The Problem

There are many data analysis packages for the different X-ray and neutron scattering techniques. These packages bring many challenges, most notably the challenges of: sustaining development and support, providing for and maintaining cross-platform deployment, utilizing multi-processor compute resources as and when necessary, and all the time trying to keep things simple for the end user! How much better it would be if the end user could have the applications they need, in one place, running on a machine that is up to the job, within an environment that is both familiar and gives structure to their workflow?

A Solution?

Here we introduce GenApp, a target-agnostic infrastructure for the creation and deployment of UIs for underlying executables which is also integrated with Apache Airavata, and illustrate its use to provide the SASSIE-Web framework for the constrained ‘atomatic’ modelling of macromolecular solution structures using SAXS/SANS and AUC data.

GenApp

GenApp is an open extensible multi-target application generation tool for the simple and rapid deployment of multi-scale scientific codes.

• An application is defined as a collection of executable modules which are then presented through a common user interface. This provides a powerful paradigm to combine both existing and new codes to perform novel workflows, or to develop new applications.

• The addition of a module in GenApp is simple, and only requires the writing of a short JSON wrapper (a module) to detail the input and output, and the editing of two JSON files, one to specify where the module should appear in the applications menu system (menu.json), and the other to specify how the application itself is to be presented (directives.json). The modules themselves can be written in any supported language, independent of the choice of the target GUI implementation. Separating the scientific code from the GUI in this way not only facilitates the linking of component modules into larger workflows and applications, but also reduces the burden in supporting legacy codes.

Module executables either run locally (most GUI applications) or, if web-based, on a web server or other resource configured within Apache Airavata.

GenApp facilitates the creation of applications as web servers or gateways. This includes remote file management and the execution and management of lengthy non-interactive jobs. The latter capability, provided through integration with Apache Airavata, allows GenApp applications to harness a range of high-performance computing resources including local clusters, supercomputers, national grids, academic and commercial clouds. Instances of GenApp web applications have been tested on XSEDE and AWS.

SASSIE-Web

The aim of SASSIE-Web is to allow experimentalists (especially novice end users) to construct modelling workflows from a set of sophisticated simulation and analytic modules and run them transparently on centrally-managed HPC resources using nothing more than a web browser.

The provision of a web interface avoids the need for end users to install and maintain large complex software on their own machines, and remote execution accelerates the computationally expensive steps of the modelling process.

The SASSIE-Web menu organizes the analysis workflow in six sets of GenApp modules.

- Tools: which includes utilities to predict scattering length densities, interpolate experimental data files when required, and extract or merge macromolecular structures;
- Build: which includes utilities to check and correct PDB-formatted coordinate files;
- Interact: which provides the 2D molecular viewer for interactive display of a specified structure;
- Simulate: which provides MC/MD programs to create representative ensembles of trial structures for testing against the data;
- Calculate: which provides a range of scattering curve calculation utilities;
- Analyse: which identifies the simulated structures that best-fit the experimental scattering data and provides for their visualisation as ensembles.

We anticipate that GenApp will be useful to generate a wide-range of scientific applications beyond the scope covered by the CCP-SAS project. Interested parties should contact: genapp-devel@biochem.utexas.edu

Reference


Sassiozzi, P. and Beltranmich, 2012) packages have already been deployed for ‘alpha’ testing as web applications hosted on our CCP-SAS server. Anyone interested in contributing to the other relevant applications should contact: joseph.burton@nih.gov