Computation Institute

# Campus Grids & Campus Infrastructures Community

Rob Gardner

Computation Institute / University of Chicago

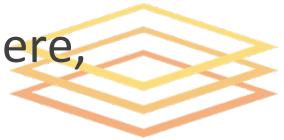
June 4, 2013

# CIC Year 1 Goals

- Development of a topical seminar series and forum highlighting concepts in the development and use of campus infrastructures (**done**)
- Convening face-to-face meetings of the OSG CIC for both infrastructure providers and domain experts/ leaders on campuses (**done**)
- Development of a campus engagement program which programmatically develops ties between research domain experts, campus infrastructure providers and the CIC. (**failed**)
- Developing a program for CIC engagement with XSEDE. (invited to meetings, **but no program**)

# CIC Year 1 Milestones

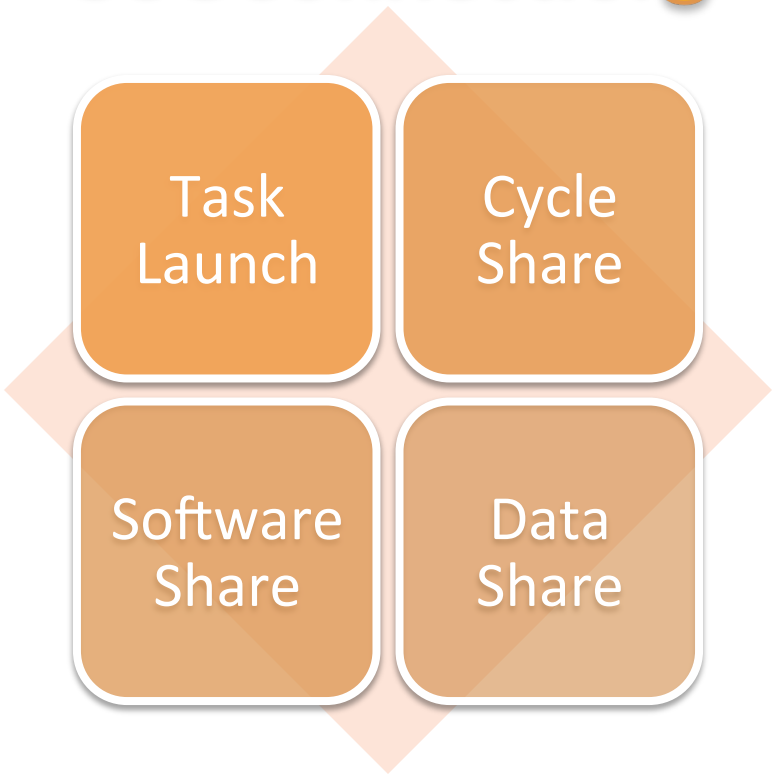
- Define the appropriate metrics for telling the campus story in OSG. We have discussed these in terms of:
  - Making distributed high throughput computing easy, visible (awareness) and ubiquitous (**failed**)
  - Finding the appropriate metric for measuring "presence" on campuses (**failed**)
  - Capture science success stories, indicating the multiplicative effects of using campus and distributed HTC resources (**failed**)
  - Classification of infrastructures with a maturity model [12] (its there, **extend?**))
- Establish the CIC Topic Seminar series as a staple for community building and knowledge sharing (**done**)
- Convene one face-to-face CIC meeting with a broad technical program compelling to the campus infrastructure providers and users (**done**)
- Promote community through use of a CIC resource center (social contacts, topical seminar materials, pointers to tools and guides) (**done**)



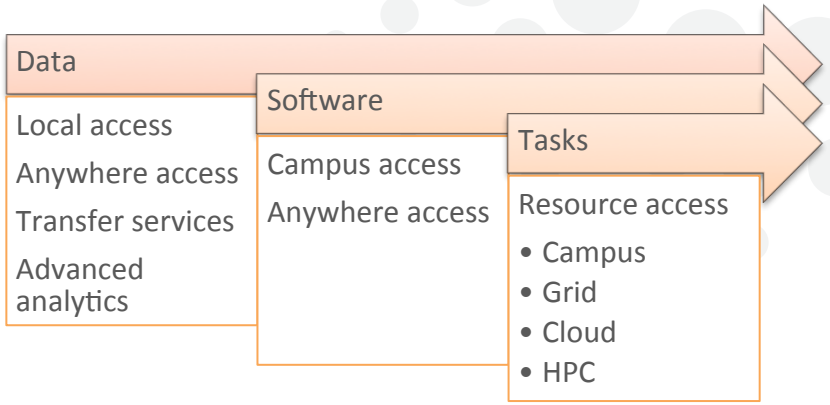
# Recasting Campus Grids as Platform of Services

Not sure who coined the Connect name.. Apologies if this diverges from the intended meaning

## "OSGConnect.org" (some traveling name)




- A very simple story
- Define a suite of service offerings for campuses
- Connecting science to resources with increasing capability



# Campus Infrastructures Maturity Model

| CGMM Level | Characteristics of organizations at this level | INCREASING CAPABILITIES |
|---|---|---|
| 1 | No organized or coordinated campus grid effort. Pockets of research computing, typically funded by individual researchers. Little or no support or documentation. | None |
| 2 | Some localized organization around campus grids. Some resource sharing at the departmental or college level. Minimal support and documentation. | BOSCO campus usage |
| 3 | Campus wide organization and/or broad visibility to campus grids. Good examples of resource sharing exist and there is some ability to utilize resources outside of the campus via partner campuses or the open science grid. Some documentation and local personnel support for campus grid users. | + OSG Connect : Jobs |
| 4 | Campus wide organization or visibility of campus grids initiatives. Widespread sharing of on and off campus resources. At least part time dedicated personnel support and some documentation for campus grid users. | + OSG Connect : Software & Data access |
| 5 | Campus grids are a 'way of life' for campus researchers, with on and off campus resources sharing the default. Mature user-facing documentation and dedicated personnel support for campus grid users. | + OSG Connect : Integrated Accounting & Informatics |

https://twiki.grid.iu.edu/bin/view/CampusGrids/DeployedCampusInfrastructures

# Capabilities

- Foundational
  - Campus identity ( → federated, grid)
  - Job management over diverse resources
  - Ubiquitous software and data access
  - Monitoring and accounting services
- Practical
  - Application best practices on d-HTC
  - Advanced workflow services
  - Advanced user interfaces

# What are the elements?

- Graduated Platform of Services
  - Campus Engagement & Identity Integration tools
  - Job management: BOSCO and its extensions + pure HTCondor
  - Distributed software access (OASIS, PALMS, PARROT)
  - Distributed data access (SRM, XRD, HTTP, SKELETONKEY)
  - Accounting and Informatics services for cycle sharing (GRATIA, CIVAIS)
- Campus Infrastructures Community
  - Forum, meetings, context to drive adoption, gather feedback, register impact
  - Tutorials, demonstrators, campus blueprints, engagements

# Extra slides

# SkeletonKey

- Motivated by UC3 users needing to have an easy way to remotely access their software and data on clusters around campus
- Designed as an easy alternative for users to manually using Parrot/Chirp
- Utilizes Parrot/Chirp but provides an easy to use configuration file

# Design Goals

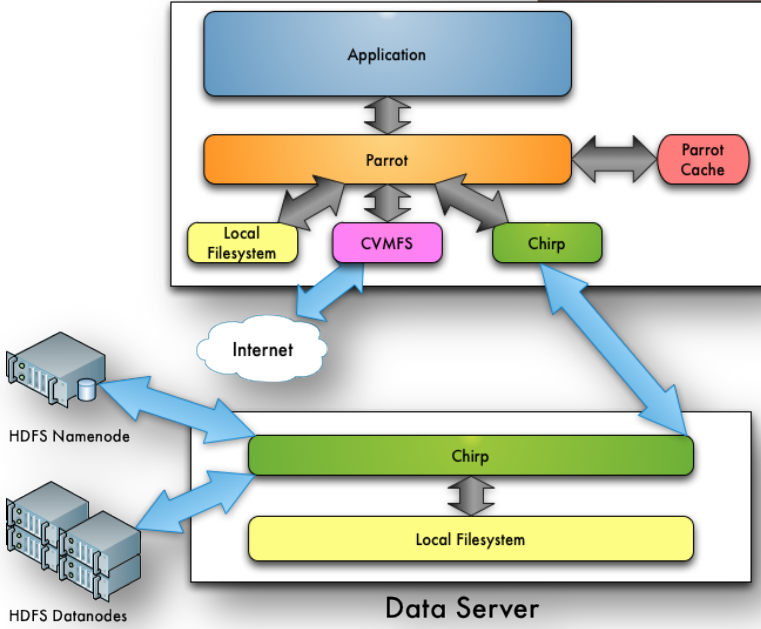
- Provide an easily understood way for users to incorporate remote software/data access into their current workflows
- Allow users to expand their computations to incorporate opportunistic usage of other campus clusters
- Eventually, allow users to expand from campus grids to using OSG opportunistically

# Current Work

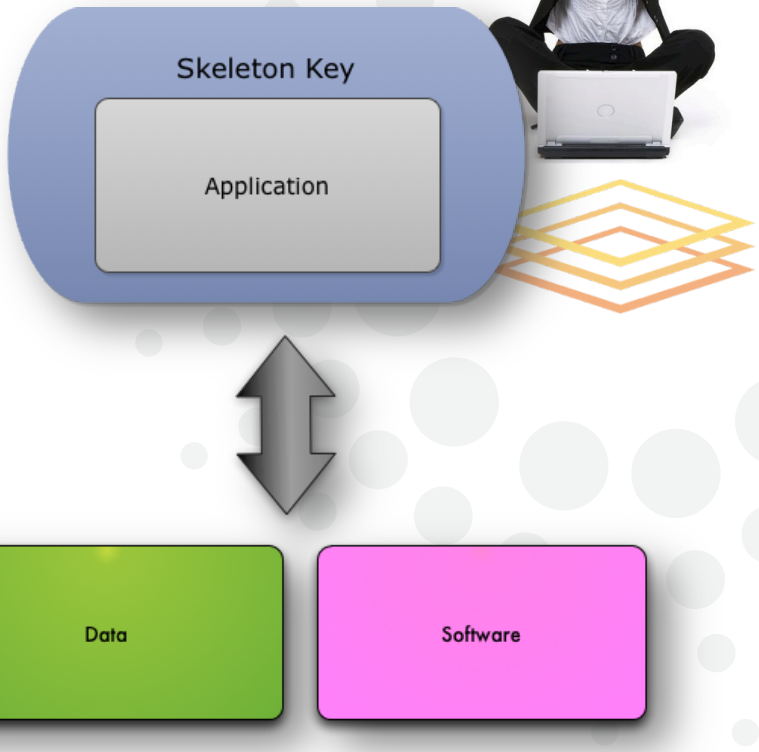
- Have initial ‘version one’ implementation
- Working with three groups to incorporate SkeletonKey in their workflows and actively utilize campus grid environment at U of C
- Incorporating user feedback into current code and updating features based on user needs

# SkeletonKey

## Directly




## With SkeletonKey



# PALMS project

- OASIS:
  - provides the infrastructure to host the software in CVMFS but users need more guidance to install the software (1) and to access it from OSG resources (2)
- Programs, Applications and Libraries Management and Setup (PALMS)

  - A system to install and manage software in OASIS
  - Simplifies the packaging and installation of different versions for different platforms
  - Helps users to setup the correct and desired environment (applications and libraries)

# PALMS software manager features

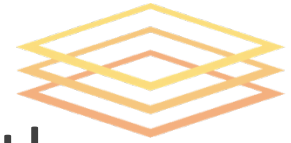
- Help packaging application and deploying it on OASIS (or into any CVMFS stratum)
- Allow installs, updates and removals of applications and libraries
- Allow multiple versions for distinct platforms
- Allow multiple versions for the same platform
- Does not require root on the OASIS server
- Can manage and solve dependencies and conflicts
- Help adapting and installing native packages

# PALMS user features

- Help select the correct version for the platform
- Provide a default version but allow to choice
- Setup the correct environment for the user shell
- Work automatically with different shells
- Add no performance penalty compared to default OASIS

# PALMS activities

- Project planning
  - Presentations and white paper
- Software development, packaging and documentation
- Deployment on OSG OASIS and on UC3
- Librarian (software manager) activities for the OSG VO



# CIVAIS project

- There is a lot of information about the operation of a Campus Infrastructure or OSG
- Processed information is more valuable than raw data
- Data and info differ by role (researcher, PI, computer center director, funding program manager, network administrator, …)
- CIVAIS: Campus Infrastructures Visual Analytics and Informatics Services
  - A analytics service collecting information form a Campus Infrastructure
  - Provides clear, concise and flexible views for users
  - And an open data platform (policy based) to stimulate derived metrics and 3rd party apps for advanced analytics

# Use Case Example (1)

- What do computing center executive/steering committees most want to know?
  - How are resources being used
  - Are they serving investing stakeholders fairly, as well as the broader university community
  - Is capacity meeting demand
  - Which technologies (processing, storage, network, visualization) are most likely to yield the most benefit to the most users
  - How do we judge the effectiveness of resource usage for advancing the scientific goals of our stake-holding partners

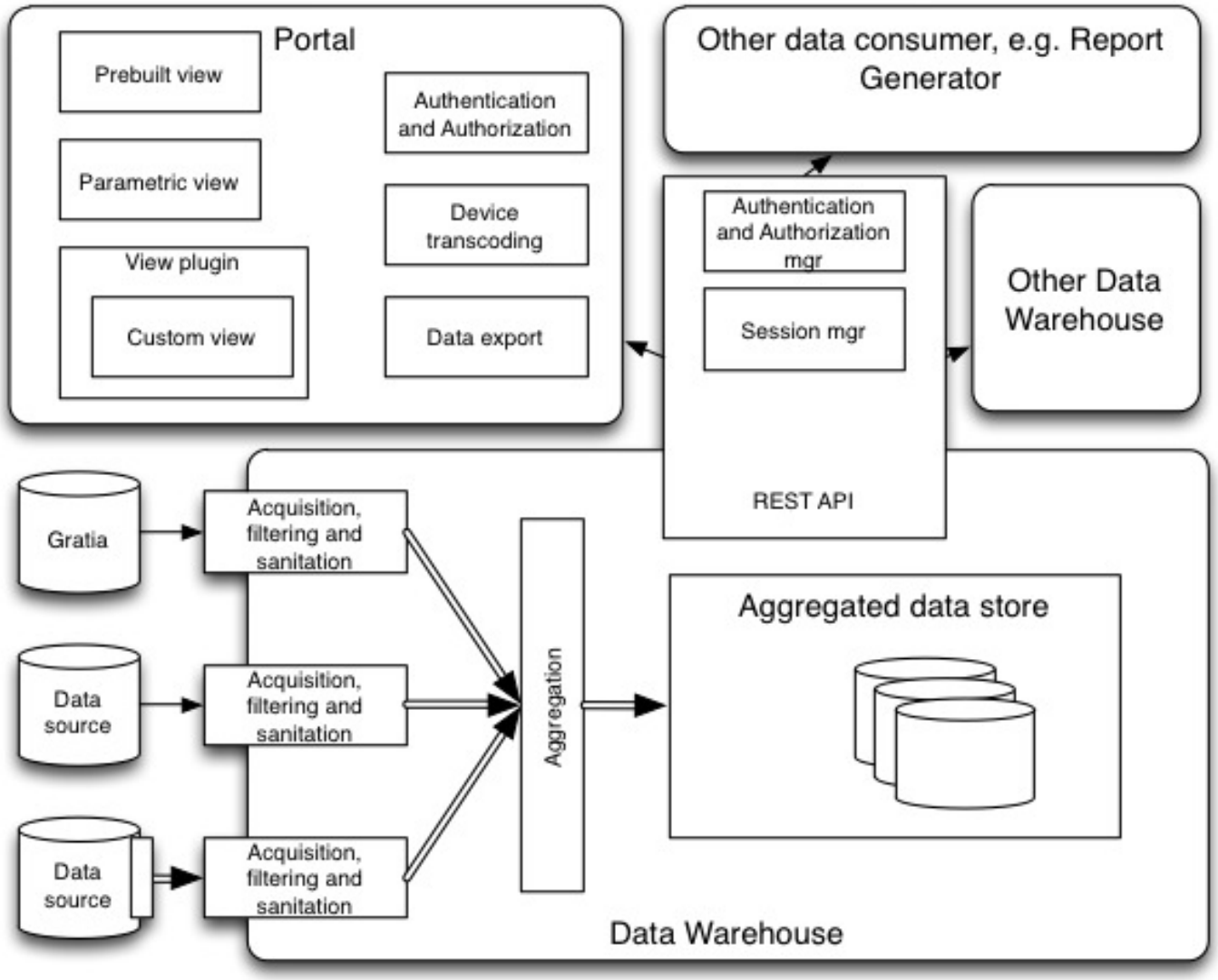
# Use Case Example (2)

- What does the OSG Executive Team most want to know?
  - How effective is the campus program in engaging new users and research communities on campuses
  - Which disciplines, outside of high energy physics, have received benefit from OSG
  - What are the impacts of OSG services and technologies on accelerating the scientific mission of our stake-holding organizations as well as the national

# CIVAIS key features

- Design starting from the user experience
- Multiple roles determining access policies and interests
- Interactive extensible web displays tailored to the role of the user
- Designed for a Campus Infrastructure
- Easy to install and deploy on a Campus
- Hierarchical reporting for a wider community (e.g. OSG CIC)
- Highly scalable
  - a single Campus reporting running on a single machine, bigger and more complex structures scaling on a distributed architecture
- Pluggable and open interface
  - Accepting multiple inputs (Gratia, message queues, etc…)
  - Ability to add custom views to the display
  - Open Data available via RESTful API

# CIVAIS architecture diagram



# CIVAIS architecture highlights

- Highly scalable and reliable data warehouse
- Multiple data inputs including Gratia server and probes and Google documents and Web forms
- Message bus for flexible and reliable communication (double arrows in the diagram)
- RESTful API for controlled data access
- Multiplatform portal using HTML5 and vector graphics for viewing, browsing and exporting data
- Standard plug-in definition for both data input and viewer extension

# CIVAIS activities

- Project planning
  - Presentations and white paper
- Project mock-ups and evaluation
- Software development, packaging and documentation
- Deployment and testing on UC3