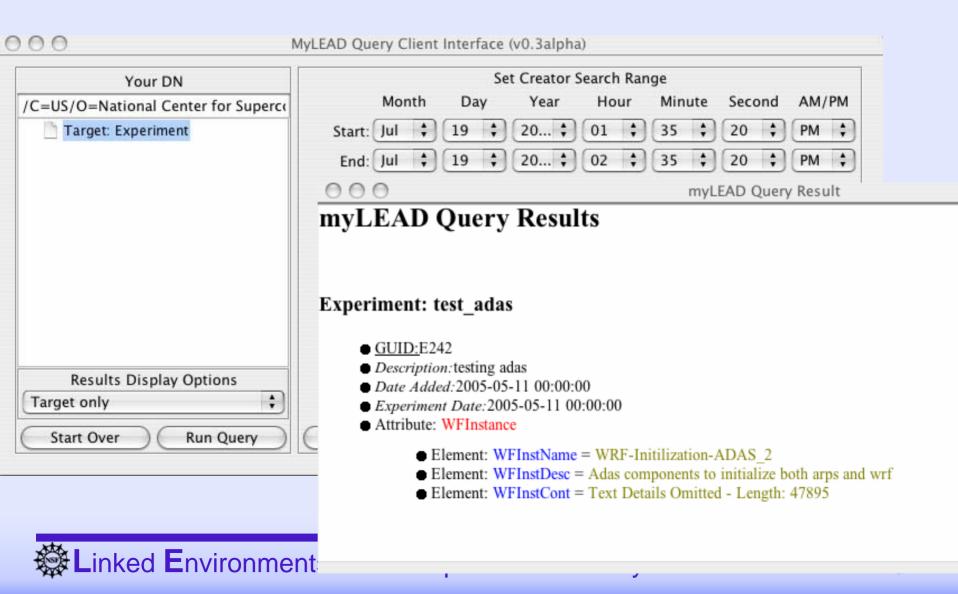


Log in and see your MyLEAD Space





Searching MyLEAD



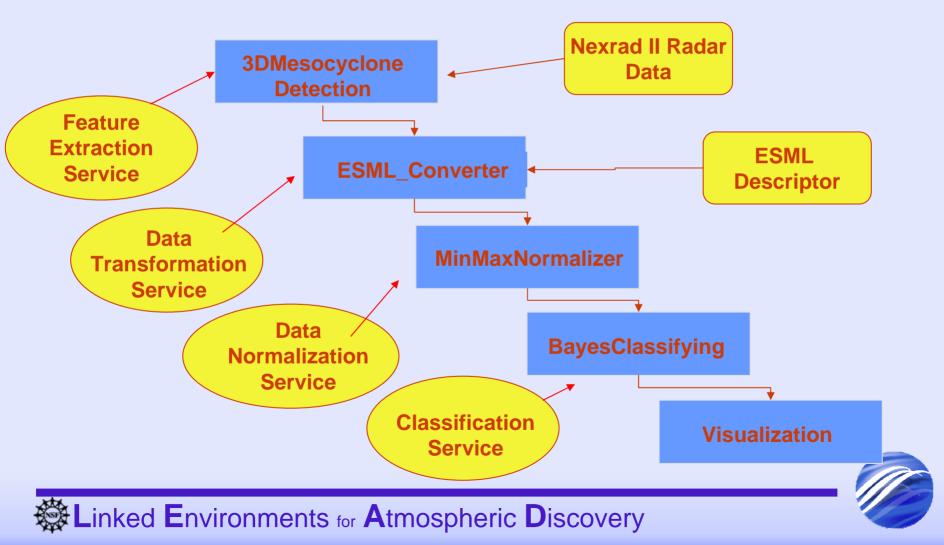
The Experiment Builder

- To review your previous experiments and create new ones
- Experiments are organized into projects
 - You can select an old one to look at,
 - Or create a new project or experiment.
 - Let's do a new experiment! (click "new")

	AD PORT		sored by the National Science	
LEAD Portal Home E	ducation and Outreach Help	Profile		
MyWorkspace Experim	ent Builder Generic Service Ti	polkit Security		
0		Experiment Builder I	Portlet	
User: Dennis Gannor	i	Project Experiments		ld Project
Experiment Name		Description	Last	Action Time
Simple_ADaM_test	This is a simple ADaM expe	eriment using a fixed data set.	2005-07-20T15:4	48:46.499-05:00
EZ_ADaM	an easy version of the AD	aM demo with canned input	2005-07-20T15:	48:47.278-05:00
1 F	AD Home EAO	Privacy	Terms of use	Contactus

Creating a workflow for Data Mining

Use ADaM services from UAH



Provide a name and description

- Next select an application from the dropdown list or create a new workflow.
- Once we have selected the app, we push "next" to add data.

LEAD	Portal Home	Education ar	d Outreach Help	Profile					
MyWo	MyWorkspace Experiment Builder Generic Service Toolkit Security								
Ø				Experime	ent Builder Portlet				
				Expe	riment Wizard				
			User: Sure	esh Marru	Project: template_workf	lows			
			Speci	fy a name, de	scription, and select workflow				
			Name:		est of the ADaM data minin				
			Description :		JAH. We are using the "AD the sample workflows.	aM			
				My Workflow					
				ADaM Dem Descripti Dataming					
						< Back	Next >	Cancel	Launch
	LEAD Home	9	FAQ	Privacy	Terms of use		Cor	ntact us	

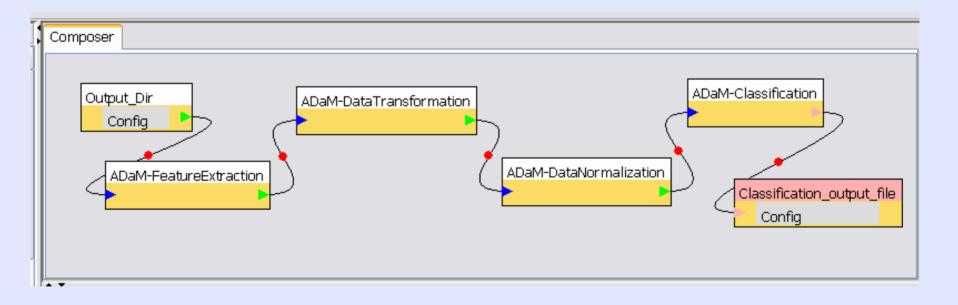
Composing the Workflow

- Graphical Composer
 - Standard drop-and-drag composer model (like Kepler and others)
 - Compiles Python or PBEL code

Workflow MyLead Component Monitor Hel Add Node Remove Node Connect/Disconnect Component List Composer 🖕 🗘 http://whitney.extreme.indiana.edu:22002/resourc 🔨 😑 🚇 http://www.extreme.indiana.edu/lead 😪 MyLEADNotificationTest Output_Dir 🏟 Terrain Preprocessor Confia 🏟 ARPS to WRF Data Interpolator 🏟 ARPS Plotting Program 🏟 3D Model Data Interpolator 伦 PostScript_to_Image_Converter ADaM-FeatureExtraction 🏟 WRF Forecasting Model 伦 WRF to ARPS Data Interpolator 🏟 ADaM-Applications 🏟 ADaM-Classification 📚 ADaM-DataNormalization 🏟 ADaM-DataTransformation Port Information Notification 📚 ADaM-FeatureExtraction ¥ 伦 WRF_Ensemble_Launcher -Selected Output Port-Selected Input Port > < **Component: Output Dir** Component: ADaM-FeatureExtraction Component Information Port: Parameter Port: Output Dir ^ Type: any Type: string Service: **Description**: This port can be connected to any type. Description: Enter the gridftp URL of the data file, the service will put the results in this location **ADaM-FeatureExtraction** Description: A set of methods for extracting features from raw data **Operation: Run**

Final Workflow

- Save it back to my lead
- Next we must bind the inputs to the workflow





Wizard understand the workflow requirements

e Education and O	PORTA	ofile	Sponsored by the National Scie		Logout Welcome, dennis gannon
periment Builder – G	eneric Service Toolk	it Security			
		Experiment	t Builder Portlet		
		Experin	nent Wizard		
	User	: Dennis Gannon	Project: ADaM Data Mining		
		Nam	e: building_a_workflow		
		Description	n: demo of composer		
		Workflov	w: ADaM test version		
		Select I	Data Stores		
	Ou	tput_Dir Enter the g	ridftp URL of the data file, the se	Select ervice will put the results in < Back Next >	this location Cancel Launch
LEAD Home	FAQ	Privacy	Terms of use	Contact us	
L inke	d Environ	nents for Atr	mospheric D iscove	ery	- 🧭

Select an output location

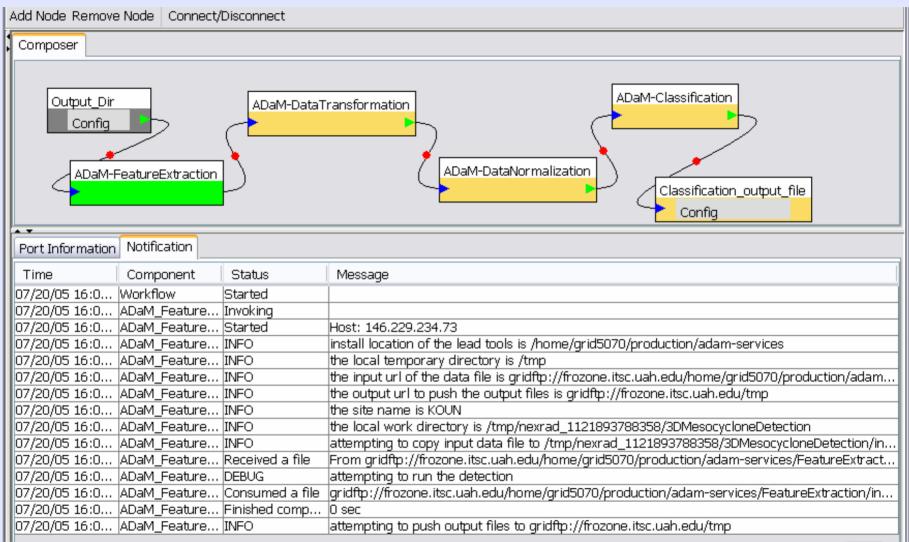
Back to Search Home		
Search for Storage Resource		Clear Find
Name	Protocol	GridFTP
Found 8 match(es)		
Query Results		
Select		
🗆 Chinkapin Tmp Storage		<u>Toggle XML</u> <u>Toggle Details</u>
☐ IU Testbed Public Space		Toggle XML Toggle Details
☐ IU Testbed Public Space		<u>Toggle XML</u> <u>Toggle Details</u>
□ UAH Testbed Public Space		<u>Toggle XML</u> <u>Toggle Details</u>
□ UNC Testbed Public Space		Toggle XML Toggle Details
🔲 Unidata Testbed Public Space		<u>Toggle XML</u> <u>Toggle Details</u>
□ OU Testbed Public Space		<u>Toggle XML</u> <u>Toggle Details</u>
□ NCSA Testbed Public Space		<u>Toggle XML</u> <u>Toggle Details</u>
Select		



Submitting the workflow

LEAD Portal Home Education an	d Outreach Help Profile					
MyWorkspace Experiment Builder	Generic Service Toolkit Security					
Ø	Experiment B	uilder Portlet				
	Experime	nt Wizard				
User: Suresh Marru		Project: template_workflo	ows			
	Name:	Datamining_Experiment				
	Description:		1 data mining sysem from UAH. We are from the sample workflows.			
	Workflow:	ADaM Demo				
	Data Stores:	Selected				
	Review ar	nd Submit				
	Please review the created workflow "Launch" button below. You may u or review in greater detail your sele	ise the "Back" button to c				
Start the workflow composer in monitoring mode prior to launching the workflow so that you can see a visualization of the workflow's progress as well as the workflow's notifications as they arrive.						
The RENCI monitoring applet provides additional additional information relating to the performance characteristics of the hosts where the workflow is executing.						
For debugging purposes only, you may want to start the notification viewer prior to launching the workflow so that you can see the workflow notifications as they arrive.						
			< Back Next > Cancel Launch			
LEAD Home	FAQ Privacy	Terms of use	Contact us			

Monitor results in real time

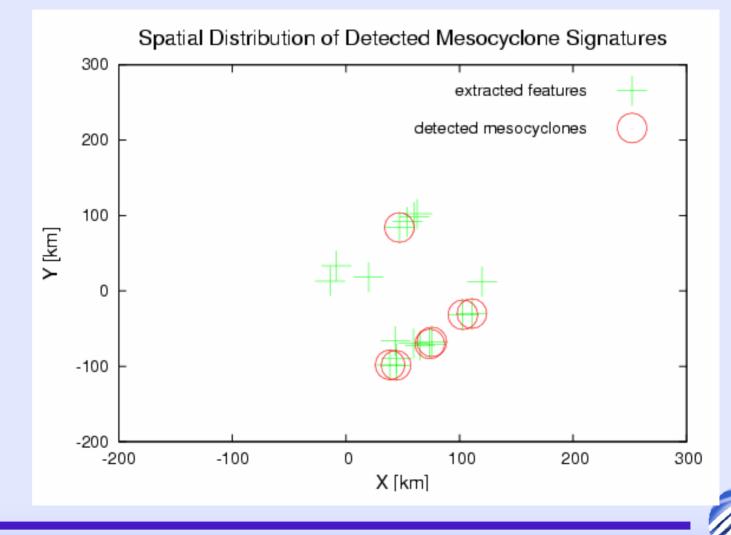




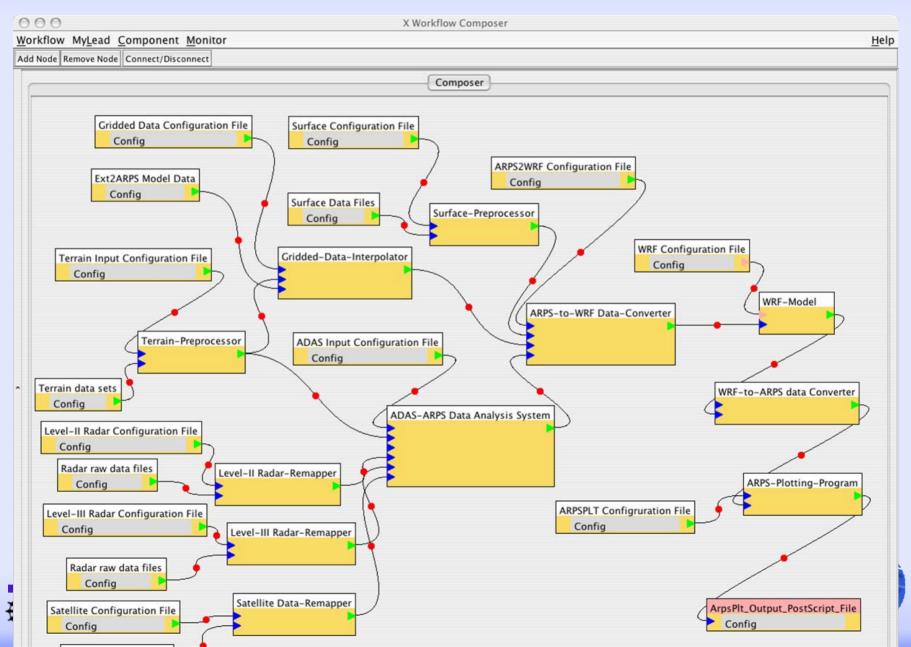
Check it out in MyLEAD

0	Experiment Builder Portlet		
User: Suresh Marru	Project: demo-d	lemo-run	
	Experiment Status Page		
Experiment Details			
Name: Datamining-demo	Last Notification Time: 2005-07-19T14:2	25:32.590-05:00	
Description: Datamining demo	Status: INFORMATION		
Workflow			
Name: ADaM Demo			
Description: Dataming services workflow			
Parame		Value	
Output_Dir	gridftp	p://frozone.itsc.uah.edu/tmp	
Notifications			
	Message	Status	Timestam
INVOKE SERVICE START		INFORMATIO	N 2005-07-19T14:25:32
APPLICATION START		INFORMATIO	N 2005-07-19T14:25:32
LEVEL: INFO MESSAGE: install location of the lead to	ols is /home/grid5070/production/adam-services	INFORMATIO	N 2005-07-19T14:25:32
LEVEL: INFO MESSAGE: the local temporary director	y is /tmp	INFORMATIO	N 2005-07-19T14:25:32
LEVEL: INFO MESSAGE: the input url of the data file		INFORMATIO	N 2005-07-19T14:25:32
gridftp://frozone.itsc.uah.edu/home/grid5070/produc			
LEVEL: INFO MESSAGE: the output url to push the o	utput files is gridftp://frozone.itsc.uah.edu/tmp		N 2005-07-19T14:25:32
LEVEL: INFO MESSAGE: the site name is KOUN			N 2005-07-19T14:25:32
	mp/nexrad_1121800848241/3DMesocycloneDetection		N 2005-07-19T14:25:32
	ta file to /tmp/nexrad_1121800848241/3DMesocyclon		N 2005-07-19T14:25:3
FILE RECEIVE DURATION Duration(Millis): 1452 FILE	UUID: testUUID	INFORMATION	N 2005-07-19T14:25:3
FILE CONSUMED FILE UUID: testUUID		INFORMATIO	N 2005-07-19T14:25:3
LEVEL: INFO MESSAGE: attempting to push output f	iles to gridftp://frozone.itsc.uah.edu/tmp	INFORMATIO	N 2005-07-19T14:25:3
FILE SEND DURATION FILE SIZE (Bytes): 1025 FILE	UUID: testUUID	INFORMATIO	N 2005-07-19T14:25:3
PUBLISH URL: http://chinkapin.cs.indiana.edu/data-	output/input.data_3Dmesocyclone_1121800848241.da	at INFORMATIO	N 2005-07-19T14:25:3
PUBLISH URL: http://chinkapin.cs.indiana.edu/data-			N 2005-07-19T14:25:3
		Сту	

Click on the output file to see visualization



Large workflows can be composed



Launch a Simulation

- Creating an experiment with the experiment builder. (Dan's big case.)
- Go to the experiment builder and
 - Look at your old experiments or create a new project
 - Name and describe it.
 - Select the application/workflow service
 - Bind your data set to it.
 - Where do you want the output to go?
 - Run it.
 - Monitor it.



Launching an Ensemble with detailed monitoring

 Lavanya Ramakrishnan from UNC will take us through this part.



Dynamic Adaptive LEAD System

- Meteorology goal
 - To provide timely and accurate forecasts using dynamic adaptation
- Computer Science goal
 - Map application requirements to resource capabilities
 - Adapt to weather as well as resource behavior

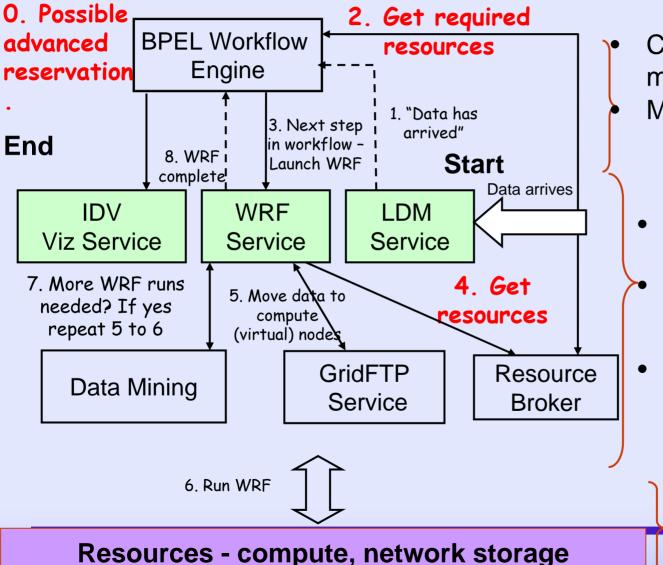
Need real time monitoring to make adaptation decisions

Adaptation Challenges

- LEAD system
 - Streaming data, respond to weather
- Proactive multilevel monitoring
 - Workflow, resource and application
- Intelligent control
 - Dynamic response to weather
 - scheduling based on priority and accuracy
 - Performance and reliability guarantees of the resources



Dynamic Workflows



- Correlate multilevel monitoring
- Multiple Decision Points

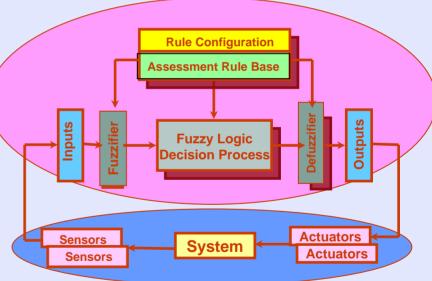
- Service Monitoring

 load of web service
- Application
 - application behavior on resources
- Resource
 - performance (e.g. CPU utilization)
 - reliability (e.g. temperature)



Leveraging Other Activities

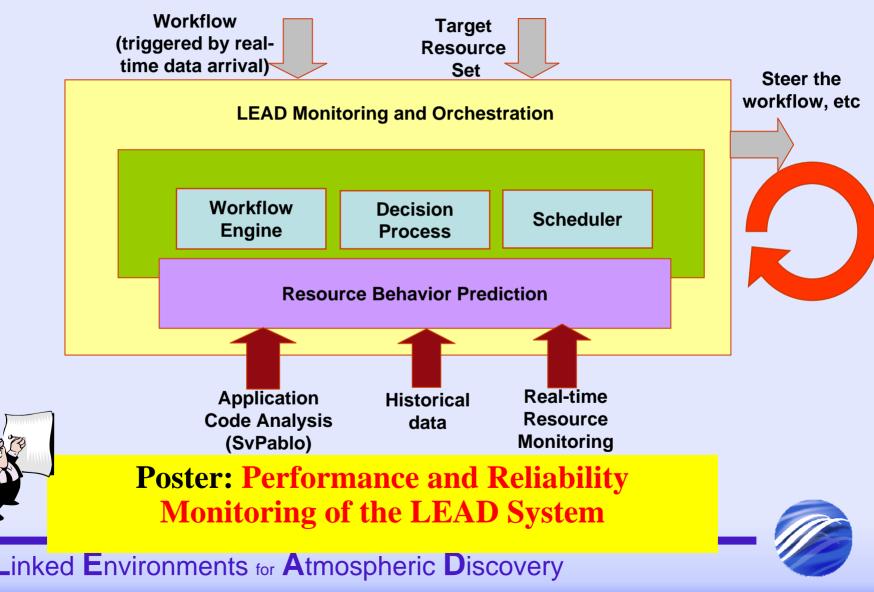
- Health Application Programming Interface (DOE LACSI)
 - standard interface for health monitoring
 - e.g. power, disk health, temperature
 - failure prediction from health data
- Network Weather Service- HAPI (NSF VGrADS)
 - leverage NMI tool infrastructure
 - integrate performance and failure indicator data
- Autopilot and SvPablo
 - control infrastructure to steer the workflow
 - source code performance data correlation



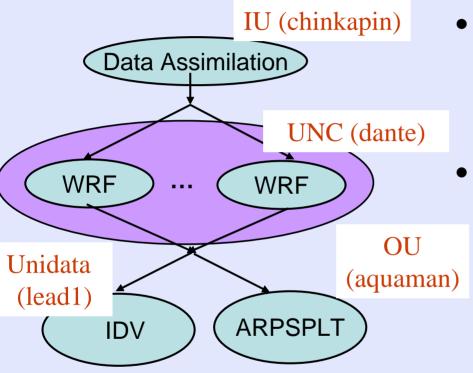
Autopilot Toolkit



LEAD Monitoring Architecture



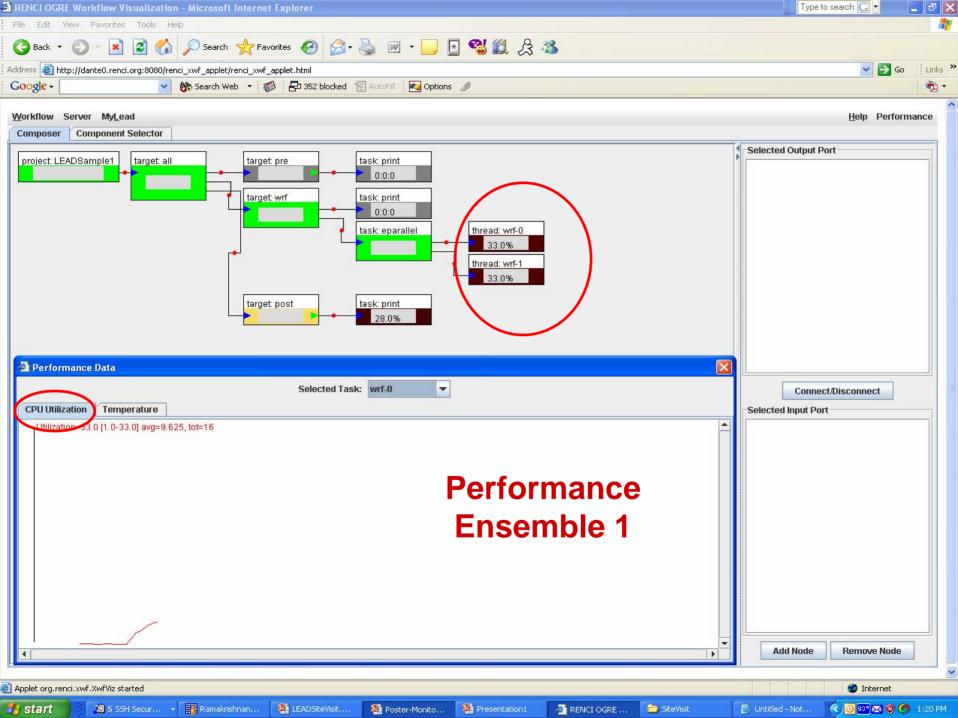
Experimental Ensemble

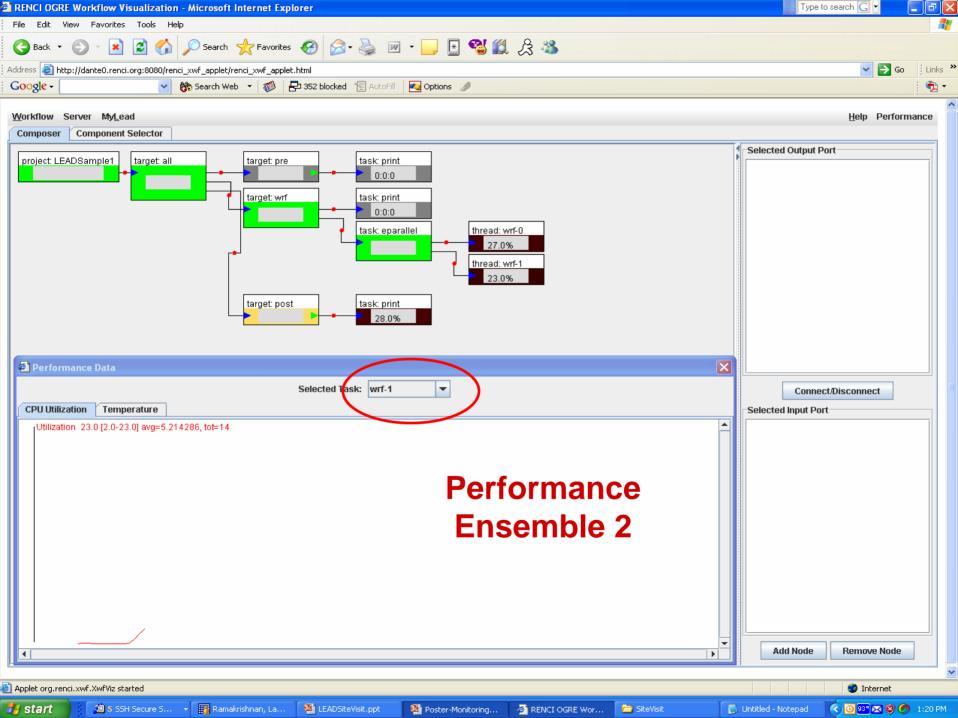


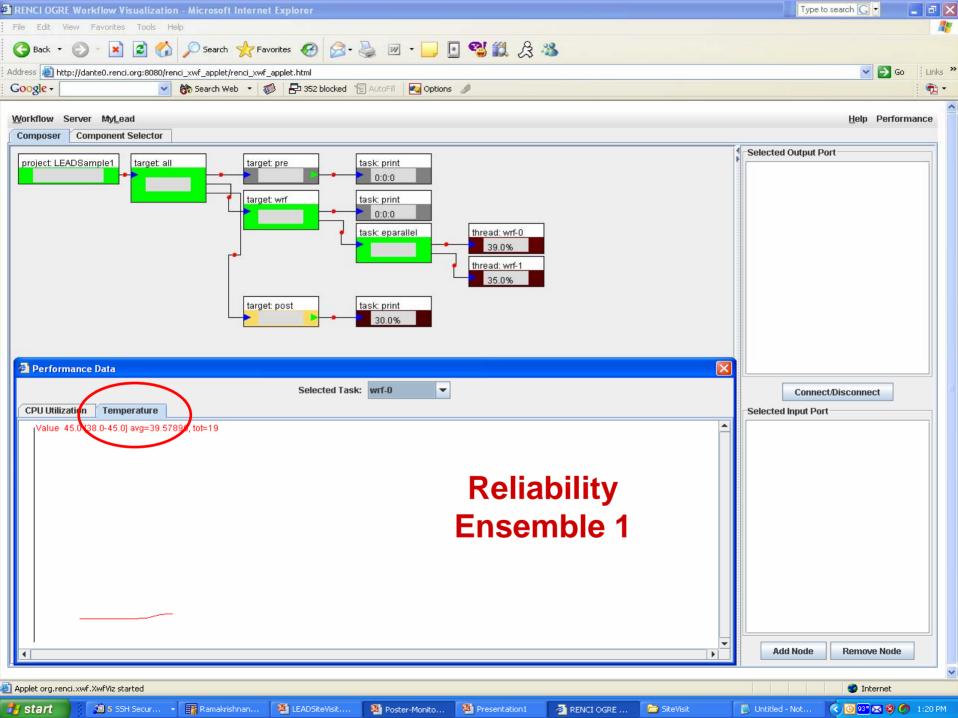
- Software components
 - IU Dynamic Workflow toolkit and Portal, NCSA OGRE, UNC Monitoring, OU WRF data
- Infrastructure
 - Testbed: IU, UNC, OU
 - Metrics
 - CPU utilization, temperature

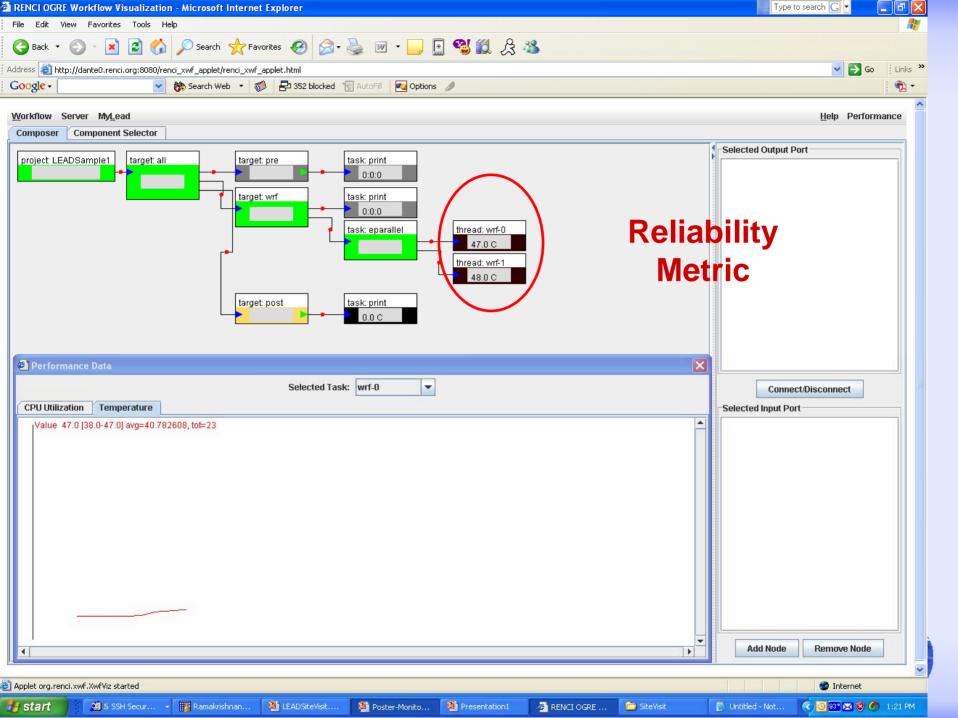
- Test Problem (ensemble = 2) on 16 nodes each
 - 60 secs timestep(34 mins),
 - 101 x 101 x 51, 27 km resolution forecast

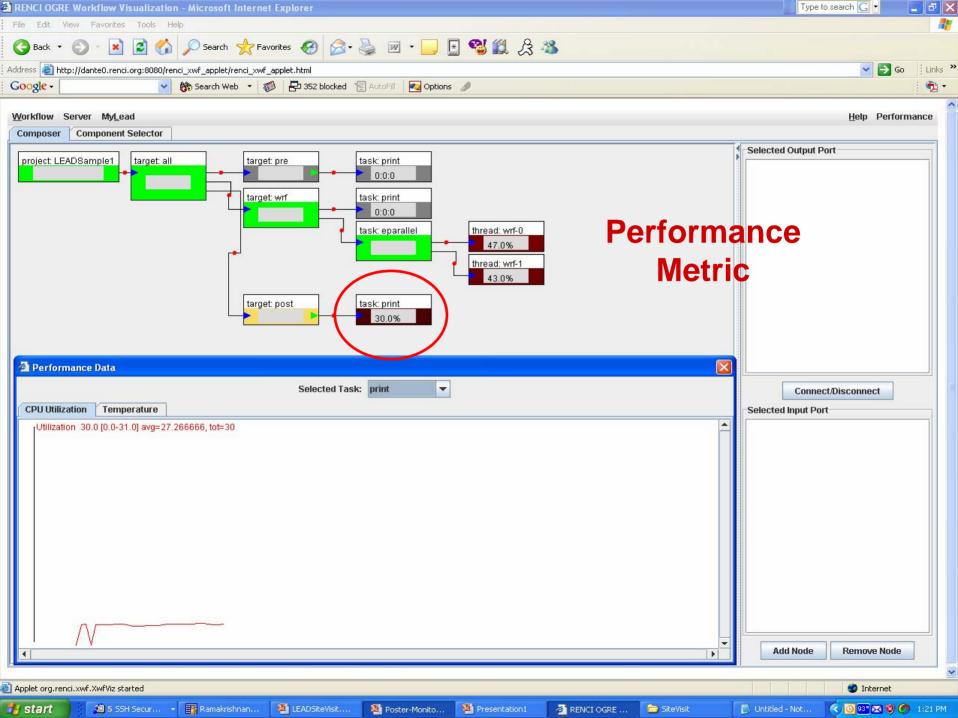


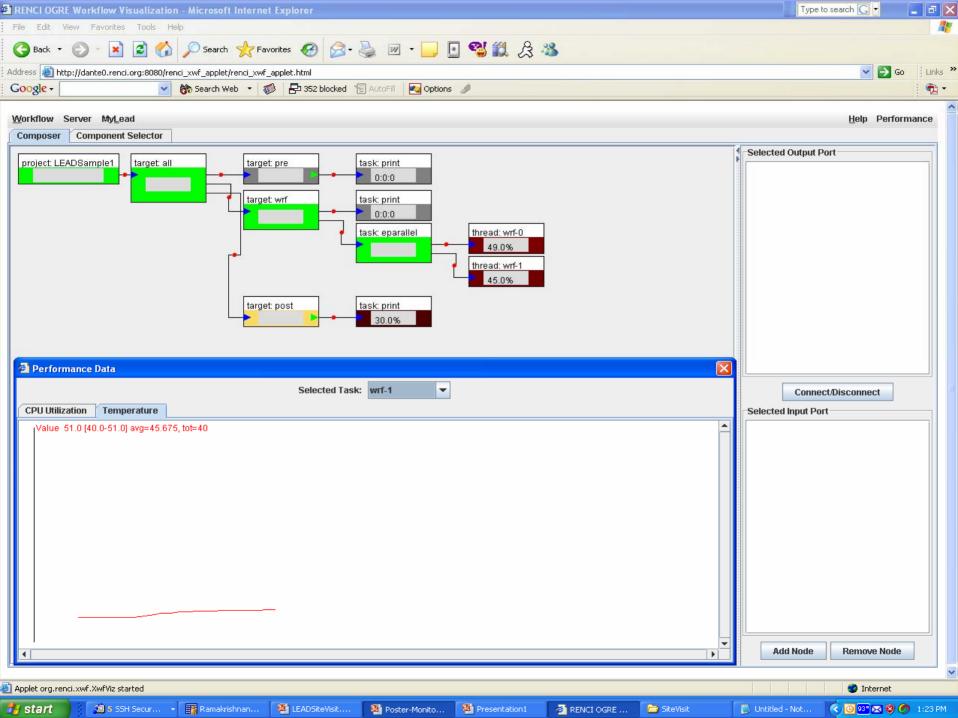


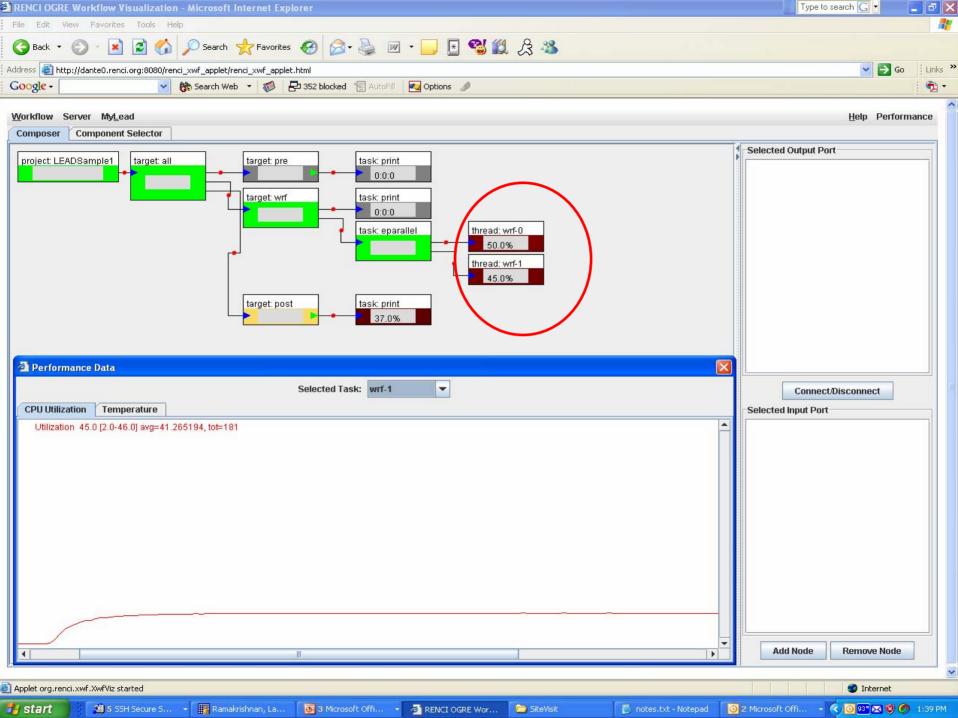


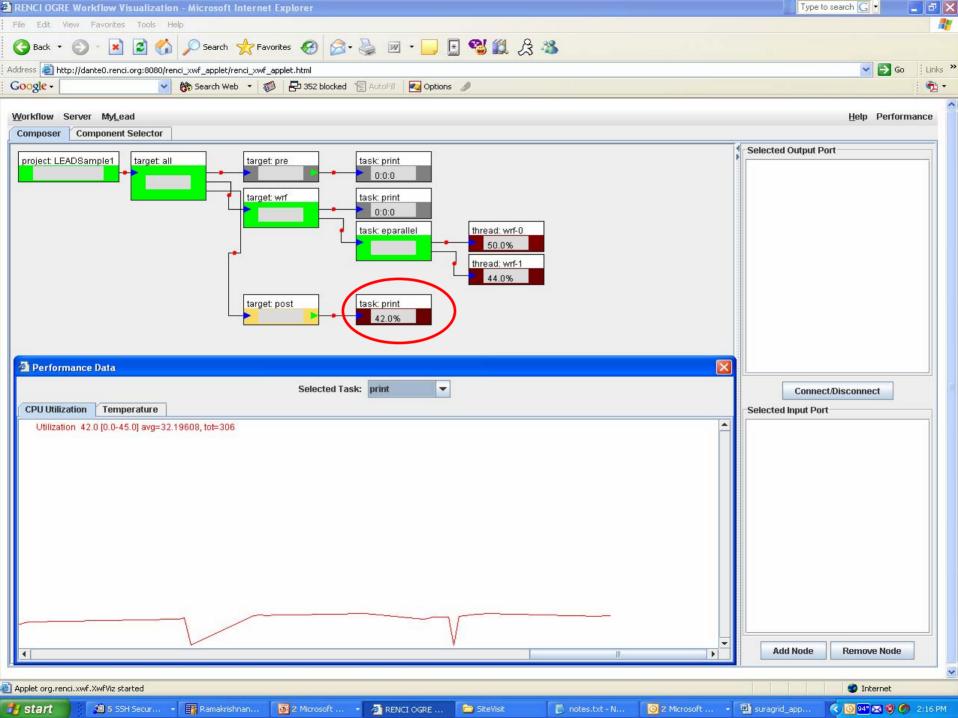


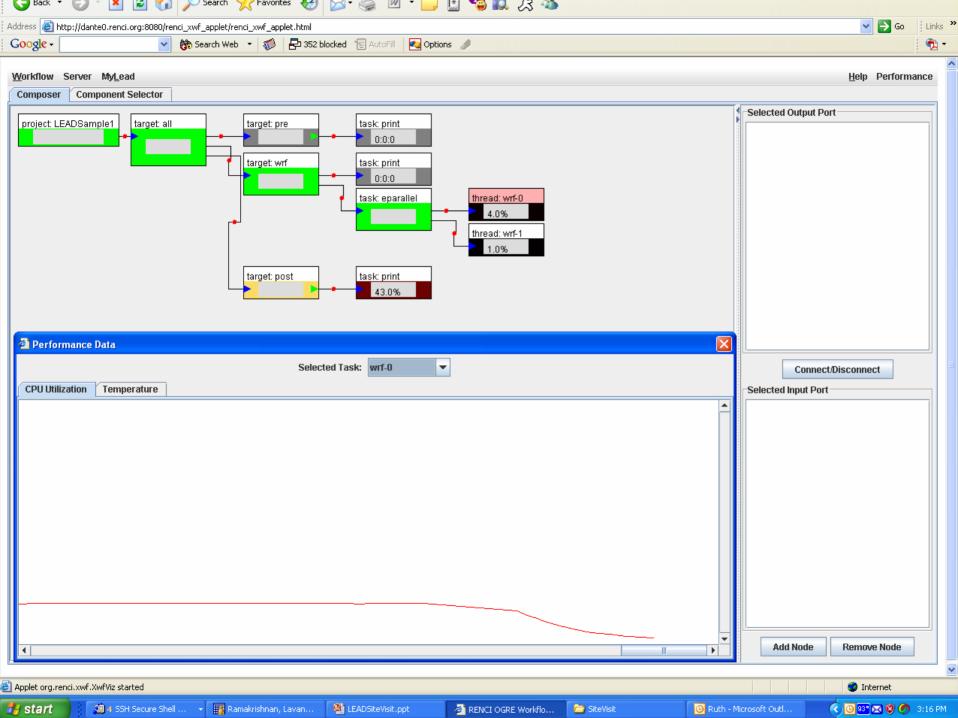


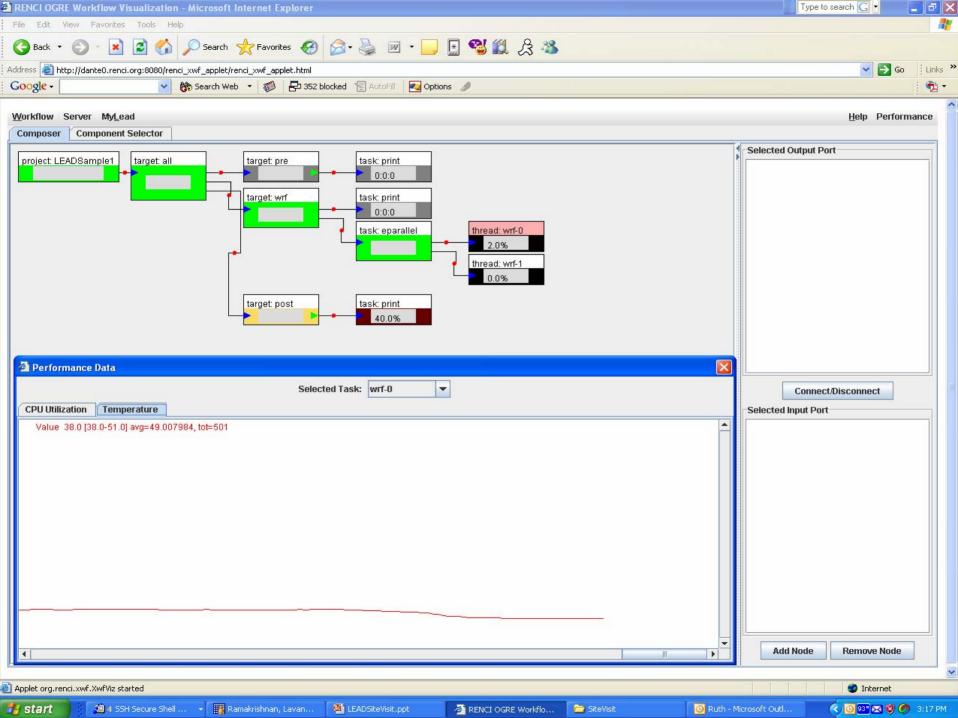












Research Issues for the Portal

- How to make it more focused on data discovery and analysis
 - and less about applications/workflow?
- How can we make it possible for the Education group to create and upload interactive learning scenarios?
 - What is the authorization model for using the scenarios?
 - How much can a student do on TeraGrid?
 - How do we conduct the best usability test?



Research issues: Workflow

- Understanding the dynamic case
 - Dynamic in both application structure and in the use of resources
- Extending the graphical language to describe dynamic workflow.
 - Is the current workflow language (BPEL) really sufficient for our dynamic workflow scenarios?
- How do we generate user interfaces for workflows
 - Experts can build workflow graphs, but basic users don't need to see the details. They just want to use them.

