

PetaShare All Hands Meeting

State of the Union

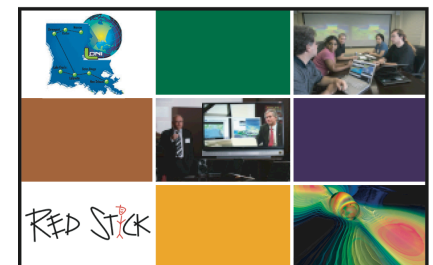
Tevfik Kosar

Center for Computation & Technology
Louisiana State University

March 3, 2008



CENTER FOR COMPUTATION
& TECHNOLOGY

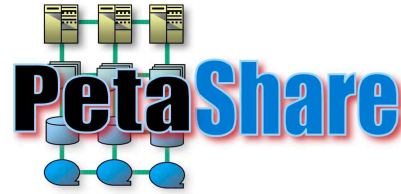


CCT: Center for Computation & Technology @ LSU

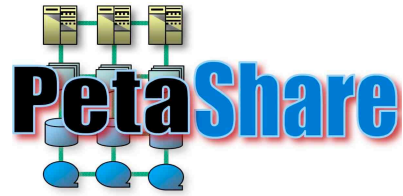
Roadmap

- ★ PetaShare Goals
- ★ Status of Development Efforts
- ★ Science Driver Presentations
- ★ Updates from Site Leads
- ★ Future Directions

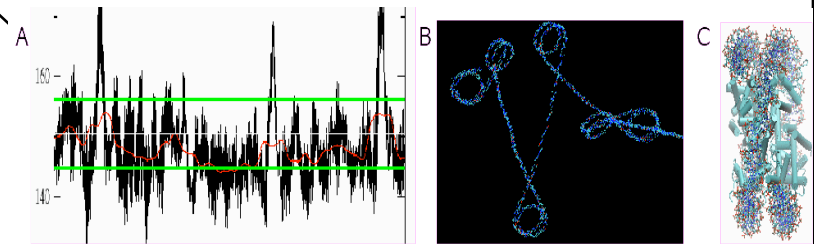
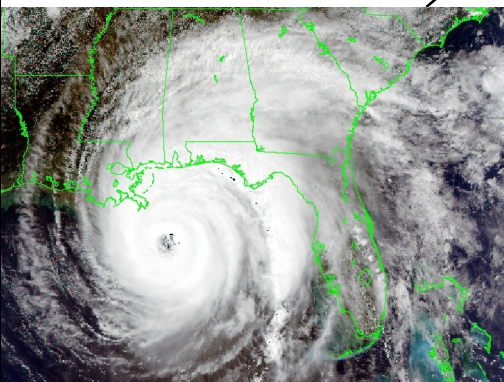
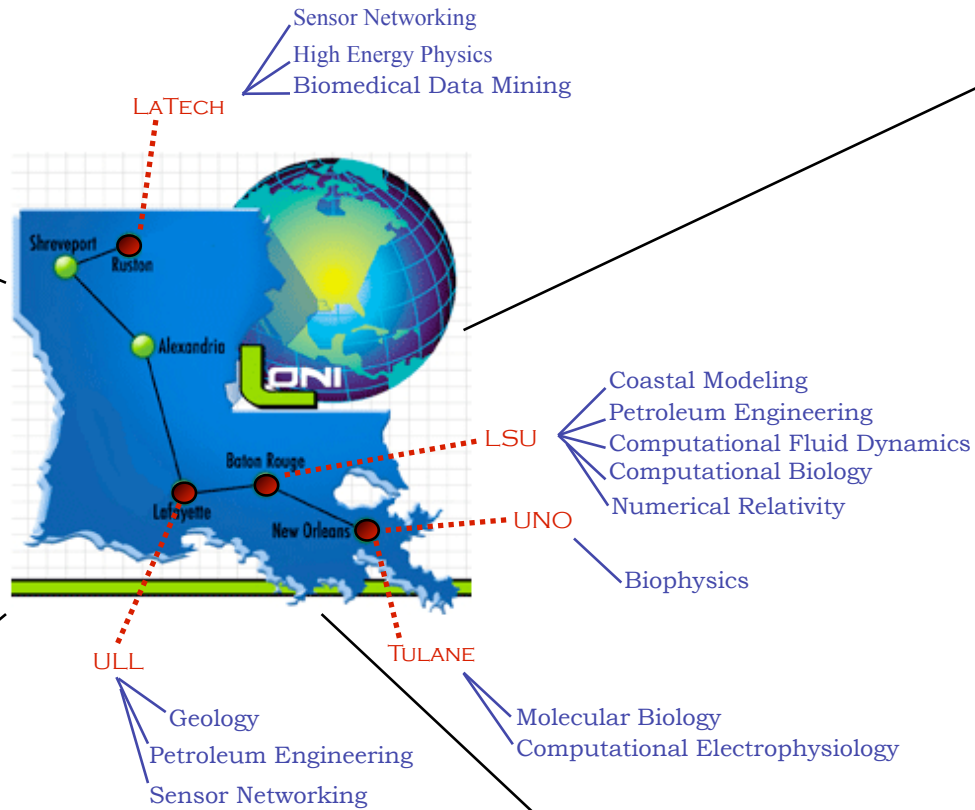
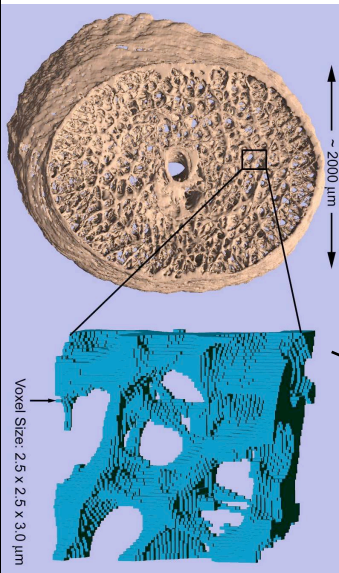




- **Goal:** enable domain scientists to focus on their primary research problem, assured that the underlying infrastructure will manage the low-level data handling issues.
- **Novel approach:** treat data storage resources and the tasks related to data access as first class entities just like computational resources and compute tasks.
- **Key technologies** being developed: **data-aware storage systems**, **data-aware schedulers** (i.e. Stork), and **cross-domain meta-data scheme**.
- **Provides** and additional 250TB disk, and 400TB tape storage (and access to national storage facilities)



Participating institutions in the PetaShare project, connected through LONI. Sample research of the participating researchers pictured (i.e. biomechanics by Kodiyalam & Wischusen, tangible interaction by Ullmer, coastal studies by Walker, and molecular biology by Bishop).



Infrastructure Overview

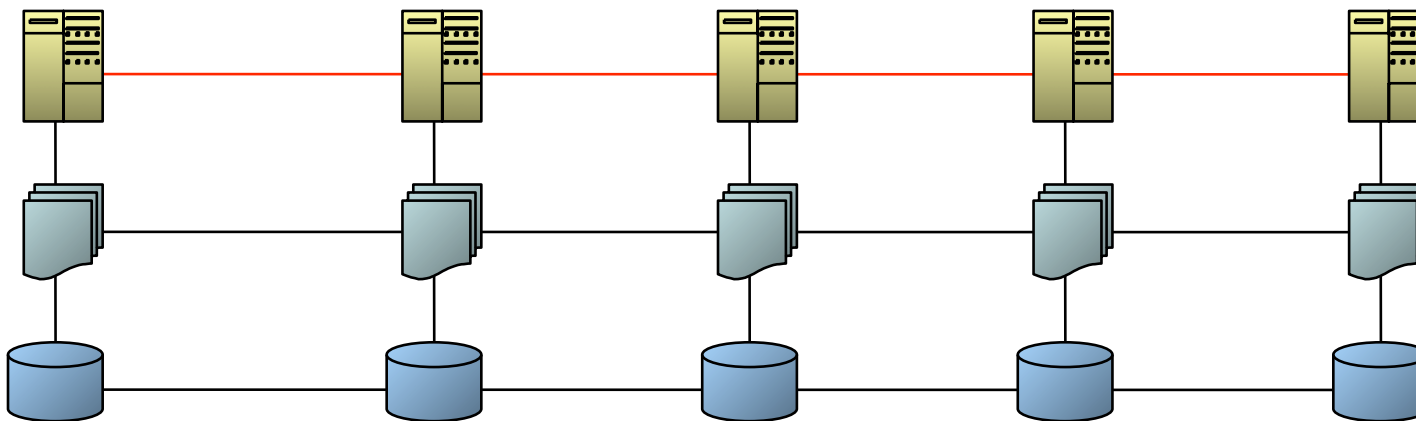
LaTech

ULL

LSU

UNO

Tulane

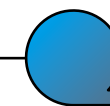
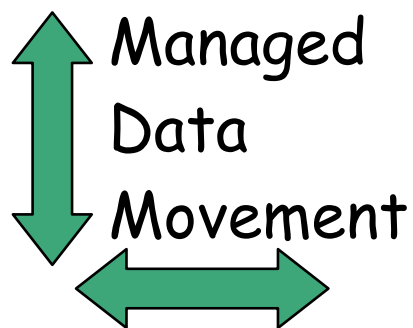


~ 100 TFLOPS

~8 TB RAM

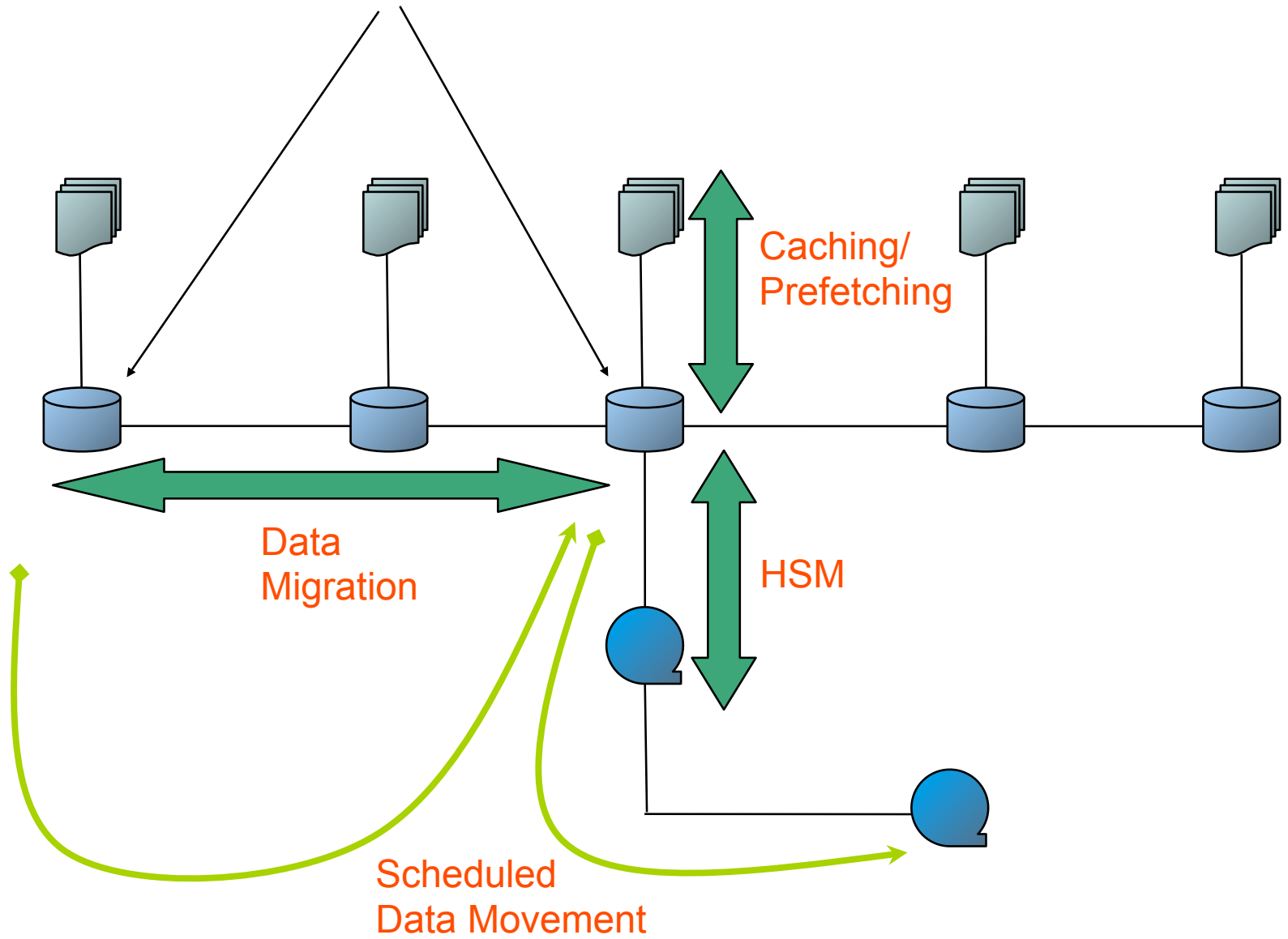
250 TB Disk

400 TB Tape



SDSC
50 TB

Replica & Meta
Data Management

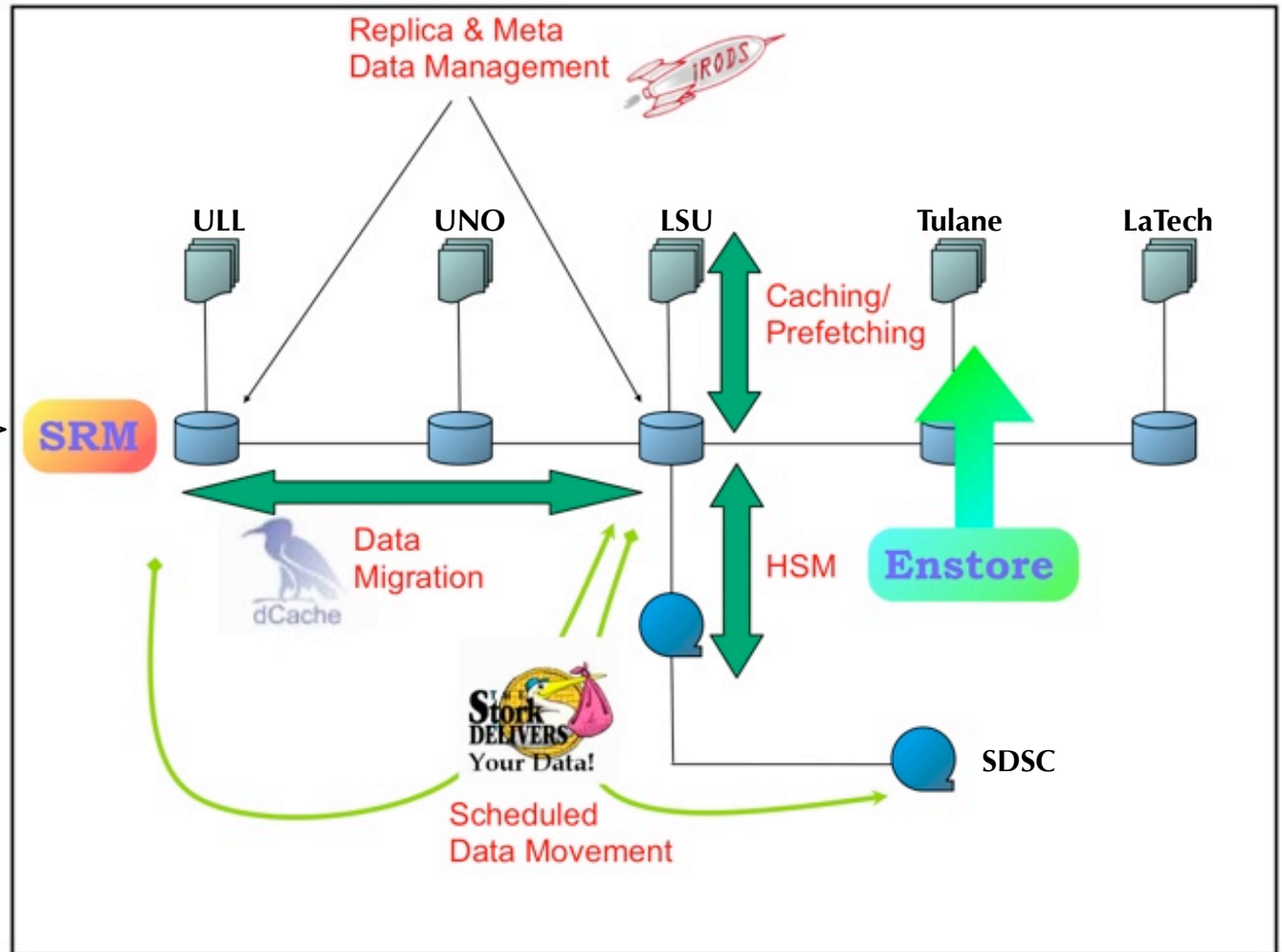


PetaShare Core

POSIX interface:

- NO need to change code
- NO relinking
- NO recompiling
- NO privileged access

petashell



Web interface:

PetaSearch

petashell

- a POSIX compatible shell interface to PetaShare

```
$ petashell
```

```
psh% cp /tmp/foo.txt /petashare/tulane/tmp/foo.txt
```

```
psh% vi /petashare/tulane/tmp/foo.txt
```

```
psh% cp /tmp/foo2.dat /petashare/anysite/tmp/foo2.dat
```

```
psh% genome_analysis genome_data -->
```

```
psh% genome_analysis /petashare/uno/genome_data
```

```
psh% exit
```

```
$
```


petashell Demo

Cross-Domain Metadata Search

- **SCOOP – Coastal Modeling**
 - Model Type, Model Name, Institution Name, Model Init Time, Model Finish Type, File Type, Misc information ...
- **UCoMS – Petroleum Engineering**
 - Simulator Name, Model Name , Number of realizations, Well number, output, scale, grid resolution ...
- **DMA – Scientific Visualization**
 - Media Type, Media resolution, File Size, Media subject, Media Author, Intellectual property information, Camera Name ...
- **NumRel - Astrophysics**
 - Run Name, Machine name, User Name, Parameter File Name, Thorn List, Thorn parameters ...

PetaSearch

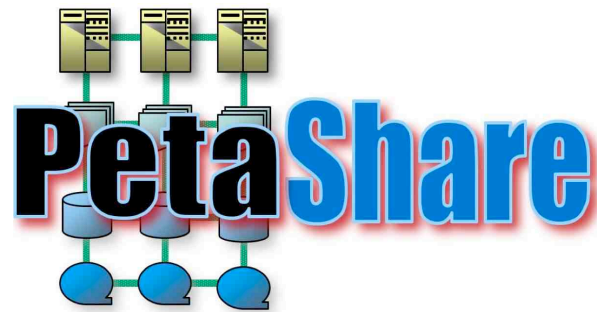
Search Keywords:

- In Archives:
- SCOOP
 - UCoMS
 - NumRel
 - Digital Media
 -

 - All Archives

Questions for Science Drivers

- Storage requirements (short & long term)
- Current data management solution & challenges
- Expectations from PetaShare



A system driven by the local needs (in LA), but has potential to be a generic solution for the broader community!

For more information on PetaShare:
<http://www.petashare.org>

The Large Hadron Collider (LHC)

