Experiences with use of GRAM in LEAD Project

Suresh Marru
School of Informatics, Indiana University

Open Source Grid & Cluster Conference
Oakland, CA
May 14th 2008

The LEAD Project

L I N K E D
ENVIRONMENTS
FOR ATMOSPHERIC
D I S C O V E R Y

K E D
NMENTS
OSPHERIC
OVERY

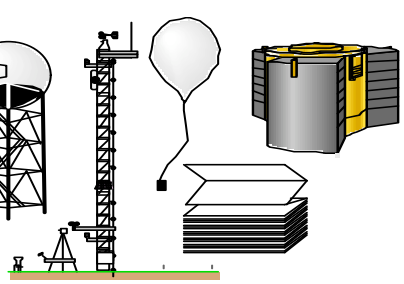
http://leadproject.org

Tera

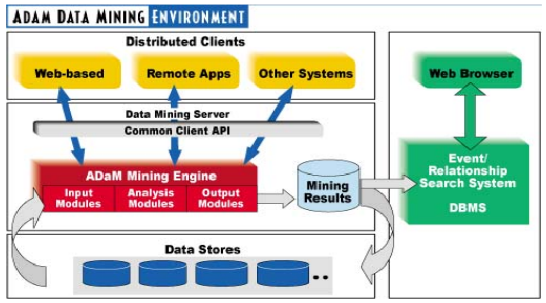
Two Principal Goals of LEAD

- **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

Example: “Optimal” Weather Prediction Using Dynamic Adaptivity



Streaming Observations



Data Mining

Storms Forming

Forecast Model

Instrument Steering

Refine forecast grid

On-Demand Grid Computing

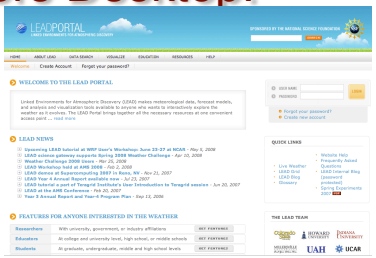
The TeraGrid

The US National Supercomputer Grid

- CyberInfrastructure composed of a set of resources (compute and data) that provide common services for
  - Wide area data management.
  - Single sign-on user authentication.
  - Distributed Job scheduling and management.

The Architecture of LEAD Gateway Services

Users Desktop.



Portal Server

Gateway Services

Proxy Certificate Server / vault

Application Deployment

Application Workflow

User Metadata Catalog

Application Events

Resource Broker

App. Resource catalogs

Replica Mgmt

Core Grid Services

Security Services

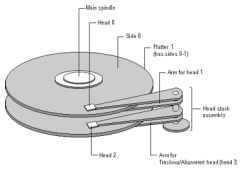
Information Services

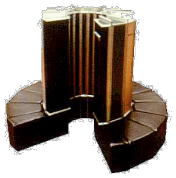
Resource Management

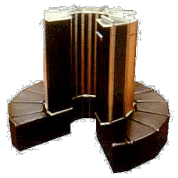
Execution Management

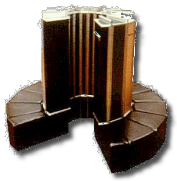
Data Services

Physical Resource Layer

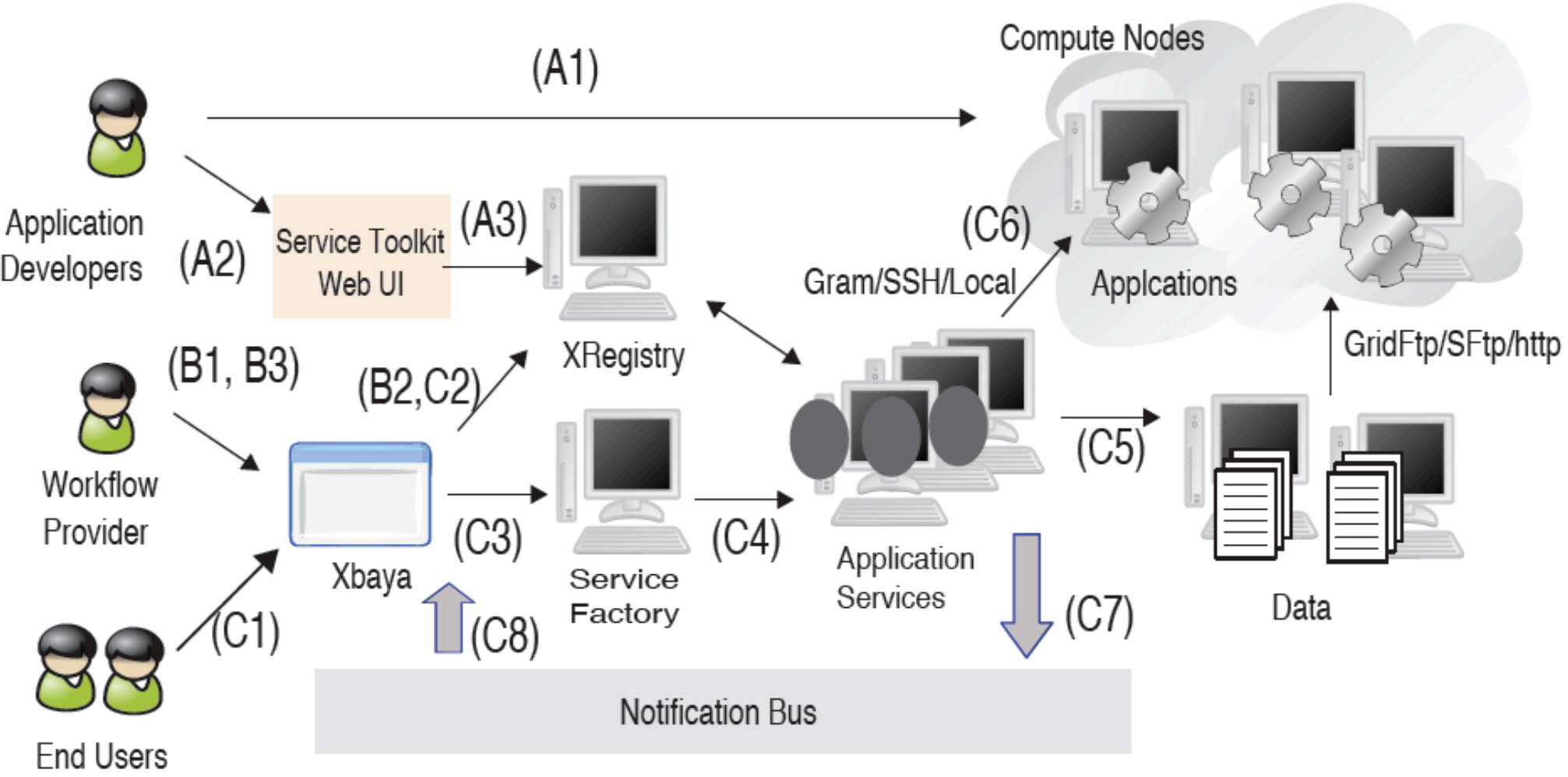








EAD Workflow Architecture using GRAM



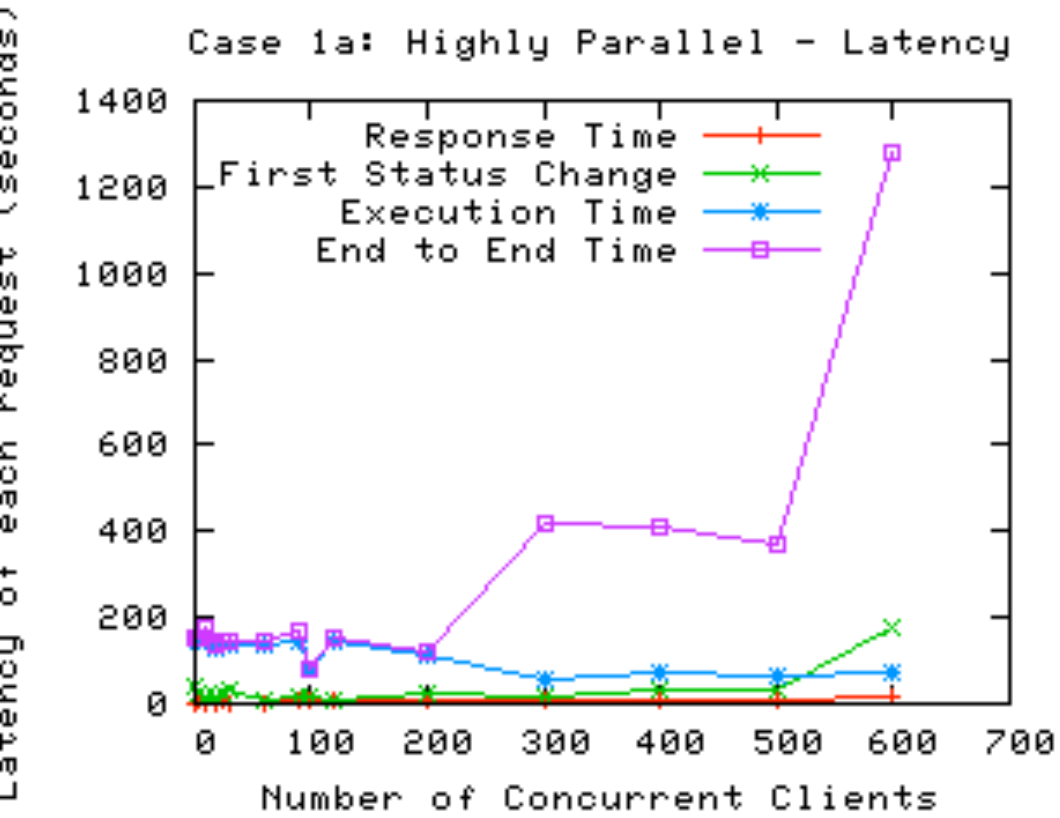
LEAD Enabling the Community

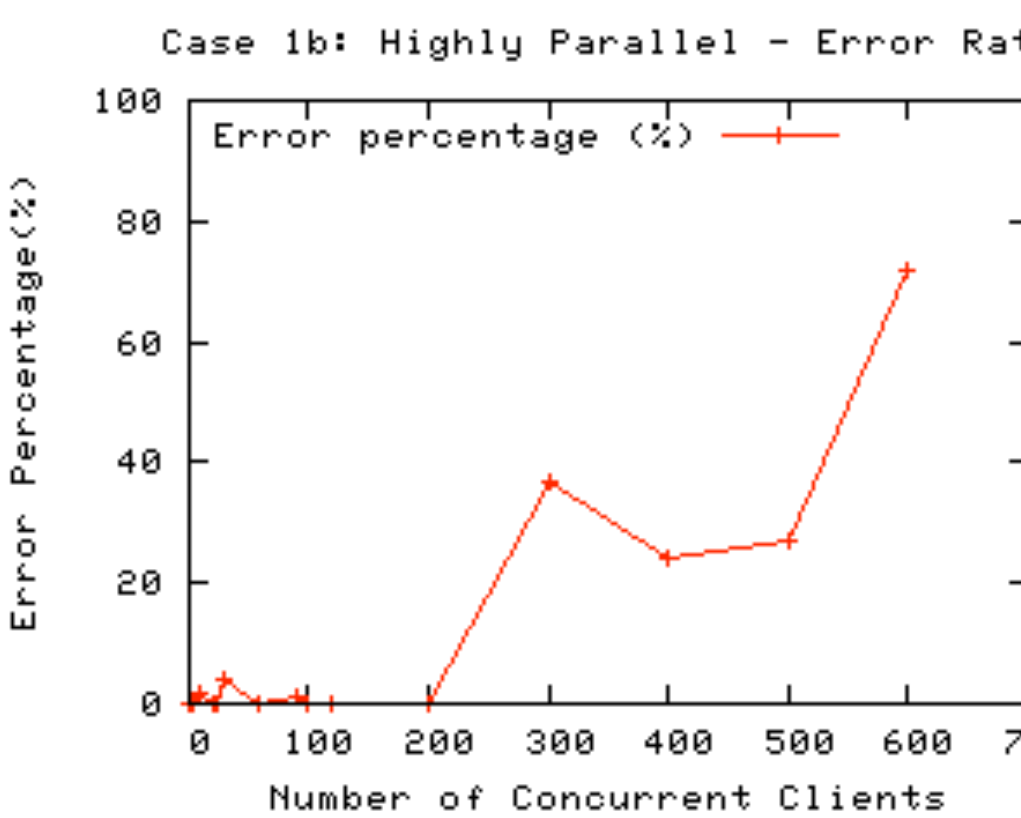
- **National Collegiate Forecast Challenge**
  - Participated by undergraduate, graduate and PhD as well as faculty and staff from atmospheric science departments and research groups.
  - Each participant forecasts the maximum and minimum temperatures, precipitation, and maximum wind speeds for select North American cities.
  - 67 participating institutions

GRAM LOAD Testing Motivation

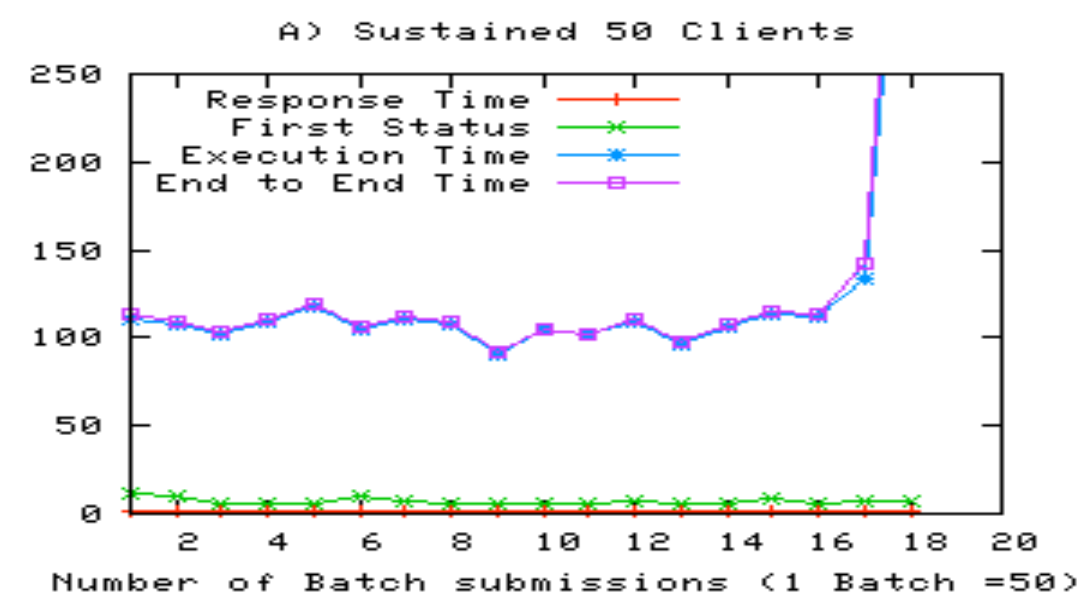
- Weather challenge users submitting 2 workflows for each forecast period accounting to 134 workflows in 4 hours.
- Each workflow comprises of 8 jobs making in 1072 jobs in 4 hours.
- GRAM services crashing and misbehaving under load.
- Coordinated with GRAM TEAM for testing and rapid deployment of bug fixes & patches and to understand the server load limit.

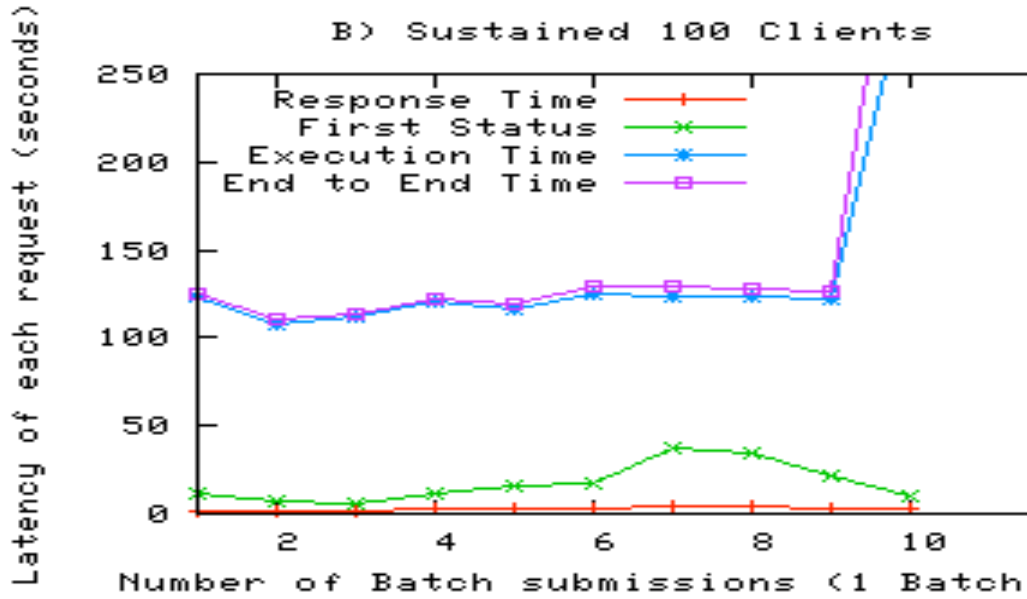
GRAM (4.0.6+) Scalability Testing

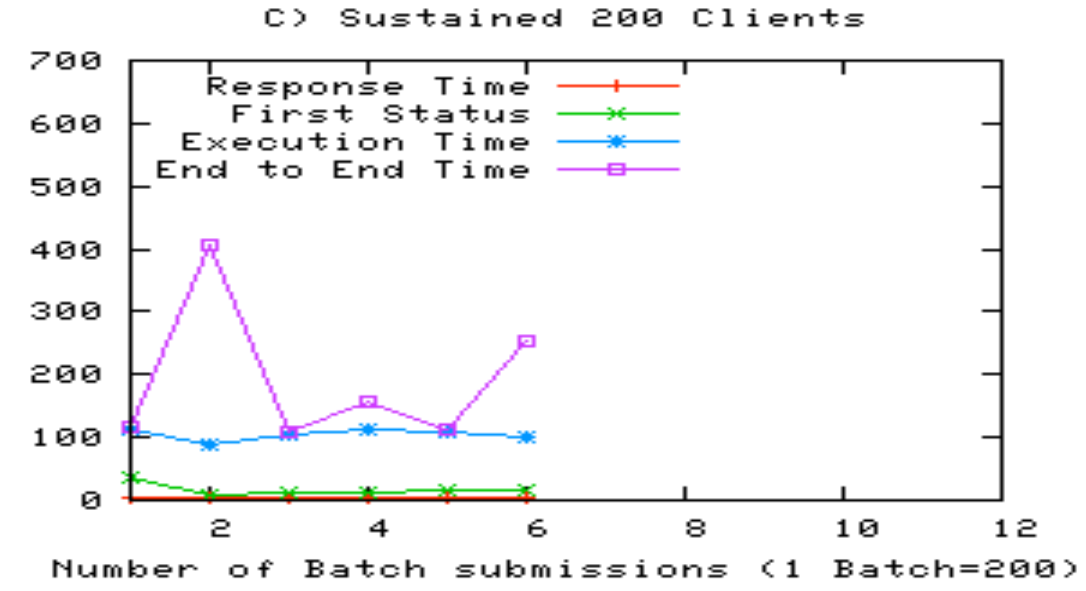


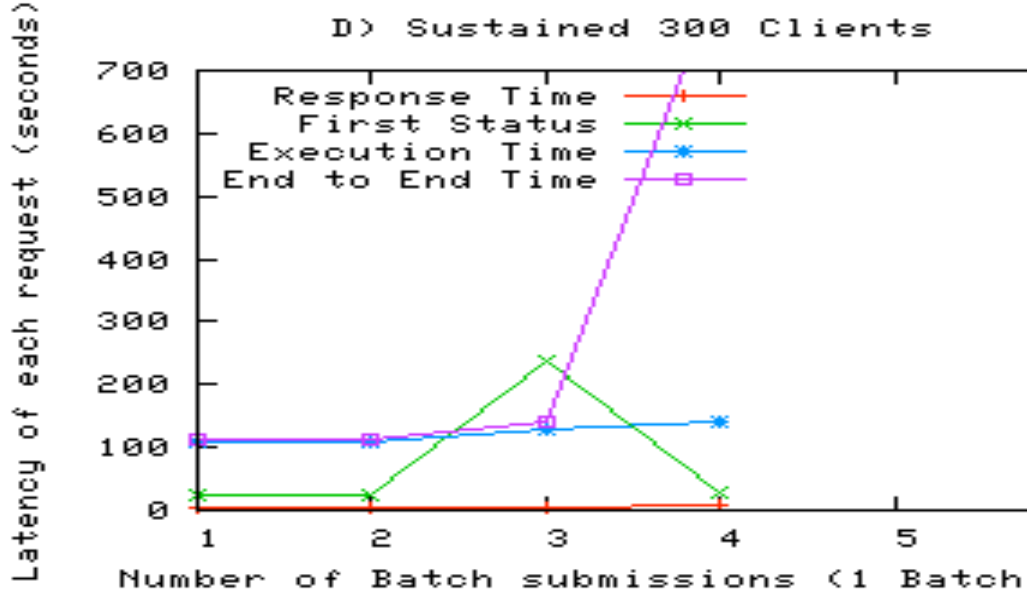


GRAM (4.0.6+) Reliability Testing









Summary of LEAD Experiences with GRA

- GRAM's tight coupling with Grid Security Infrastructure provides seamless integration for Gateway Architectures.
- Provides uniform interface and job description with multiple cluster resource managers.
- Web Service Interface to Job Management.
- Scalability and Reliability is improving but still needs lot more improvement to avoid the bottle neck in front of clusters with thousands of cores.