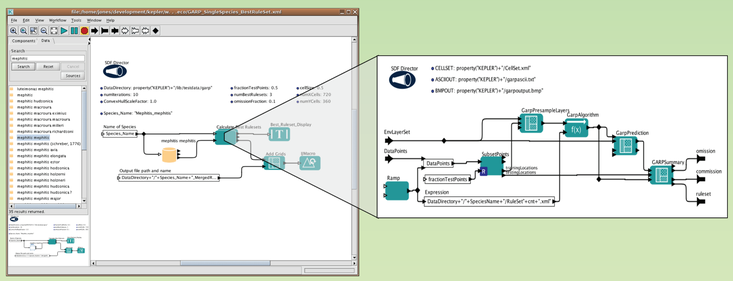
Promoting Community Contributions with Highly Configurable Component Based Software
A Kepler Architecture

Aaron Schultz, Matthew Jones, Timothy McPhillips, Sean Riddle

Overview

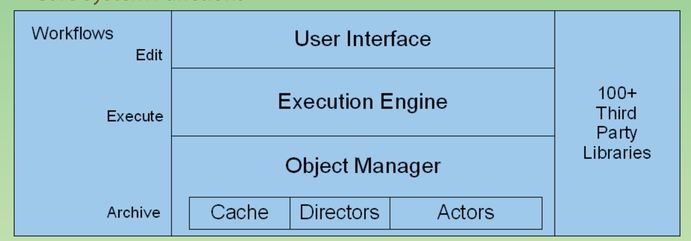
Kepler allows scientists to utilize a wide variety of data stores and analysis tools through graphical editing and execution of scientific workflows. As Kepler popularity grows a more modular architecture for providing and configuring application functionality is needed. Community contribution and user experience will benefit from adopting a well supported standards based approach for application componentization.



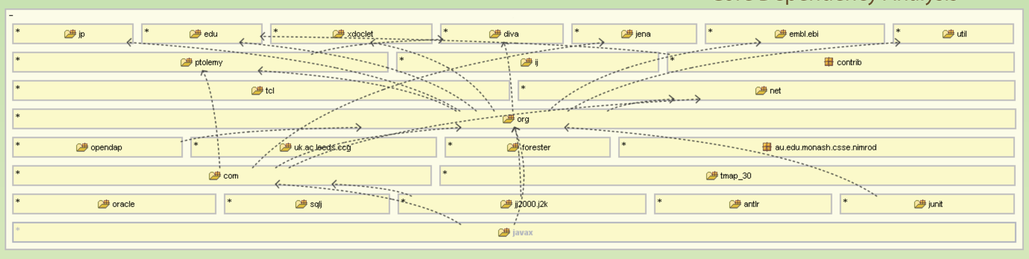
Existing Architecture

The core Kepler software provides a graphical user interface, execution engine and object manager for edidting, executing and archiving scientific workflows. The object manager provides caching of actors and data greatly enhancing run time efficiency. Many third party jars are used to provide system functionality and many projects have contributed actors and functionality to the Kepler application.

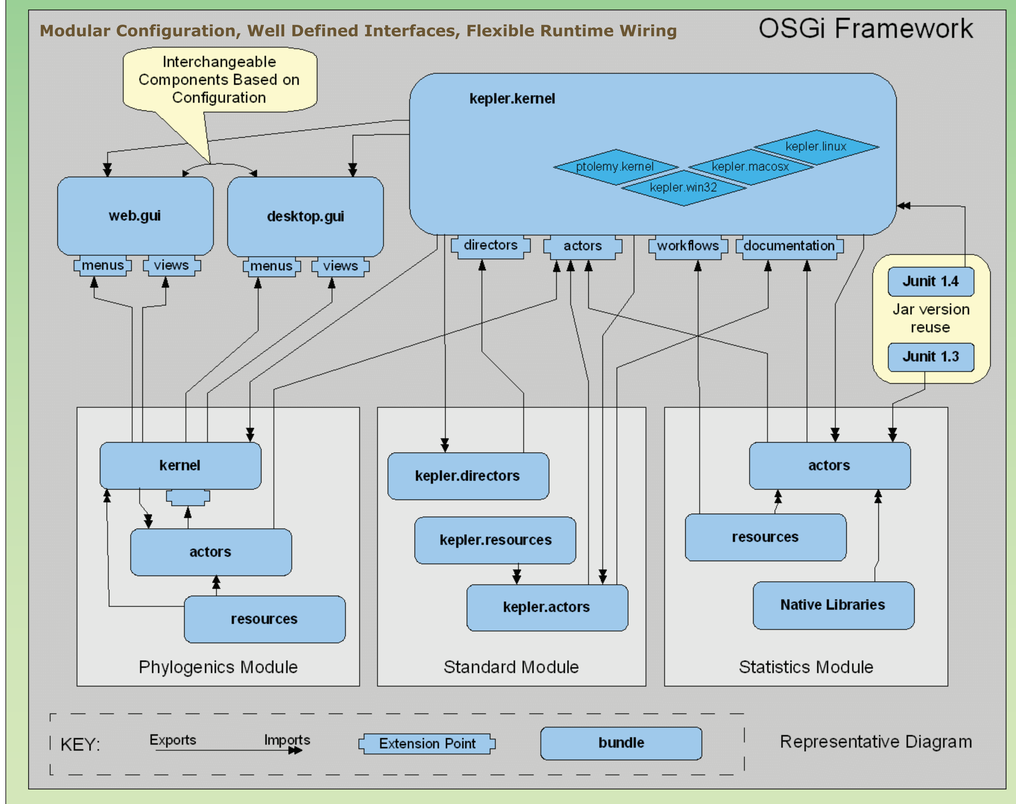
Core System Functions



Core Dependency Analysis



Existing Jar and package dependencies have been analyzed and converted to OSGi standard dependency format for experimentation. A modular approach will require more visibility and attention to dependency management.



Proposed Architecture

Configurable standards based components with well defined interfaces are wired together using the OSGi standard for Java modularization and extension.

Allows scientists to easily
- find and add functionality
- remove functionality
- configure functionality

Allows developers to easily
- modify existing code
- create extensions
- package extensions
- version extensions
- share extensions

OSGi Technology

Open Services Gateway Initiative

OSGi is an open framework for connecting many independently developed Java packages in one virtual machine. Bundles are jar files that contain a manifest that follows the OSGi standard. Any OSGi compatible framework can then load, interconnect and cache these bundles at runtime to operate as one unified application. Bundles contribute to the application by extending well defined extension points and can use packages from other bundles by importing packages exported by those bundles. OSGi is the underlying technology used by the very popular and growing Eclipse IDE and has been under development since 1999.

Benefits:
- provides for parallel development by many distributed teams
- allows end users to easily manage functional configuration
- separates code namespaces
- allows for using multiple versions of the same library
- developer awareness of program architecture and dependencies
- open source, well developed, large user community

Drawbacks
- core developer overhead for managing dependencies
- core developers must be knowledgeable of OSGi technology
- additional infrastructure for managing and distributing bundles

Kepler is based upon work supported by the National Science Foundation under awards 0225676 for: SEEK, 0722079 for Kepler/CORE, 0619060 for REAP and by the Department of Energy under Contract No. DE-FC02-01ER25486 for SciDAC/SDM, and by DARPA under Contract No. F33615-00-C-1703 for Ptolemy.