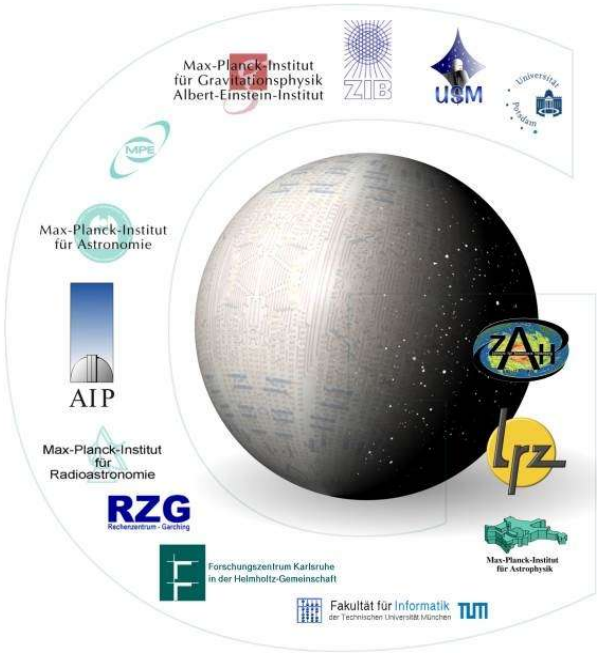


A Cactus User Portal with GridSphere

Thomas Radke
Max Planck Institute for Gravitational Physics
Albert Einstein Institute (AEI)

Astro-RG Workshop @ OGF20
9 May 2007, Manchester, U.K.

Astrophysical Background

- AEI is one of the leading research institutes in the international gravitational wave astronomy community
- Scientists in AEI's Astrophysical Relativity division study
  - the detection and measurement of GWs using laser interferometry (GEO600, LIGO, LISA)
  - the prediction of corresponding waveform templates from simulated astronomical events of GW emission, using *Cactus* as computational framework for their numerical simulations
- the eScience group at AEI supports the physicists in their every-day research by providing taylored software solutions using modern CS and Grid technologies

Cactus Metadata Management

- Users run many Cactus jobs on various different resources
  - What simulations were done in the past, by me and my collaborators ? Which parameters were used ?
  - What simulations are currently running where ?
  - Where are the output files and other data for a specific simulation stored ?

  ➔ A *Cactus metadata management system* would help scientists to better keep track of all their simulations.
- Cactus code development is very complex

  ➔ An *automated integration tests mechanism* would help Cactus programmers to better maintain the toolkit.

The AstroGrid-D Approach

1. Definition and flexible description of application-specific Cactus metadata
2. Collection of Cactus metadata
   - from within running simulations
   - from remote testing/validation procedures such as automated Cactus integration tests
3. Metadata management by an information service with persistent datastore backend and query engine
4. Querying, retrieval, and presentation of metadata to the end-user through an integrated portal user interface

Description of Cactus Metadata

- to generically describe Cactus simulation metadata we use the *Resource Description Framework (RDF)*
  - W3C standard technology for building Semantic Webs
  - extensive set of powerful web service tools and open source development libraries available
- RDF tools must agree on a vocabulary to process metadata
- various RDF vocabularies already exist to describe certain standard types of metadata classes
- for our application-specific metadata we had to extend those and defined a *Cactus metadata RDF schema*

Collecting Cactus Metadata

- Cactus has code modules (thorns) which transparently
  1) collect metadata from the running simulation
     - predefined static simulation metadata gathered at startup
     - user-defined dynamic metadata periodically updated during runtime
  2) generate an intermediate RDF/XML document from it
  3) send it off to an external *Stellaris* information service
- metadata collected by a *Cactus integration tests* Perl script
  - analyses the output from individual Cactus unit tests: `checkout, configure, build, build-utils, testsuites`
  - installed as nightly cron job on selected production machines

A Cactus Portal with GridSphere

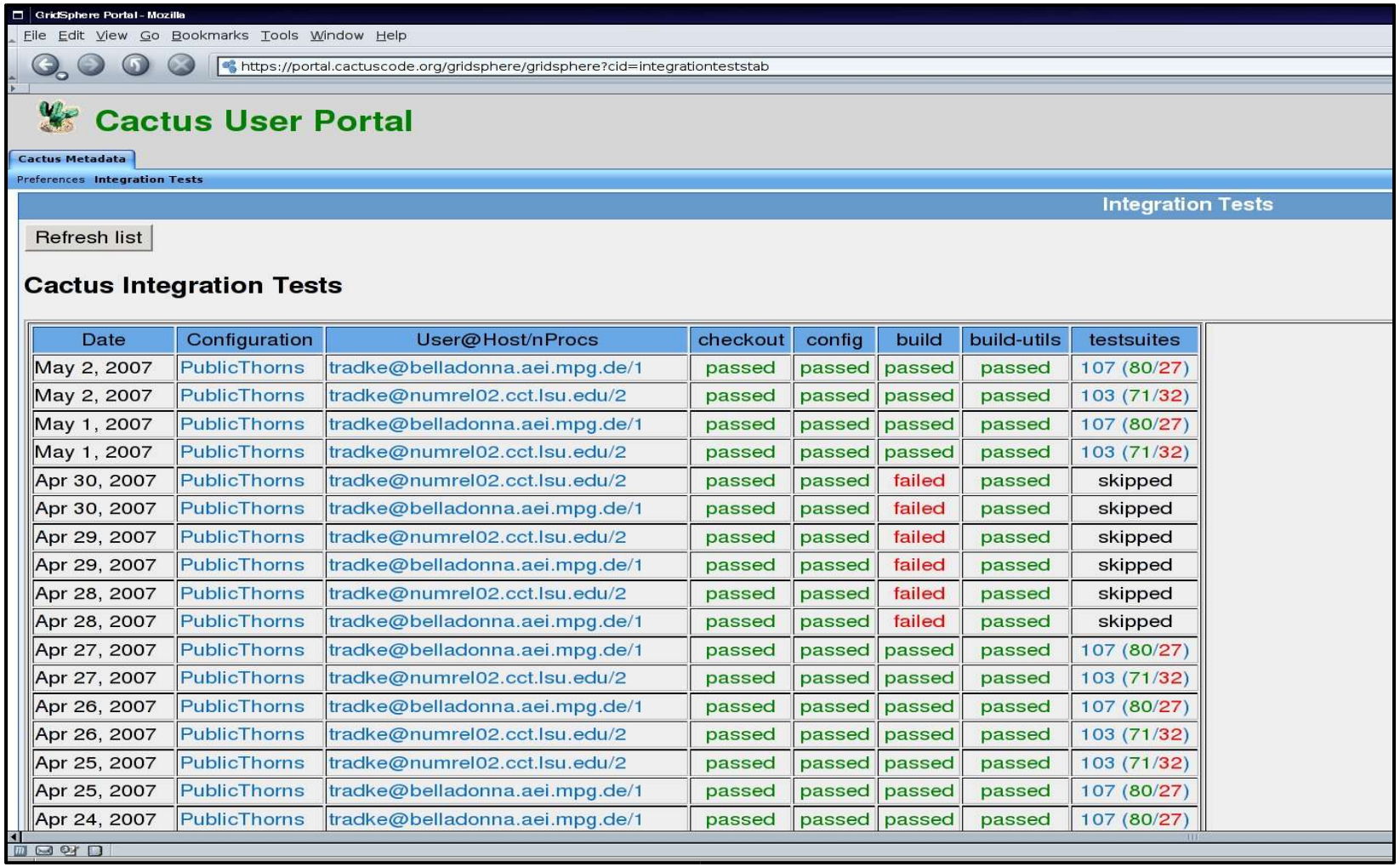
A *Cactus User Portal* was developed to provide an integrated end-user interface to query, retrieve, and present simulation metadata. It can be accessed with any standard web browser.

- implemented with the *GridSphere portal framework*
  - widely used framework; 100% open source; maintained at AEI
  - compliant with JSR 168 standard for web portlet development
    - portlet interoperability between different portal servers
- a built-in set of core portlets provides standard functionality
  - secure user authentication (eg. with Grid certificates)
  - data persistence and session management
  - personalised user profiles (themes, subscription to portlets)

Metadata Presentation in the Portal

- For Cactus, a new portlet was written for *GridSphere* which
  - dynamically generates a SPARQL query to search for a specific set of Cactus simulation metadata (Java)
  - contacts an external *Stellaris* information service and retrieves the results from a query (JENA framework)
  - presents the query results in different interlinked views (JSP)
- users can interact with the portlet and set individual preferences to refine their metadata queries
  - search for simulations by configuration name, user, execution host (all parameters given as regular expressions)
  - limit the list of returned query results to the first N entries

Automated Cactus Integration Tests



Date	Configuration	User@Host/nProcs	checkout	config	build	build-utils	testsuites
May 2, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	passed	passed	107 (80/27)
May 2, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	passed	passed	103 (71/32)
May 1, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	passed	passed	107 (80/27)
May 1, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	passed	passed	103 (71/32)
Apr 30, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	failed	passed	skipped
Apr 30, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	failed	passed	skipped
Apr 29, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	failed	passed	skipped
Apr 29, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	failed	passed	skipped
Apr 28, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	failed	passed	skipped
Apr 28, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	failed	passed	skipped
Apr 27, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	passed	passed	107 (80/27)
Apr 27, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	passed	passed	103 (71/32)
Apr 26, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	passed	passed	107 (80/27)
Apr 26, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	passed	passed	103 (71/32)
Apr 25, 2007	PublicThorns	tradke@numrel02.cct.lsu.edu/2	passed	passed	passed	passed	103 (71/32)
Apr 25, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	passed	passed	107 (80/27)
Apr 24, 2007	PublicThorns	tradke@belladonna.aei.mpg.de/1	passed	passed	passed	passed	107 (80/27)

Summary and Outlook

- In AstroGrid-D we developed a *Cactus user portal* to help scientists in managing their Cactus simulations.
- We make use of standard Web and Grid technologies
  - AstroGrid-D's *Stellaris* information service
  - the *GridSphere* portal framework

  but hide most of their complexity from the end user.
- An automated *Cactus Integration Tests* procedure was deployed as a production-mode portal application.

in the near future: integration of other services

- job monitoring tools, data archiving systems
- application-specific online monitoring/steering services

Links to Further Information

- Cactus User Portal
  - public user portal with guest user account
    `https://portal.cactuscode.org`
- Numerical Relativity Portal
  - production portal for the NumRel groups at AEI and CCT
    `https://portal.aei.mpg.de`
- Cactus homepage: `http://www.cactuscode.org`
- GridSphere homepage: `http://www.gridsphere.org`
- AstroGrid-D homepage: `http://www.gac-grid.org`