

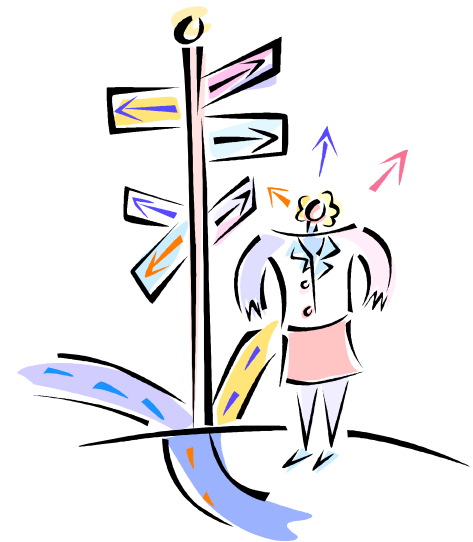
Fault Tolerance and Recovery for Grid Workflow Systems

Anirban Mandal
Gopi Kandaswamy
Dan Reed

VGrADS Workshop
April 2007

Presentation Outline

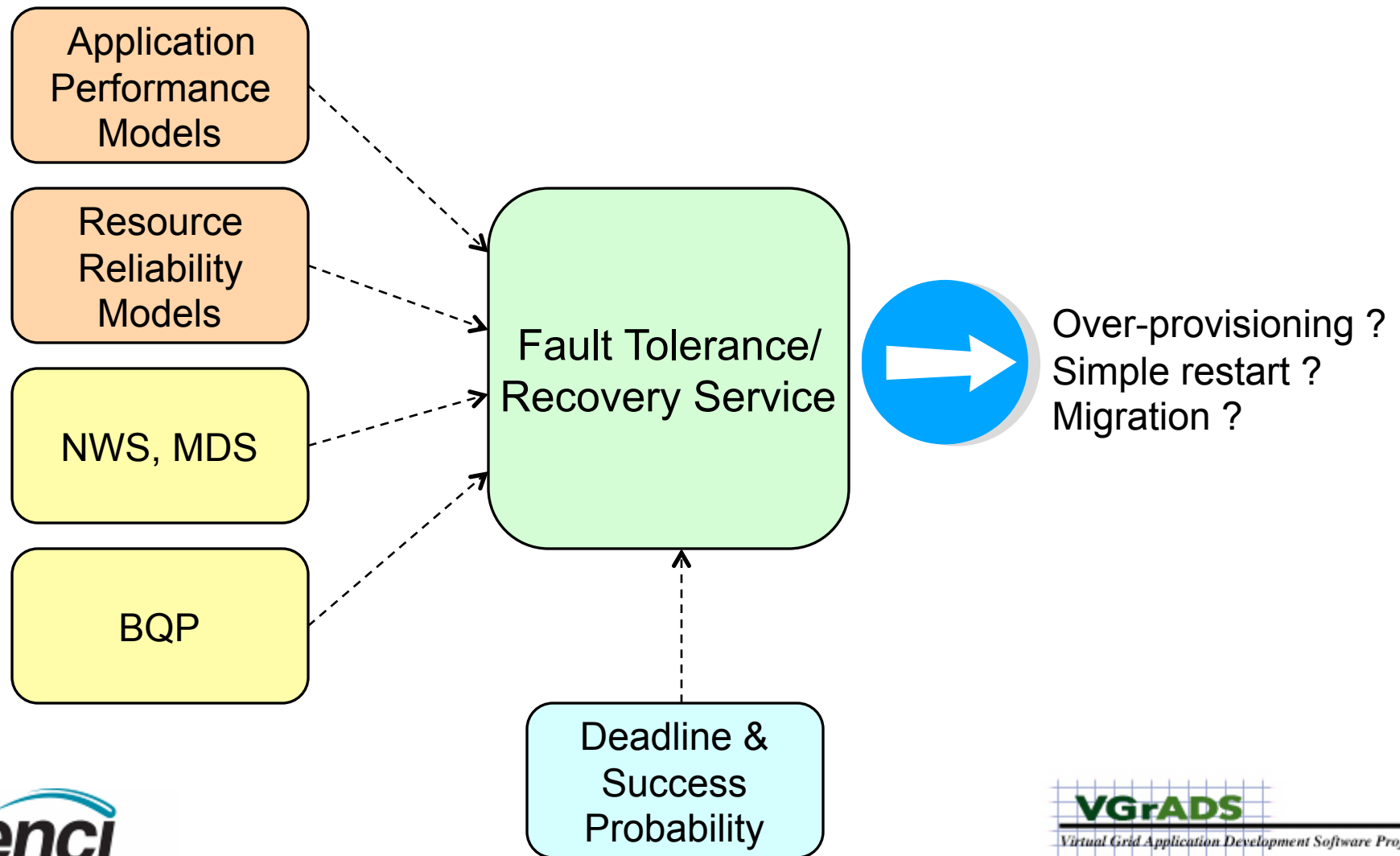
- Motivation and rationale
- Fault tolerance and recovery service
- Fault tolerance and recovery algorithms
 - over-provisioning, simple restart, migration
- Proposed next steps
- Demonstration
 - LEAD-VGrADS integration



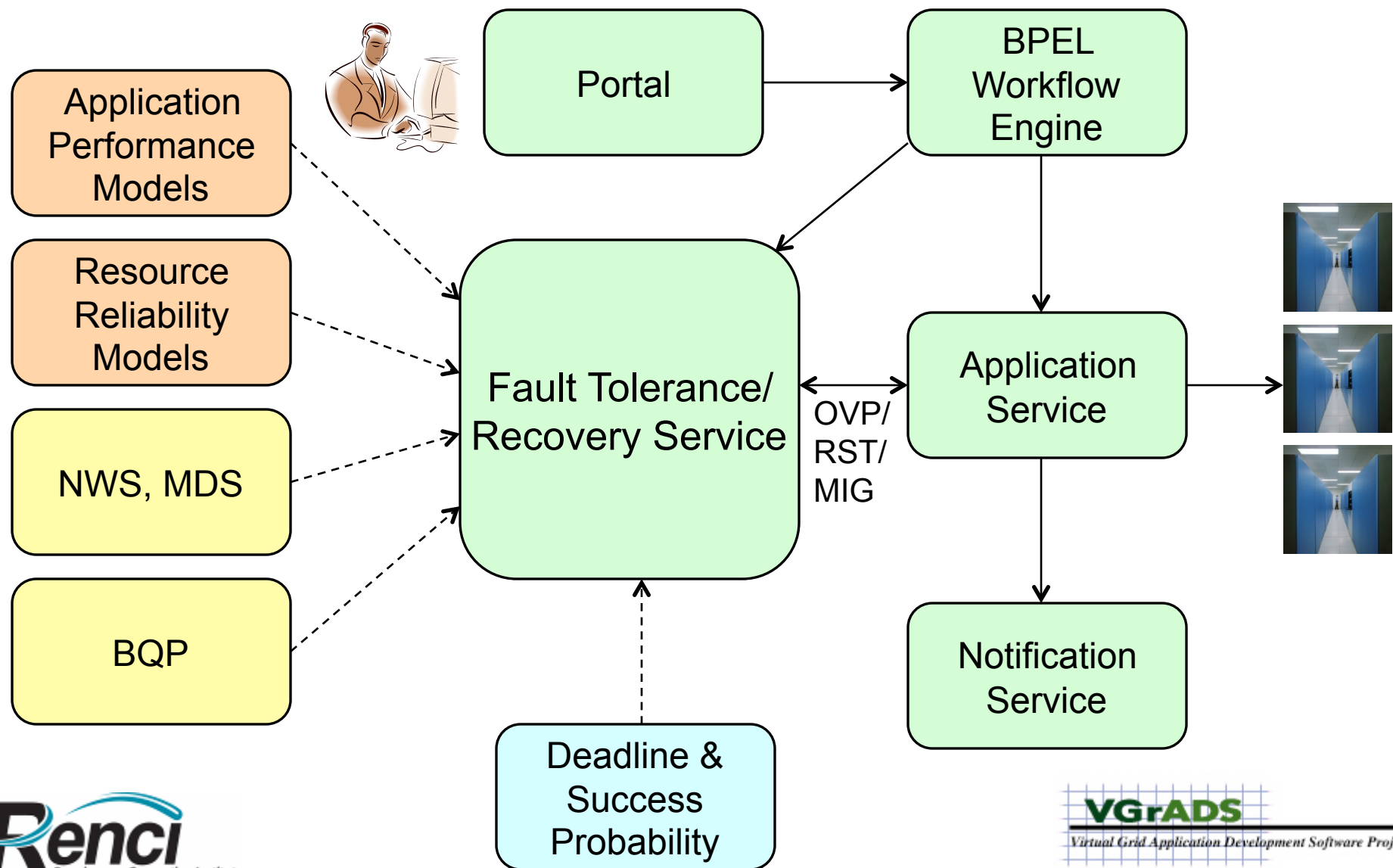
Motivation

- **Reliability and performance are related**
 - failure is the limiting case of poor performance
 - both involve measures of behavior over time
- **Large, complex workflows are sensitive to failures**
 - faults are the norm, rather than exceptions
 - distributed systems, services and resources
 - completion “guarantees” are problematic
 - workflow completion is probabilistic in the presence of faults
- **Many time-critical workflows are deadline driven**
 - severe weather events, disaster response, ...

Fault Tolerance/Recovery (FTR) Service



Simplified Architecture



FTR Algorithms

- **Notation**

p_i : one hour failure probability of resource i

h_i : expected execution cost of application on resource i

- queue wait time
- expected computation time
- expected communication time

x : required success probability

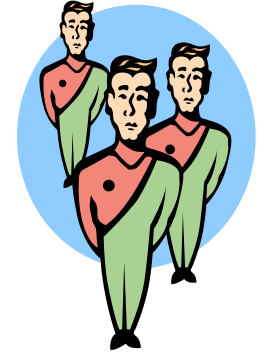
d : required deadline

- **If the reliability function is linear, probability of failure**

$$m_i = \max (p_i * h_i, 1)$$

- **Resource i represents (queue, #nodes) combinations**

Over Provisioning



- Find
 - degree and resources for over-provisioning
- Number of application copies
 - meet a deadline d with a success probability x
- Solve the following optimization problem

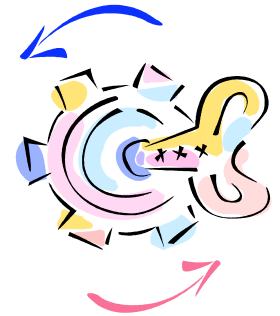
For given $[1..M]$ resources, find a partition $P = \{s_1, s_2 \dots s_n\}$ of $[1..M]$ such that

$$1 - m_{s_1} * m_{s_2} * \dots * m_{s_n} \geq x \wedge |P| \text{ is minimum } \wedge \min \{h_{s_1} \dots h_{s_n}\} \leq d$$

↑
Probability of failure

↑
Minimum number of resources meeting deadline

Simple Restart



- Among the current available resources ...
 - (universe – faulty resources)
- Find the best resource that
 - meets the deadline and success probability
- Mechanism can also be used for
 - runtime resource selection
- Two FTR mechanisms based on simple restart
 - restart with retries
 - restart if there is no progress

Migration



- Find the best migration path
 - more complex optimization problem
 - need to optimize orderings
- Solve the following optimization problem

Find a partition $P = \{s_1, s_2 \dots s_n\}$ and ordering $\{s_1 > s_2 > \dots s_n\}$ such that

$$1 - m_{s_1} * m_{s_2} * \dots m_{s_n} \geq x \wedge |P| \text{ is minimum } \wedge (t_{s_1} + \dots t_{s_n} + q) \leq d,$$

where $t_{s_i} = (p_{s_i} * h_{s_i} > 1) ? 1/p_{s_i} : h_{s_i}$ and q is the migration overhead

Proposed Next Steps

- **Modeling resource reliability**
 - draw data from job log files in the resources
 - maintain constant updates of reliability estimates
- **Reliability Information Service (viz. BQP)**
 - “What is the probability that a job, when submitted to a queue at a resource will die before completion because of a failure?”
- **Balancing multiple mechanisms**
 - given resource constraints
 - e.g., over-provisioning and migration
 - quantifying resource “wastage” with common metric

Proposed Next Steps (2)

- **Develop runtime rescheduling strategies**
 - **requires following changes in VGrADS software**
 - ability to re-acquire slots
 - ability to relinquish unnecessary slots
 - ability to accommodate FTR directives
 - ability to schedule remaining workflow
 - **requires following change in workflow engine**
 - ability to run a re-scheduled workflow

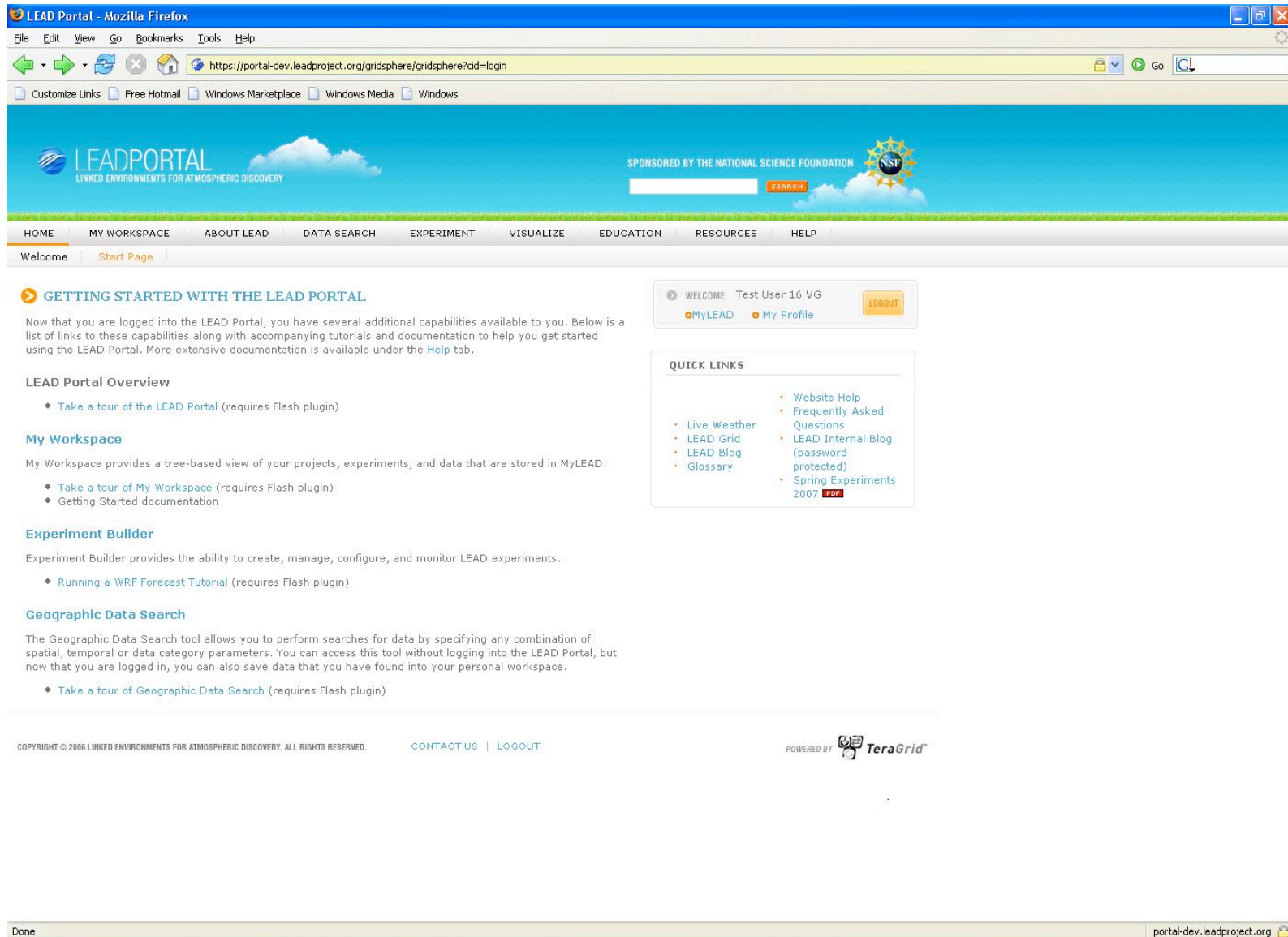
Proposed Next Steps (3)

- **Deadline estimates**
 - given a deadline for the entire workflow
 - find deadlines for individual applications
 - requires critical-path analysis

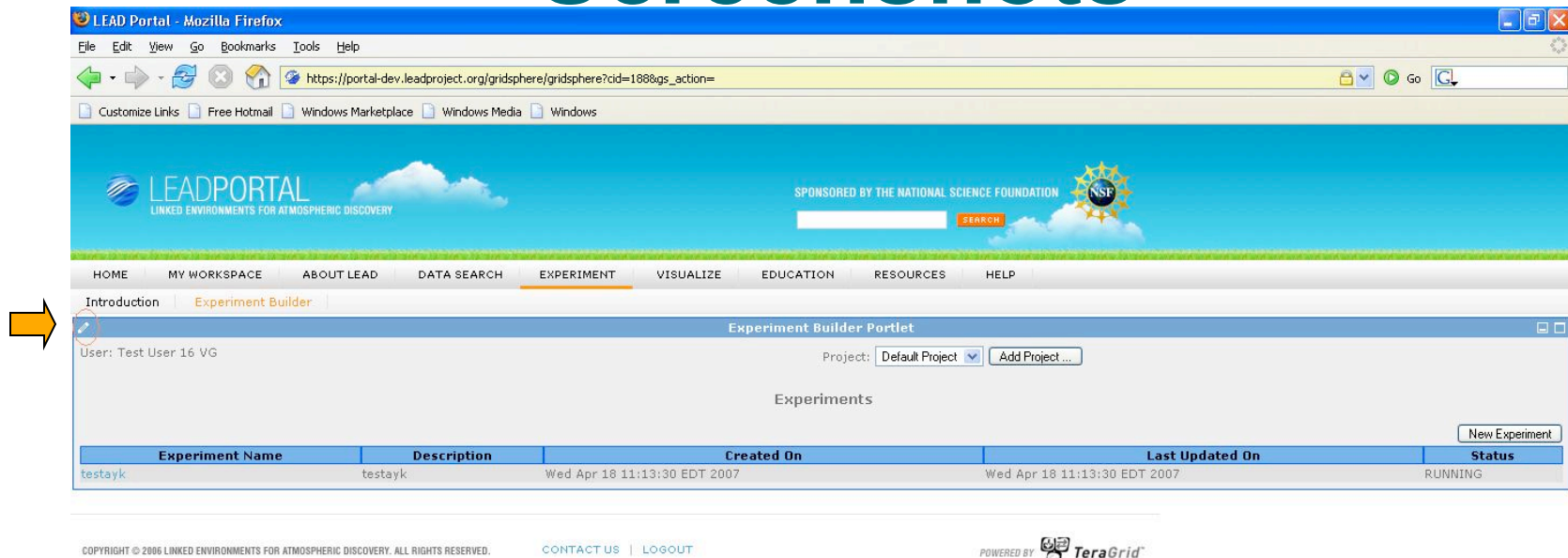
FTR Demonstration

- **Launch workflow from the LEAD portal**
 - build new experiment
 - select forecast region and workflow
 - fetch data
- **FTR service manages application execution by**
 - choosing the fault-tolerance mode (OVP, RST)
 - invoking application service(s) as per current mode
 - monitoring for failures
- **Workflow composer shows workflow progress**
 - status of application currently running
 - workflow execution status via notifications etc.

Screenshots



Screenshots



LEAD Portal - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://portal-dev.leadproject.org/gridsphere/gridsphere?cid=188&gs_action=

Customize Links Free Hotmail Windows Marketplace Windows Media Windows

LEADPORTAL
LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY

SPONSORED BY THE NATIONAL SCIENCE FOUNDATION

SEARCH

HOME MY WORKSPACE ABOUT LEAD DATA SEARCH EXPERIMENT VISUALIZE EDUCATION RESOURCES HELP

Introduction Experiment Builder

Experiment Builder Portlet

User: Test User 16 VG Project: Default Project Add Project ...

Experiments

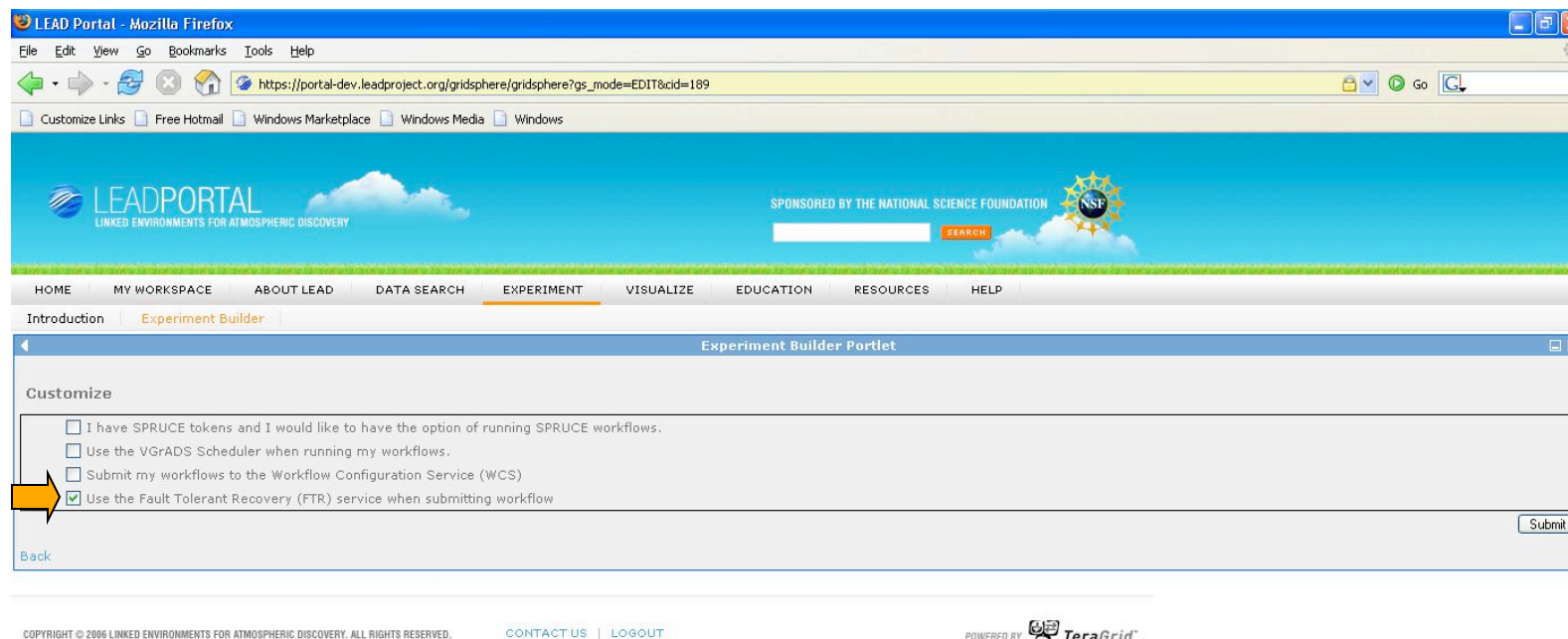
Experiment Name	Description	Created On	Last Updated On	Status
testayk	testayk	Wed Apr 18 11:13:30 EDT 2007	Wed Apr 18 11:13:30 EDT 2007	RUNNING

New Experiment

COPYRIGHT © 2006 LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY. ALL RIGHTS RESERVED. CONTACT US LOGOUT

POWERED BY TeraGrid

Screenshots



Screenshots

The screenshot shows the LEAD Portal web application in a Mozilla Firefox browser window. The address bar displays the URL: https://portal-dev.leadproject.org/gridsphere/gridsphere?cid=188&gs_action=. The page header includes the LEADPORTAL logo with the tagline "LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY" and a search bar. The navigation menu contains links for HOME, MY WORKSPACE, ABOUT LEAD, DATA SEARCH, EXPERIMENT (highlighted), VISUALIZE, EDUCATION, RESOURCES, and HELP. Below the navigation menu, there are sub-links for Introduction and Experiment Builder. The main content area is titled "Experiment Builder Portlet" and contains an "Experiment Wizard" section. The wizard prompts the user to "Specify a name, description, and select workflow". The "Name" field is filled with "test2" and the "Description" field is empty. Below these fields, there are two tabs: "My Workflows (0)" and "Sample Workflows (5)". The "Sample Workflows (5)" tab is selected, showing a workflow diagram for "NAM Initialized WRF V2.2 Forecast". The diagram consists of several interconnected components: "CrossCuttingConfigurations" (Config), "Terrain_V5_2_7_Preprocessor", "WRF_Static_V5_2_7_Preprocessor", "NAMInitialData" (Config), "NAM_Initial_Conditions_V5_2_7_Interpolator", "NAM_Lateral_Boundary_V5_2_7_Interpolator", "ARPS2WRF_V5_2_7_Interpolator", "WRF_Forecast_V2_2_Model", and "WRF_Output_Files" (Config). The workflow starts with "CrossCuttingConfigurations" feeding into "Terrain_V5_2_7_Preprocessor" and "WRF_Static_V5_2_7_Preprocessor". "NAMInitialData" feeds into "NAM_Initial_Conditions_V5_2_7_Interpolator". "NAM_Lateral_Boundary_V5_2_7_Interpolator" feeds into "ARPS2WRF_V5_2_7_Interpolator". "ARPS2WRF_V5_2_7_Interpolator" feeds into "WRF_Forecast_V2_2_Model", which finally outputs to "WRF_Output_Files".

Screenshots

The screenshot shows the LEAD Portal - Mozilla Firefox browser window. The address bar displays the URL: https://portal-dev.leadproject.org/gridsphere/gridsphere?cid=188&gs_action=. The page header includes the LEADPORTAL logo with the tagline "LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY" and a sponsorship notice from the NATIONAL SCIENCE FOUNDATION (NSF). The navigation menu includes links for HOME, MY WORKSPACE, ABOUT LEAD, DATA SEARCH, EXPERIMENT (highlighted), VISUALIZE, EDUCATION, RESOURCES, and HELP. Below the navigation menu, the "Experiment Builder Portlet" is displayed, featuring an "Experiment Wizard" section with the following details: User: Test User 16 VG, Project: Default Project, Name: test2, Description: test workflow, and Workflow: NAM Initialized WRF V2.2 Forecast. The "Model Domain Configuration" section includes a "Region Type Selection" with radio buttons for Regional 5km, Regional 20km (selected), and CONUS 20km. The "Forecast Start Time" section shows "Now (in other words, run a forecast using the most recent data available)" selected, with a "Start Date" of 2007/04/20 and a "Start Hour" of 02. The "Forecast Duration" is set to 6 hours. A map of North America is shown with a green box indicating the model domain. The "Forecast Domain" section displays the center latitude as 38.0000 and center longitude as -98.0000. The "Map Layers" section includes checkboxes for Radar Sites, Google Streets (on top), and Nexrad Doppler (checked).

Questions ?