DOSAR State of the Organization

DOSAR Workshop VII
University of Texas, Arlington
March 30-31, 2006

Joel Snow
Langston University

State of the Organization Outline

- Identity and Origins
- Goals
- Accomplishments
- Present and Near Future Resources
- Ongoing Activity
- Future Directions
- Conclusion

DOSAR Identity & Origin

- Distributed Organization of Scientific and Academic Research
  - A community and campus based grid organization
- Initially organized for distributed production and analyses in the DØ experiment (DØSAR)
- Playing leadership roles in realization of computing grids in corresponding campuses and states

DOSAR Consortium

- **University of Texas at Arlington**
- **Louisiana Tech University**
- **Langston University**
- **University of Oklahoma**
- **Universidade Estadual Paulista, São Paulo,Brazil**
- **Tata Institute of Fundamental Research, Mumbai, India**
- **University of Mississippi**
- University of Kansas
- Kansas State University

Primary Goals of DOSAR

- **Communicate and disseminate accumulated experiences with real large-data analyses to the grid community for the benefits of future HEP experiments and society as a whole.**
- **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.**
- **Exploit grid projects and international research collaborations to develop a highly trained technical workforce within the member regions.**

What did we want to accomplish in DØ?

- **Construct end-to-end service environment in a smaller, manageable scale**
- **Train and accumulate local expertise and share them**
- **Form a smaller group to work coherently and closely**
- **Draw additional resources from variety of funding sources**
- **Promote interdisciplinary collaboration**
- **Increase intellectual resources for the experiment**
  - **Enable remote participants to more actively contribute to the collaboration**
- **Form a grid and use it for DØ and LHC experiments**
  - **MC production - Raw data processing - Analysis**
- **Promote and improve IAC's group stature**

Short DOSAR History I

- Formation of DØ Southern Analysis Region: Apr. 2003
  - DOSAR DØ MC Production begins
  - McFarm job manager developed at UTA
- Activation of DØSAR Grid for MC: Apr. 2004
  - FNAL's SAMGrid infrastructure
  - Grid3 at UTA & OU
- Transition to Distributed Organization of Scientific and Academic Research, DOSAR: Apr. 2005
  - reflects migration to LHC and broader concerns

Short DOSAR History II

- Three DOSAR sites participate in DØ data reprocessing: May – November 2005
  - UTA, OU, UNESP
- ATLAS data challenges – UTA and OU
- DOSAR VOMS installed at UTA: May 2005
- DOSAR registered as a VO in OSG: July 2005
- Tier 2 Sites
  - UTA/OU/LU/UNM (ATLAS), TIFR (CMS), UNESP (CMS)

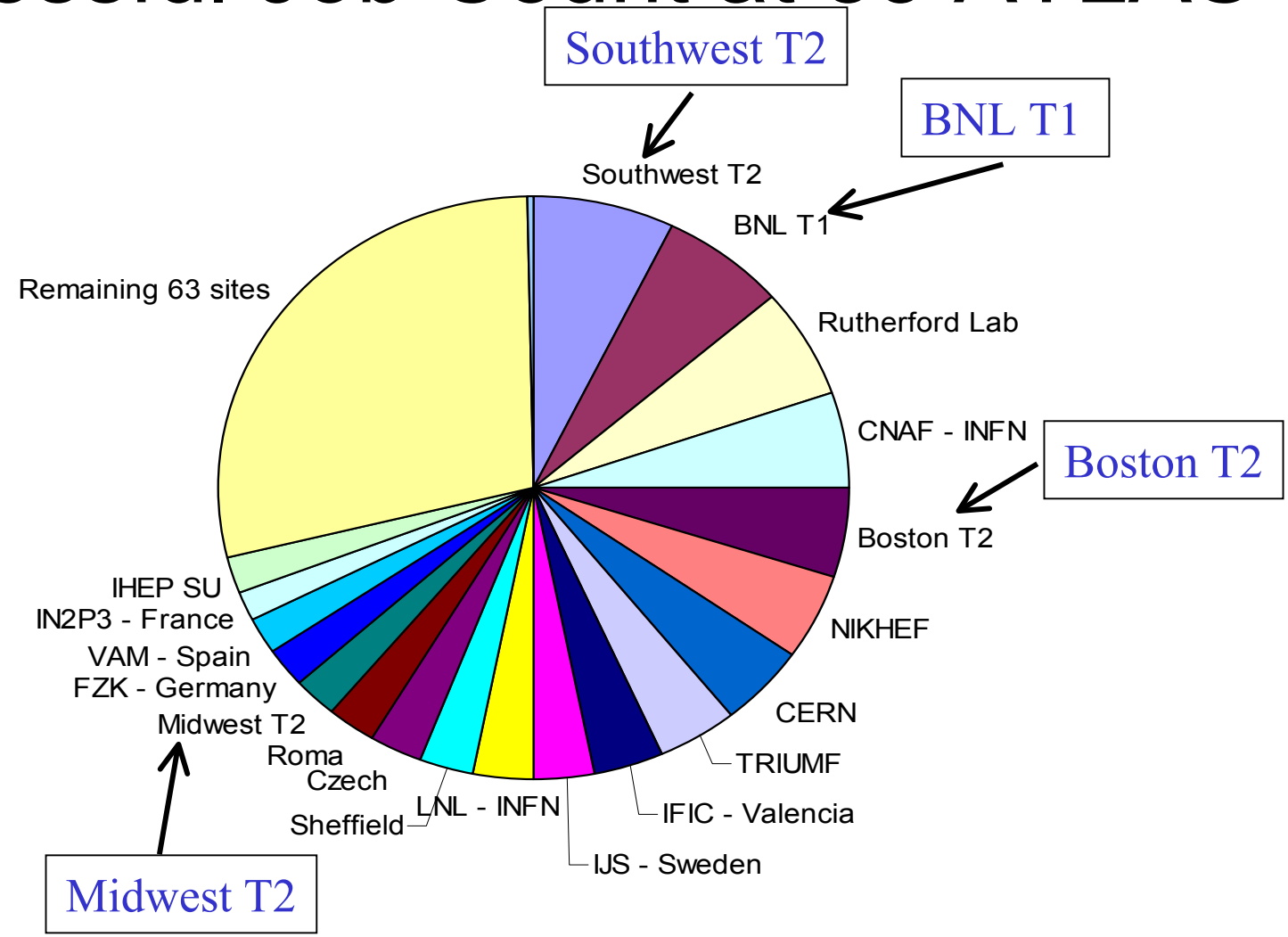
Some Successes in Funding at DOSAR

- **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 Brazil**
- **EPSCoR funds for OU & LU: 2004**
  - **Compute resources for IAC**
  - **Human resource for further development**
- **LTU as part of LONI wins support from State of LA: 2005**
- **OU, LU, UTA, UNM won a joint ATLAS Tier 2 site: 2005**

DOSAR DØ MC and Reprocessing

Institution	Inception	NMC (TMB) x10^6	NRepro. x10^6
LTU	6/2003	1.1	0
LU	7/2003	2.6	0
OU	4/2003	4.6	14.9
Tata	6/2003	3.5	0
SPRACE	4/2004	7.1	9.6
UTA-HEP	1/2003	4.2	0
UTA-RAC	12/2003	11.0	30.1
Total As of	3/6/06	34.1	54.6

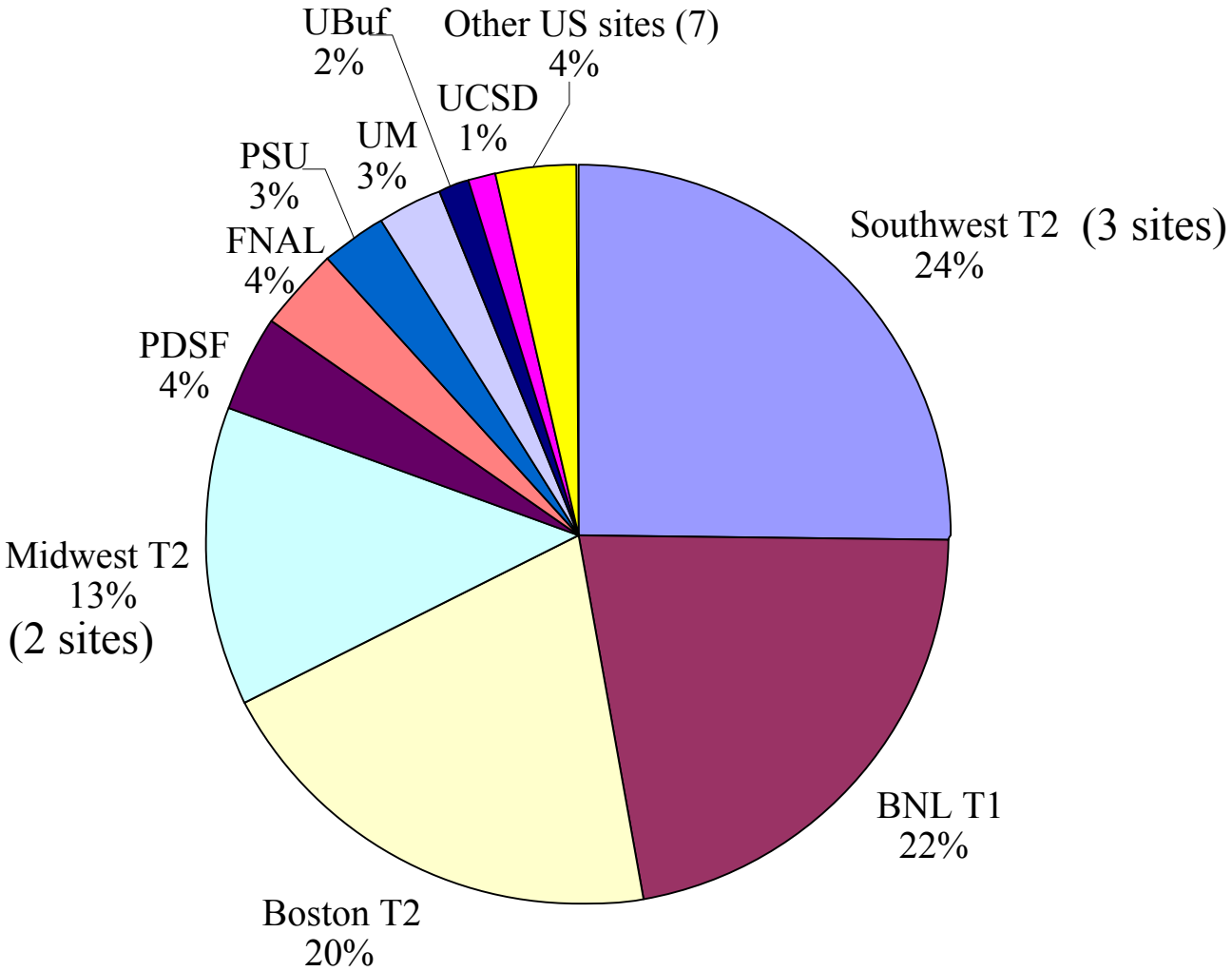
Rome Grid Production
Successful Job Count at 83 ATLAS sites



3 out of top 5 contributions in ATLAS were from U.S.

Slide courtesy of K. De

U.S. ATLAS Grid Production (Rome/DC2 combined)



20 different sites used in the U.S.

ATLAS Tier 2's played dominant role

Slide courtesy of K. De

DOSAR Computing Resources

- ATLAS Southwest Tier 2 Center
  - UTA, OU, LU, UNM
- Desktop Clusters
  - OU, LU
- Dedicated Clusters
  - LTU, OleMiss, TIFR (Tier 2), São Paulo (Tier 2), OU
- Opportunistic Clusters
  - OU (Topdawg, Condor Pool) , LTU (SuperMike)

DOSAR Dedicated Computing Resources

Institution	Core-GHz	Memory **(GB)**	Storage **(TB)**
Langston	26	11	2.5
LTU	23	15	0.8
OU	297 + [540]	200 + [270]	8.5 + [5.0]
São Paulo	313 + [410]	91 + [128]	14.4 + [1.2]
TIFR	221		11.0
Ole Miss	163		10.0
UTA	1120	704	55.0
Total	2140 + [950]	1006 + [398]	102.2 + [6.2]

[ ] - new resources within 1 year

DOSAR Non-Dedicated Resources

Institution	Core-GHz	Memory **(GB)**	Storage **(TB)**
LTU (@ 50%)	512	256	
OU (@ 25%)	[3277]	[2048]	[10.0]
UTA (@ 50%)	384 + [205]		45.0 + [9.0]
Total	448 + [921]	128 + [512]	22.5 + [7.0]

LTU : SuperMike at LSU

OU: TopDawg

UTA: DPCC

[ ] - new resources within 1 year

DOSAR Networking

- UTA & LU presently 1Gbs
  - UTA will connect to NLR (10Gbs) via LEARN, when?
  - OU will connect to NLR within 6 months
- SPRACE at 622Mbs $\rightarrow$ 1Gbs
- LTU to NLR via LONI
- OleMiss 155Mbs (?)
- LU 100Mbs $\rightarrow$ 1Gbs
- TIFR $\rightarrow$ 622Mbs (needs improvement)

DOSAR Projects I

- DØ MC production (SPRACE,OU,LU,LTU,Tata)
- DØ primary data processing (OU)
- SamGrid – OSG gateway SAM station at OU
- PANDA – ATLAS job production system (UTA,OU)
- ATLAS Tier2 (UTA,OU,LU)
- CMS Tier2's (São Paulo, TIFR)

DOSAR Projects II

- Physics Analysis – regional ATLAS meeting (UTA,OU,LU,LTU,UMN,SMU,UTD)
  - Grid enabled analysis tools (UTA)
- State Grid & Networking initiatives
  - TX: HiPCAT, LEARN
  - LA: LONI
  - São Paulo: Brazilian Grid
  - OK: OCHEP

DOSAR Projects III

- OSG (UTA,OU,SP)
  - Deployment, Production, & Integration
  - Add other sites and resources
- Advanced international networking and its applications (SP, TIFR)

DOSAR in

MonALISA
MONitoring Agents using a Large Integrated Services Architecture

DOSAR Future

- Actively engage in LHC experiments
  - UTA, OU, LU, and LTU participating actively in ATLAS distributed production and analysis
  - Ole Miss, SP, TIFR in the process of working with CMS
- Actively participate in OSG as a community based, grass-root VO
  - Participate in OSG activities
    - Install the software stack at remaining sites (LCG at TIFR?)
  - Fully utilize the LHC and LC connections
- Continue promoting interdisciplinary collaboration and grid interoperability
- Actively participate and lead grid computing efforts in the respective states
- Get engaged in employing grid computing technology for education and clinical applications

Conclusions

- DOSAR contributed to DØ production significantly
  - MC, reprocessing, primary processing
- Network infrastructures for members significantly improved
- Became a VO in OSG
- Actively participating in LHC experiments
- Use DOSAR for DØ and LHC data analyses
  - Producing presentable results from DØ data analyses in the regional grid
- Fully utilize DOSAR beyond HEP
- Play larger leadership role in state-wide grid initiatives
- Promote grid technologies on campus and in our regions
- Participate in global grid efforts as a VO