



Interinstitutional Cyberinfrastructure

A VO Case Study

James Leous

Information Technology Services
Penn State



Cyberinfrastructure

Cyberinfrastructure integrates hardware for computing, data and networks, digitally-enabled sensors, observatories and experimental facilities, and an interoperable suite of software and middleware services and tools.

— NSF Report on Cyberinfrastructure



Research as Process

- literature search/forming the team



Research as Process

- literature search/forming the team
- writing the proposal



Research as Process

- literature search/forming the team
- writing the proposal
- funding



Research as Process

- literature search/forming the team
- writing the proposal
- funding
- data collection



Research as Process

- literature search/forming the team
- writing the proposal
- funding
- data collection
- data processing



Research as Process

- literature search/forming the team
- writing the proposal
- funding
- data collection
- data processing
- publish



Research as Process

- literature search/forming the team
- writing the proposal
- funding
- data collection
- data processing
- publish
- archive



Research as Process

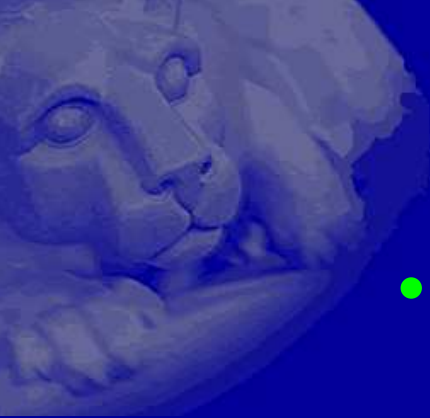
- literature search/forming the team
- writing the proposal
- funding
- data collection
- data processing
- publish
- archive
- REPEAT



S&E Indicators 2008

S&E research is an increasingly collaborative activity. Between 1988 and 2005, the share of publications with authors from multiple institutions grew from 40% to 61%.

- Coauthored articles with only domestic institutions in the bylines grew from 32% to 41% of all articles.
- Articles with institutions from multiple countries an indicator of international collaboration and the globalization of science grew from 8% to 20%.



More S&E Indicators 2008

- Between 1995 and 2005, coauthorship with foreign authors increased by 9 percentage points for authors in the federal government and academia.
- Of the S&E fields, astronomy had the highest rate of international coauthorship in 2005, at 58%, well above the U.S. national average of 27% across all fields.



How Do We Support This?

Increasingly we support CI with Web applications or Web interfaces to existing computational clusters, instruments, observatories, experiments, and archives.

- Lit Search — Refworks, zotero, del.icio.us
- Research Information Systems — Kualu Research, home grown
- Proposal — Wikis, Google Docs
- Data Management/Processing — Web portals, Web CMSs
- Archives — Fedora, DSpace, custom



COmanage

COmanage moves the identity and access management (IAM) out of the individual tools and into the collaboration itself.

COmanage can tie together:

- mail lists
- video/audio conferencing software
- group/event calendaring
- wikis
- shared file space
- repositories/archives?

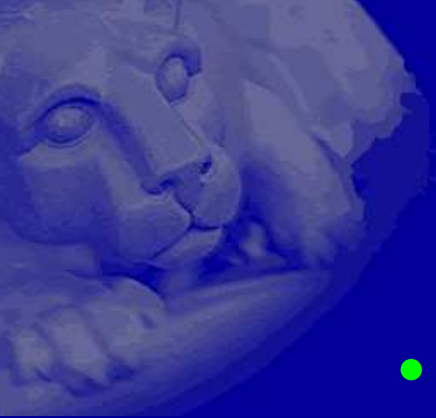


Advantages of COmanage

COmanage provides a platform for the integration of many Web applications integrating them into the Collaborative Organization (CO).

COmanage enables a CO by importing an individual attributes from their own institutions (although it could be just at one site), and integrating them with CO specific attributes.

How? “attribute ecosystem voodoo” — see Ken’s talk



Disadvantages of COmanage

- Some tools don't lend themselves to “moving identity and access management” out of them.
- What about institutions who don't participate?



Case Study – LIGO

The Laser Interferometer Gravitational–Wave Observatory (LIGO) is:

- One of the largest and most ambitious projects ever funded by NSF
- Composed of 600 scientists from 45 institutions
- A VO since 1997
- Managed by the LIGO Scientific Collaboration (LSC)
- PI institutions are MIT, CalTech, Penn State, UW–Milwaukee, and Stanford



LIGO Challenges

- Data rates of 0.5 PB/yr across 3 LIGO detectors ($> 1\text{TB/day}$)
- many institutions provide shared infrastructure (*e.g.* computational clusters, wikis, instrument control/calibration)
- Just starting to think about identity and access management (IAM)
- International collaboration — ACIGA, GEO, VIRGO
- A typical researcher has dozens of accounts.



LIGO CO Work Plan

- Penn State Gravity Team implemented LIGO roster based on LDAP and Kerberos
- LIGO “Hackathon” in January very productive
- LIGO Authentication and Authorization working group identified three focus areas: “ssh,” “Web,” and “grid.”
- Installed Kerberos, Shibboleth IdPs, several SPs, Shib enabled Confluence, Grouper talking to LIGO roster (as local attribute store), and CManage



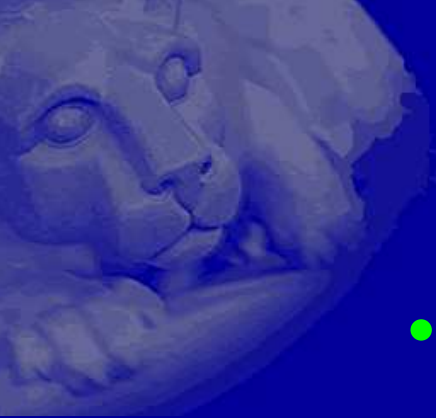
LIGO CO Work Plan

- Installed SPNEGO Web server
- Prototyped several methods for using Web to obtain Grid credentials using the COmanage infrastructure
- LSC Authentication and Authorization working group plan approved 4/2008.



Next Steps

- Leverage evolving LIGO IAM infrastructure



Next Steps

- Leverage evolving LIGO IAM infrastructure
- Establish permanent instance of LIGO COmanage



Next Steps

- Leverage evolving LIGO IAM infrastructure
- Establish permanent instance of LIGO COmanage
- Encourage remaining institutions to join InCommon



Next Steps

- Leverage evolving LIGO IAM infrastructure
- Establish permanent instance of LIGO COnmanage
- Encourage remaining institutions to join InCommon
- Detect Gravity Waves?