

DOSAR: State of Organization

Pat Skubic
University of Oklahoma

DOSAR Workshop at ISU

April 5-6, 2007

Outline

- History of DOSAR
- Goals and Accomplishments
- Contributions to OSG
- Statement of Work submitted to OSG
- Conclusions

What is DOSAR?

- Stands for “Distributed Organization for Scientific and Academic Research”
 - Community and campus based grid organization
 - Its primary goal is spearheading grid use
 - DOSAR in Korean is the God of Martial Arts
- DOSAR stems from the DØ Remote Analysis effort
 - Groups’ efforts in DØ simulation and reconstruction production
- Working closely with other disciplines
- Each group plays a leadership role in realization of computing grids in corresponding campuses and states

DOSAR History

- DØ Remote Computing Era
 - DØ Remote Analysis Model Proposed: Nov. 2001
 - Proposal for RAC accepted and endorsed by DØ: Aug. 2002
 - UTA awarded MRI for RAC: June 2002
 - Formation of DØ Southern Analysis Region: Apr. 2003
 - DOSAR DØ MC Production begins
 - Activation of 1st US RAC at UTA: Nov. 2003
 - Formation and activation of DOSAR Grid for MC: Apr. 2004

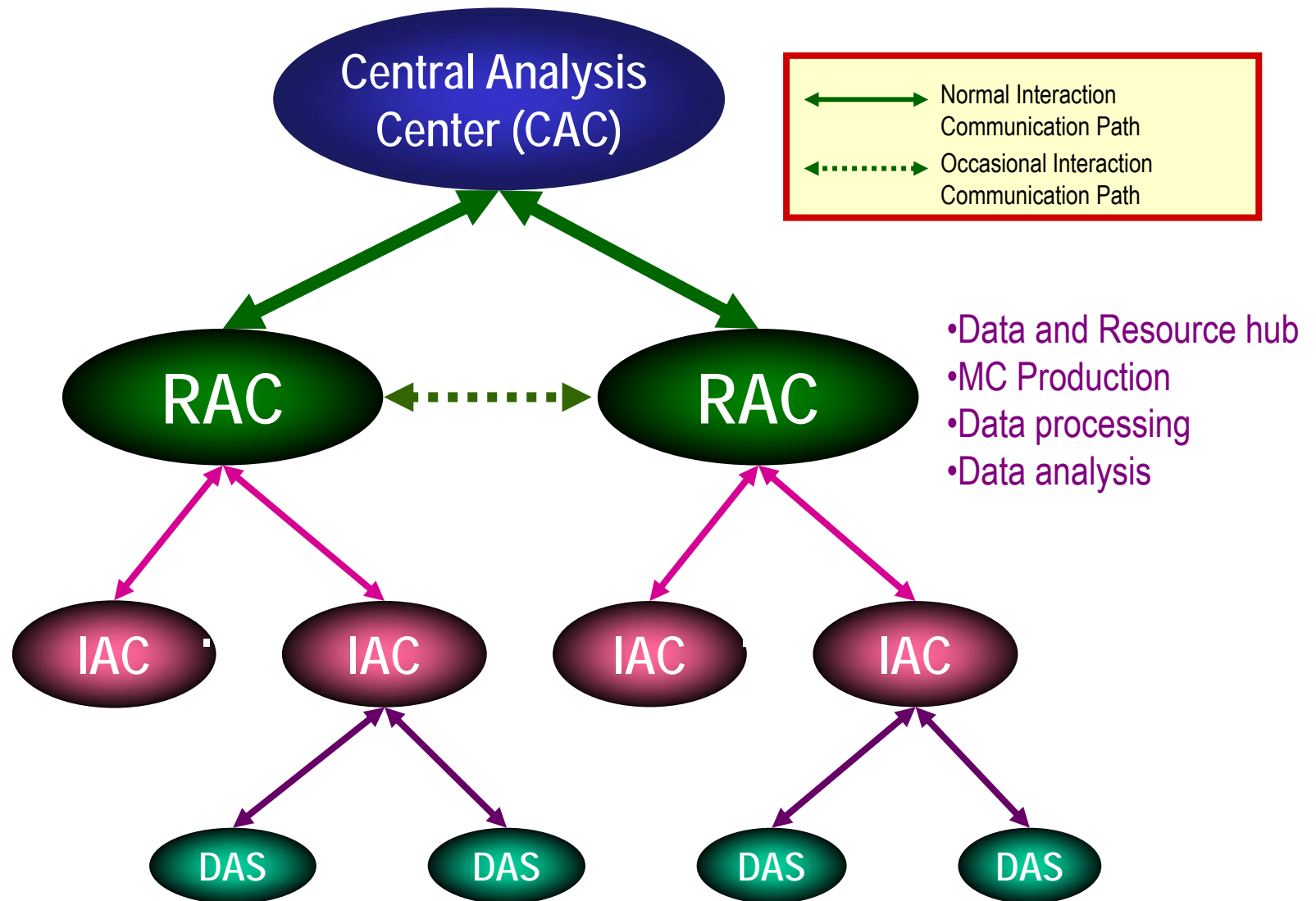
DØ Remote Analysis Model (DØRAM)

Fermilab

Regional Analysis Centers

Institutional Analysis Centers

Desktop Analysis Stations



April 5, 2007

DOSAR Workshop at ISU
P. Skubic

DOSAR History

- Beyond the DØ experiment boundary era
 - Transition to Distributed Organization for Scientific and Academic Research, DOSAR: Apr. 2005
 - Active engagements with LHC experiments begun
 - DOSAR VOMS installed at UTA: May 2005
 - DOSAR registered as a VO in OSG: July 2005
 - ATLAS distributed production and analysis system, Panda, implemented at OU and UTA: Jan. 2006
 - All groups engaged in LHC experiments

DOSAR Consortium

➤ First Generation IAC's

- ✓ University of Texas at Arlington
- ✓ Louisiana Tech University
- ✓ Langston University
- ✓ University of Oklahoma
- ✓ Tata Institute (India)

• Second Generation IAC's

- Cinvestav, Mexico
- ✓ SPRACE - São Paulo Regional Analysis Center, Universidade Estadual Paulista; Brazil
- ✓ University of Kansas
- ✓ Kansas State University

• Third Generation IAC's

- University of Arizona, Tucson
- Rice University
- Ole Miss, University of Mississippi
- Iowa University State University
- Louisiana State University
- Oklahoma State University

Primary Goals of DOSAR and Achievements

- Harness for common grid use a diverse set of human and computing resources previously unavailable
 - LTU brought in Super-Mike and additional resources for DØ and ATLAS
- Empower offsite remote users with desktop data analysis capability as if they are at the experiment
- Prepare all involved institutions to perform data analysis using grid technology on DØ and future HEP experiments such as the LHC experiments, CMS and ATLAS
 - OU, LU, LTU & UTA are members of SWT2 physics analysis group
 - SPRACE plays a leading role in CMS remote analysis as Tier 2 Center

Primary Goals of DOSAR and Achievements

- Collaborate to use cutting edge grid technology to promote a wide range of interdisciplinary and educational activities within the member regions
 - UTA plays leadership role in HiPCAT, Texas Grid community; Leading BioTex grid, working with chemists, geologists and medical professionals
 - LTU Working as a leading institution in LONI, Louisiana Grid community
 - OU and LU have been working toward creating “The Oklahoma state grid”, working closely with OSCER
 - SPRACE leads Brazilian national grid effort w/ funds!!!

Primary Goals of DOSAR and Achievements

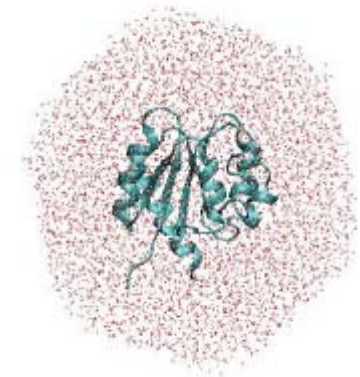
- Strongly participate in global grid efforts such as OSG or EGEE to contribute to the development of grid resources and technology, utilizing a mixture of dedicated and desktop resources.
 - Met with OSG leadership to discuss DOSAR's contribution
 - We bring direct contact to states and campuses built upon the strong collaboration between groups working closely together
 - Contributed to grid software development such as PanDA (Used in ATLAS MC production and protein molecular dynamics – see next slide)

Protein molecular dynamics on OSG using CHARMM

- Solution: Use PanDA and a set of custom management scripts.

The Scheduler Interface

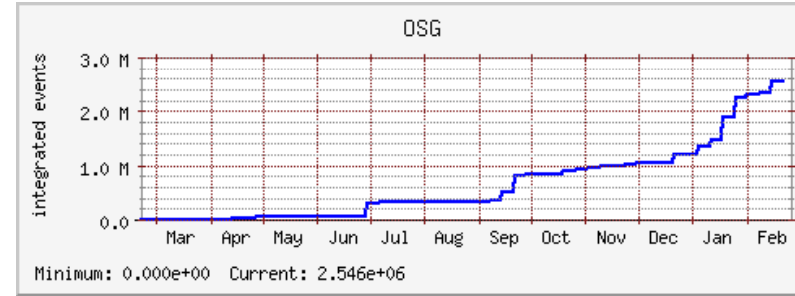
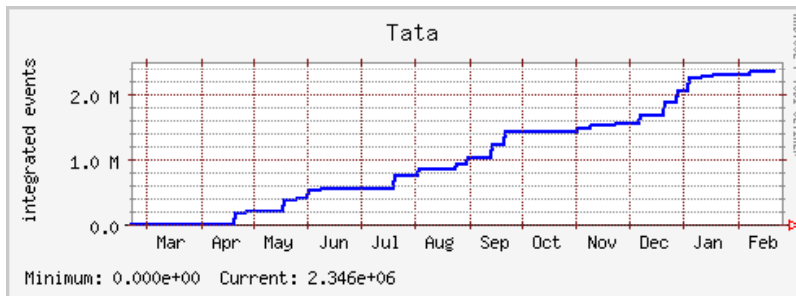
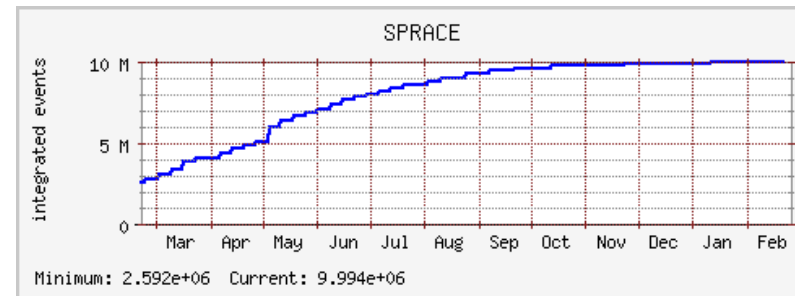
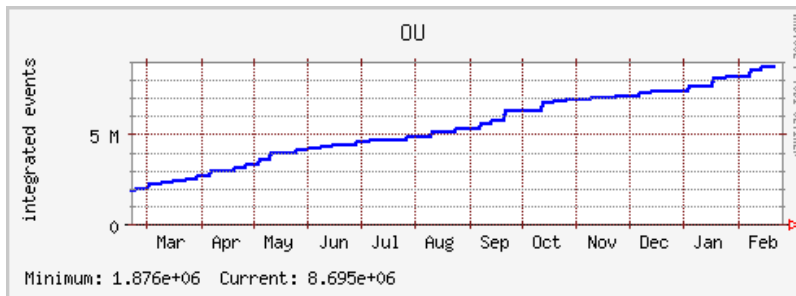
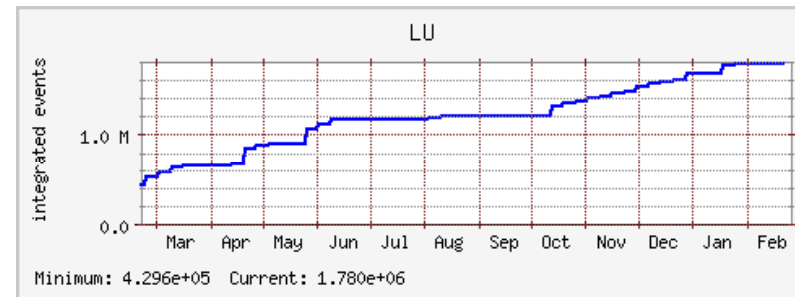
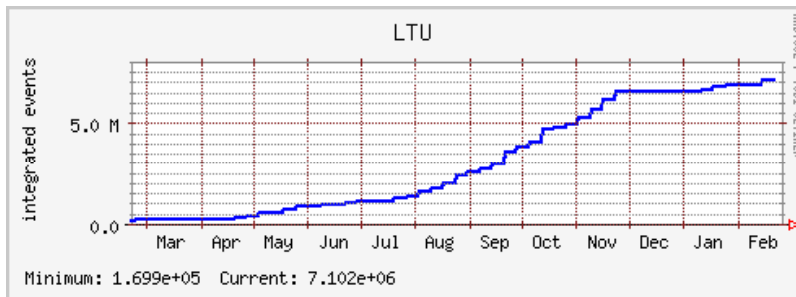
- We use the PanDA front end.
- We also use TestPilot and run our own pilot scheduler for maximum control.
- Users can track jobs via a Web interface.



Primary Goals of DOSAR and Achievements

- Exploit grid projects and international research collaborations to develop a highly trained technical workforce within the member regions.
 - Each institution provides enormous opportunities to students from other disciplines to work in DOSAR
 - Created exchange programs for CSE students
 - DØ, OSG and ATLAS
 - 10 CSE students graduated from the exchange programs and play leadership role in the grid community

D0 MC Production




April 5, 2007



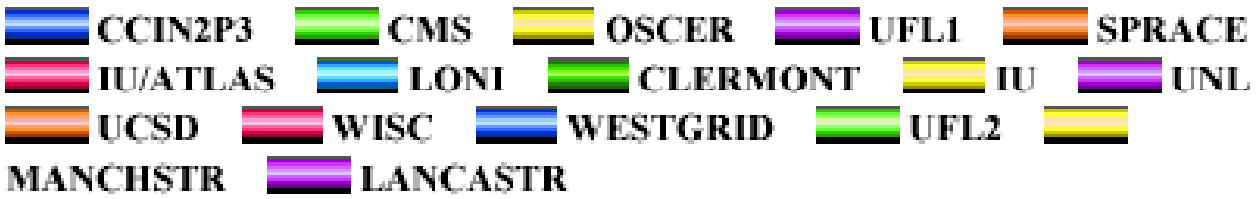
DOSAR Workshop at ISU
P. Skubic

DØ Data Reprocessing

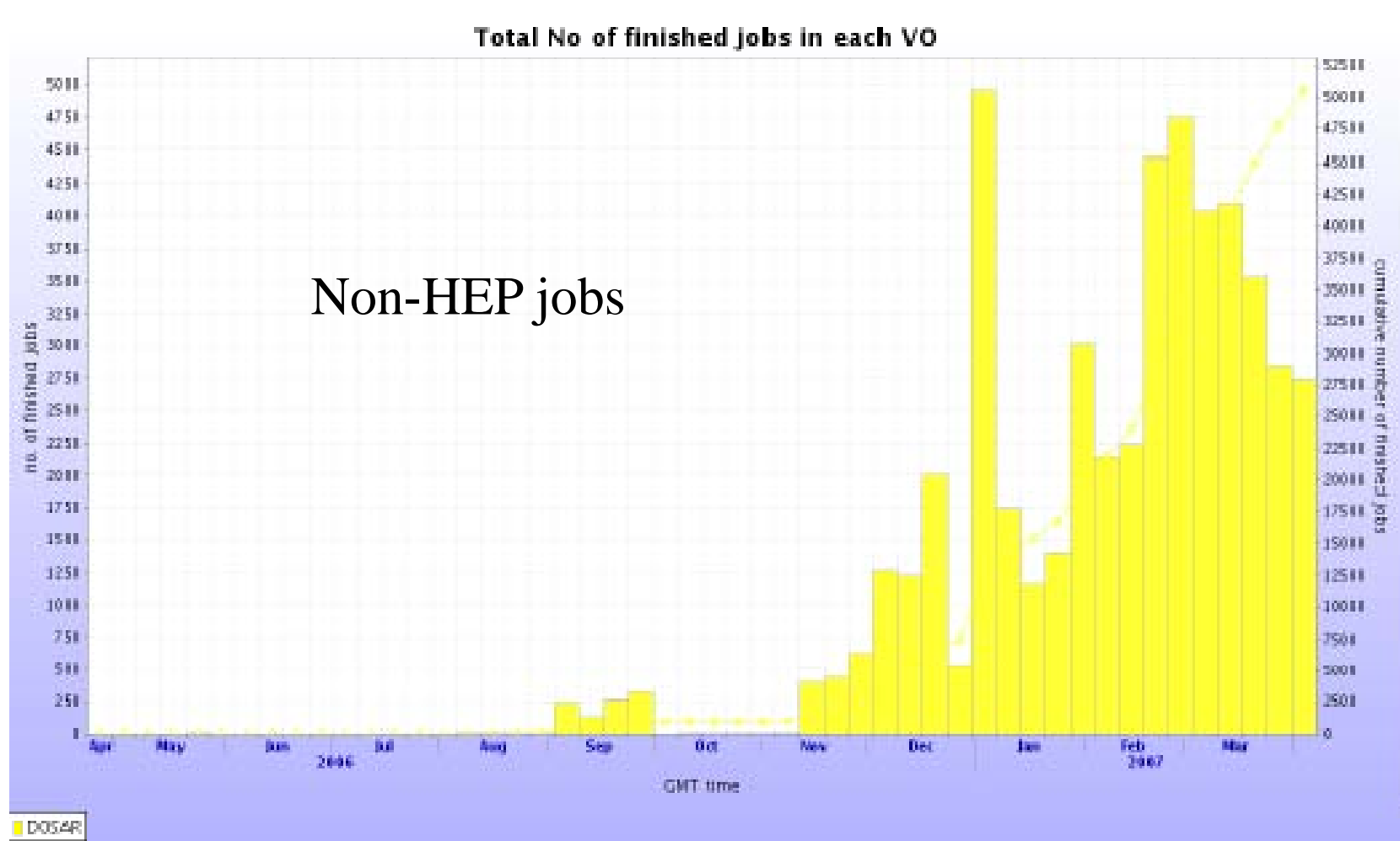
P17 Reprocessing Status as of 19-Apr-2006 (Remote sites only)

Processed Events	840090682	
Sites		

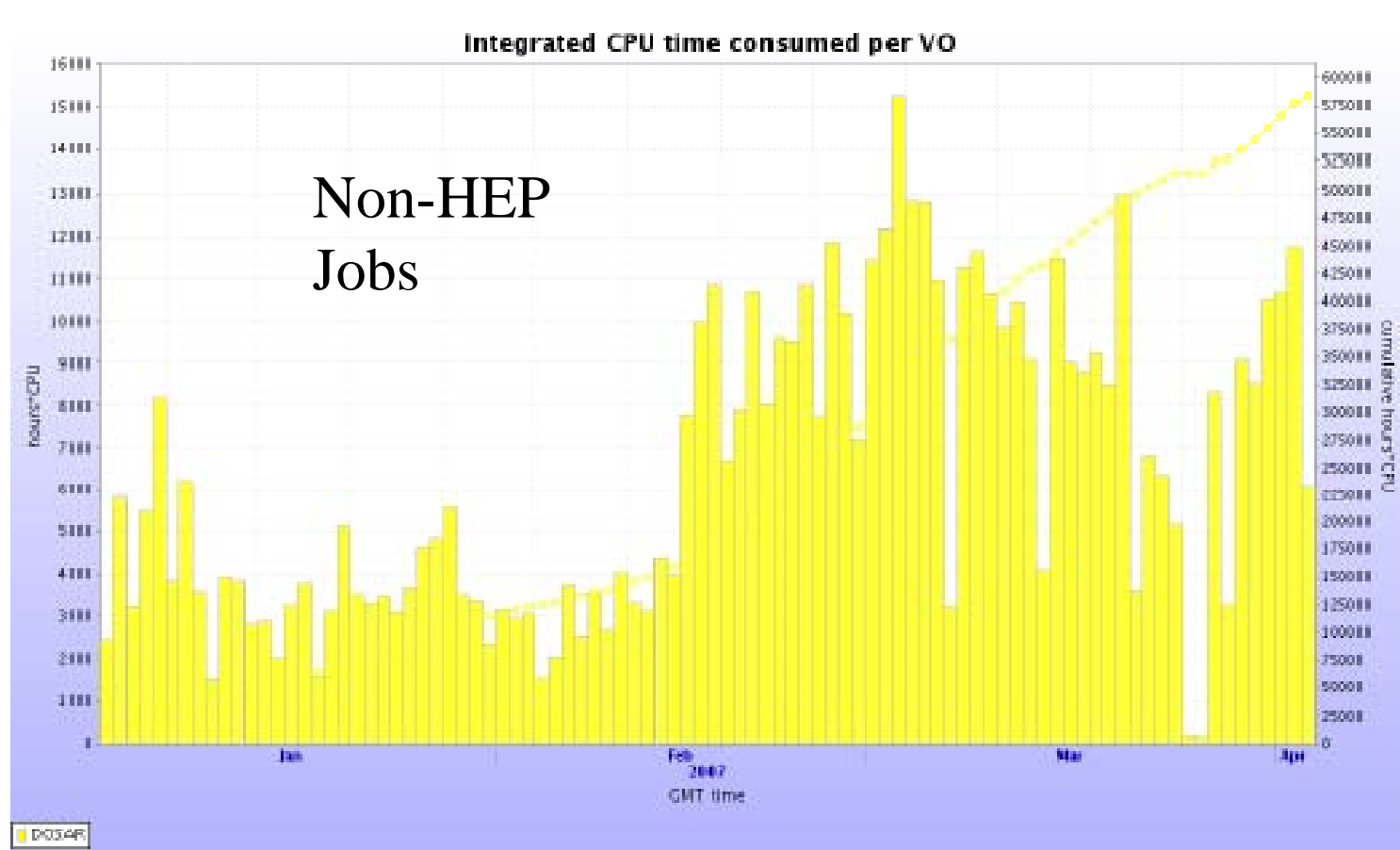
P20.07.01 Reprocessing Status as of 03-Apr-2007 (Remote sites only)

Unmerged Events	168409275	
Merged Events	108037557	
Contribution		

Finished jobs in DOSAR VO (one year)

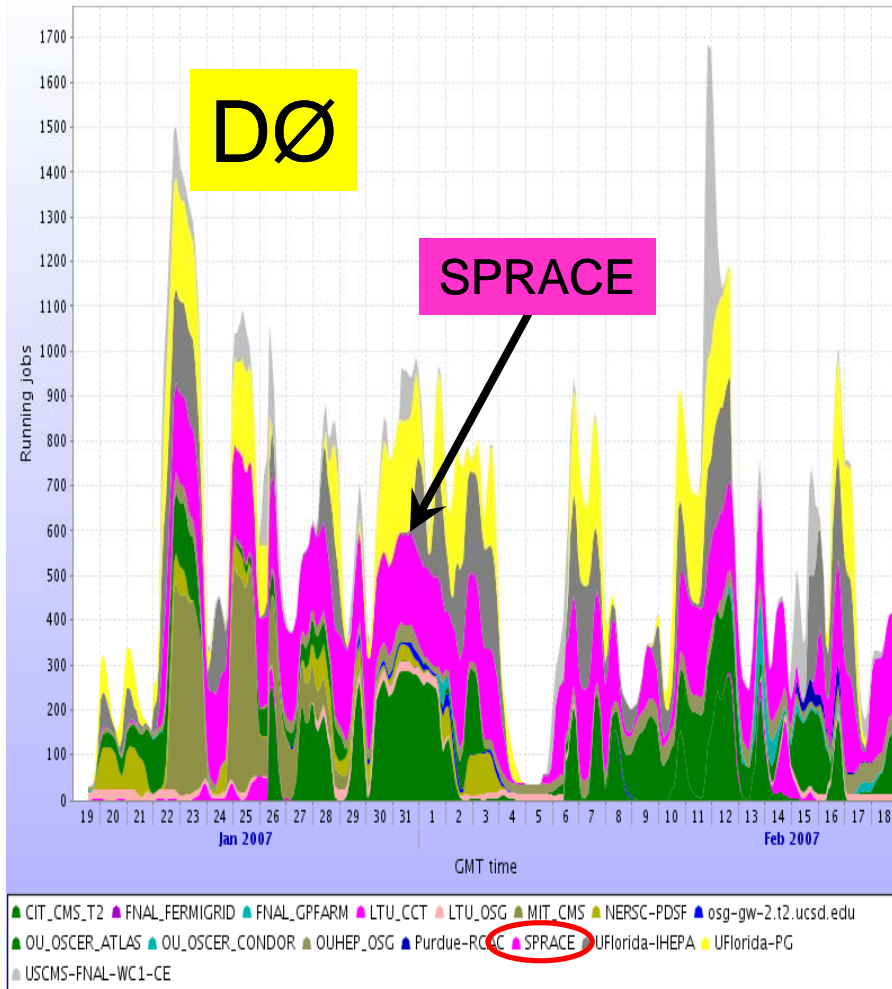


CPU time used by DOSAR VO (3 mo.)



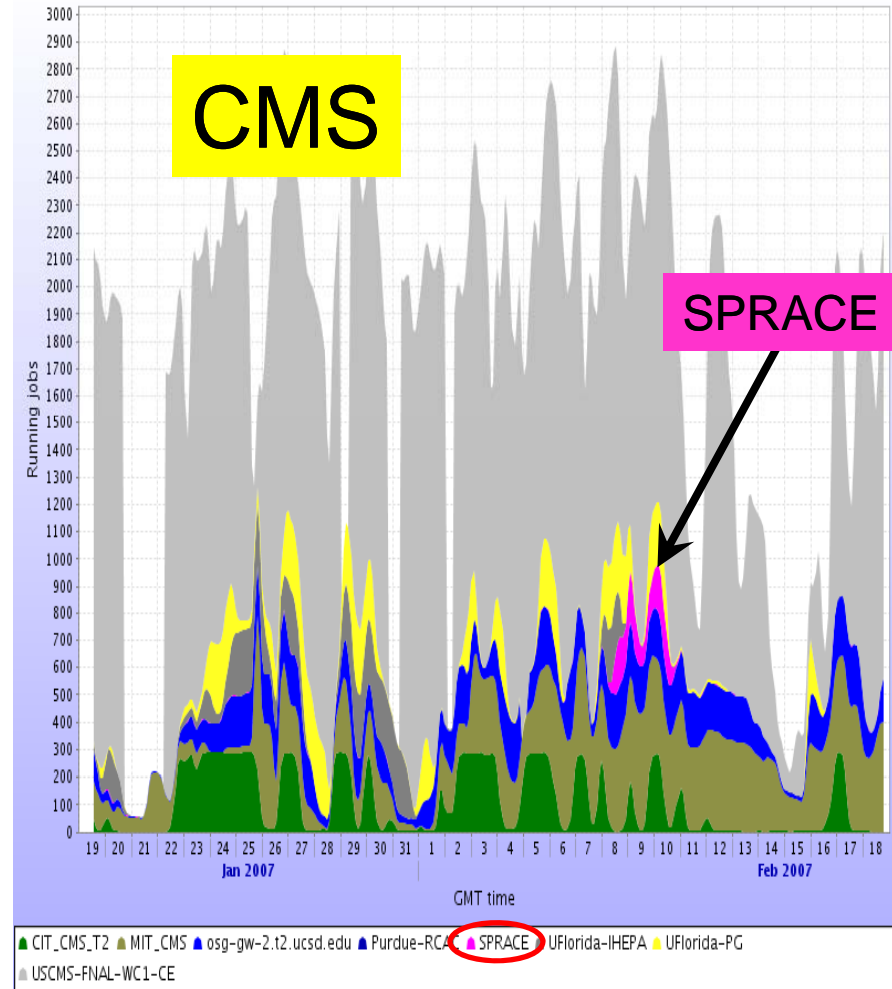
MonALISA (one month)

Jobs status for DZERO VO



April 5, 2007

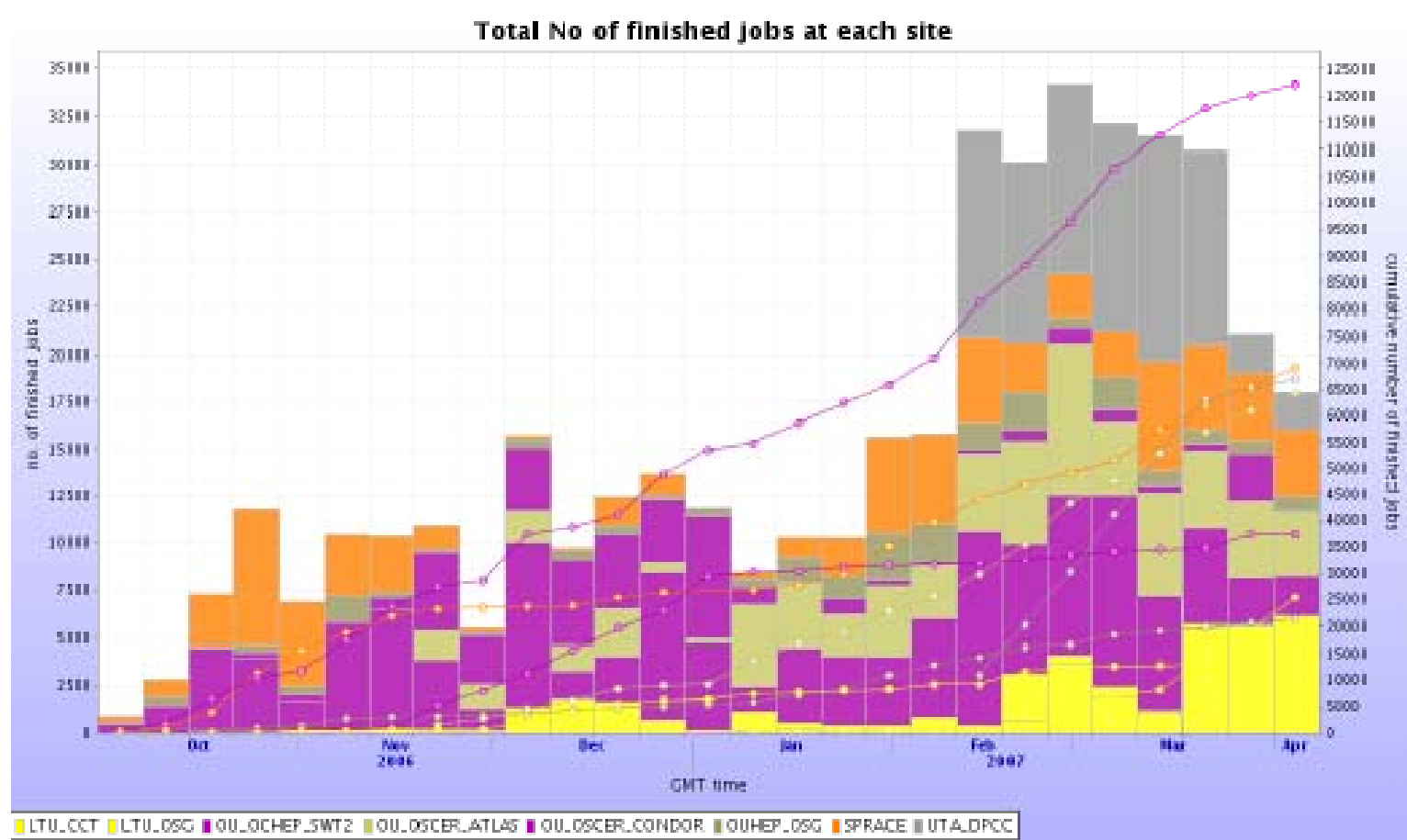
Jobs status for CMS VO



DOSAR Workshop at ISU
P. Skubic

16

DOSAR: Finished jobs by site (6 mo.)



April 5, 2007

DOSAR Workshop at ISU
P. Skubic

17

Working with OSG

- We accomplished virtually everything we wanted in DØ
 - J. Snow (LU) is the leader in MC production
- Bring fresh expertise to OSG
- Expertise in monitoring solutions such as Ganglia and MonALISA → UTA working on Panda ATLAS monitoring together with CMS experiment
- LTU working on first implementation of HAOSCAR (High Availability Open Source Cluster Application Resources) w/ OSG stack

Contributions to OSG

- Testing of framework, middleware, and user interfaces.
- Active participation in OSG integration and deployment activities.
- Partner with high-speed optical network initiatives
- Help implement and utilize grid computing technology for educational use.
- Participate and test grid-based HEP data analysis and disseminate the experience to OSG

Statement of Work submitted to OSG

- ISU: CONDOR pool development using MS Virtual Server
- LTU/LSU: Grid production; Deployment of high availability computing algorithms using OSG; LONI
- SPRACE: Grid production, data reprocessing
- OU: CONDOR pool development using co-Linux; Grid production; Desktop cluster & OSG deployment
- UTA: MonALISA monitoring development; Grid production; HiPCAT

DOSAR Strategy

- Maximally exploit existing software and utilities to enable as many sites to contribute to the DØ and LHC experiments
 - Continue to participate in DØ MC and Reco activities in opportunistic manner
 - Focus on LHC experiments
 - Support OSG move into the global grid framework
 - Engage in and contribute significantly in OSG
- Engage in realization of computing grid beyond HEP
 - Work closely with campus and state computing people to bring grid onto campuses
- Bring DOSAR specific computing jobs to the grid
- Want to make grid experience more powerful

Some Successes in Funding at DOSAR

- Funds from NSF MRI for UTA – RAC: 2002
 - Construction of the first U.S. university based RAC
- EPSCoR + University funds for LTU – IAC: 2003
 - Increase IAC compute resources
- Brazilian National Funds for SPRACE: 2003 & 5
 - Construction of an extensive RAC for
- EPSCoR funds for OSU, OU & LU: 2004
 - Compute resources for IAC
 - Human resource for further development
 - Renewal submitted
- LTU at part of LONI wins support from State of LA: 2005
- OU, LU and UTA, together with UNM, won a joint ATLAS Tier 2 site: 2005
- LTU won a joint MRI funds: 2006

Typical Example: São Paulo Regional Analysis Center

	Phase 1 (2004)	Phase 2 (2005)	Phase 3 (2006)
CPU	50	115	242
Power (TeraFlop)	0.24	0.62	1.44
Storage (TB)	4	12	20
Personnel (FTE)	1	1	1.5

Conclusions

- DOSAR is an example of a successful grid organization
 - Critically important asset to DØ
- All groups actively engaged in LHC experiments
 - Yet DOSAR crosses the experimental boundary
- Using DOSAR resources for DØ and LHC data analyses and production
- Closely engaged in OSG activities
- Continue Desktop Cluster deployment, CONDOR pools
- Provide leadership role in state-wide grid efforts
- Expanding membership and scope