Application Generation with GenApp

Emre H Brookes

Science Gateway Seminar

1 May 2015
Outline

- CCP-SAS
- GenApp Framework
- Future tasks
Outline

- CCP-SAS
- GenApp Framework
- Future tasks
Paul Butler (PI)  
University of Tennessee Knoxville  
NIST Center for Neutron Research  
Emre Brookes  
University of Texas Health Science Center San Antonio  
Jianhan Chen  
Kansas State University  
Joseph Curtis  
NIST Center for Neutron Research  
Tom Irving  
Advanced Photon Source  
+ other collaborators

CCP-SAS Project
http://www.ccp-sas.org

Stephen Perkins (PI)  
University College London  
David Barlow  
Kings College London  
Karen Edler  
University of Bath  
Richard Heenan & Steve King  
ISIS Pulsed Neutron & Muon Source  
David Scott  
Nottingham University  
Nick Terrill  
Diamond Light Source  
+ other collaborators

Collaborative Computational Project  
for advanced analyses of structural data in chemical biology and soft condensed matter

Grant No.  
SI2/CHE-1265821

Grant No.  
EP/K039121/1
CCP-SAS

**Collaborative Computational Project** for advanced analyses of structural data in chemical biology and soft condensed matter

**SI2-CHE** cyberinfrastructure project addressing Grand Challenges in the Chemical Sciences

**Small Angle Scattering**

Lab X-ray ($\lambda=0.1 - 0.2$ nm)

Neutron reactor ($\lambda=0.1 - 1$ nm)

Storage Ring ($\lambda=0.03 - 0.4$ nm)

Brookes, E. H.  
Science Gateway Seminar – 1 May 2015
Biological solution SAS data processing

**Guinier Plot**
- Overall size: $R_g$
- Molecular weight: $I_0$
- Aggregation, Hydration, Ion distribution, etc

**Kratky Plot**
- Folded / unfolded conformation

**PDDF**
- Structural Info in Real Space: $D_{max}$, Shape, $R_g$, etc

**SAXS**
- Bead-/DR-model Structure

**Rigid-body Modeling**
- Complexes Reconstructed Using sub-units w/ Known Structures

**Global Restraints**
- Before
- After High Resolution Structures with Accurate Global Shape

**Low Resolution Structures**

CCP-SAS

- SASSIE http://www.smallangles.net/sassie
  - Joseph Curtis et al.
  - PYTHON
  - includes wrapped binary executables
- SCT/SCTPL/HYDRO http://www.ucl.ac.uk/smb/perkins
  - Steve Perkins et al.
  - Structural Immunology Group at University College London
  - FORTRAN
- US-SOMO http://somo.uthscsa.edu
  - Emre Brookes et al.
  - C++/Qt
  - includes wrapped binary executables

attract others …

Considerations

- **Ease of deployment** in an ever-evolving software environment landscape
- Legacy and frequently specific lab developed codes
- Labs frequently cannot afford a dedicated software team nor the cost in time and funding
An Open Extensible Multi-Target Application Generation Tool for Simple Rapid Deployment of Multi-Scale Scientific Codes

Brookes, E. H. XSEDE 14 Atlanta

- JSON
- Global directives & application menu and configuration
- Module I/O definitions
- Module messaging
- Language assembly descriptions
Outline

- CCP-SAS
- GenApp Framework
  - Framework user perspective
  - Execution model & messaging
  - File management
  - Job management
  - Brief demo (if time & network allows)
- Future tasks
Outline

- CCP-SAS
- GenApp Framework
  - Framework user perspective
  - Execution model & messaging
  - File management
  - Job management
  - Brief demo (if time & network allows)
- Future tasks
extremely basic instructions to wrap an application

- from your base directory
  - edit directives.json
  - edit menu.json
  - edit modules/*.json
  - create application executable json wrappers in bin/
    - NB: executable **must** only output valid json to stdout
  - run $ genapp.pl
  - html output should be available via web interface
    - eg
      - [http://somo.chem.utk.edu/genpptest](http://somo.chem.utk.edu/genpptest)
  - if you get forbidden errors in the web browser, make sure everything is group readable by user apache

Credit: Joseph Curtis – 1st User
Directives

```
{
    "title" : "genapptest",
    "application" : "genapptest",
    "version" : "2.0",
    "languages" : [ "html5", "qt3", "qt4", "qt5" ],
    "footer" : "genapp environment",
    "footersize" : "2.5rem",

    "executable_path" : {
        "html5" : "/share/apps/genapp/genapptest/bin",
        "qt3" : "/home/ehb/genapptest/bin",
        "qt4" : "/home/ehb/genapptest/bin",
        "qt5" : "/home/ehb/genapptest/bin",
    },

    "docroot" : {
        "html5" : "/share/apps/genapp",
        "qt3" : "/tmp",
        "qt4" : "/tmp",
        "qt5" : "/tmp",
    },

    "helper" : {
        "txt" : "emacs",
        "pdb" : "rasmol"
    },

    "appconfig" : "/home/ehb/genapptest/appconfig.json"
}
```


Application configuration

```json
{
    "mail" : {
        "admin" : "someone@someplace.net",
        "from" : "someotherplace.net",
        "feedback" : "somefeedback@someplace.net",
        "smtp" : {
            "host" : "some-smtpserver.domain.net",
            "user" : "some-user@somewhere.net",
            "password" : "encrypted-pw"
        }
    },
    "hostip" : "160.36.200.43",
    "messaging" : {
        "wsport" : 8080,
        "zmqhostip" : "1.2.3.4",
        "zmqport" : 12342,
        "udphostip" : "1.2.3.4",
        "udpport" : 12343
    },
    "resources" : {
        "local" : "",
        "compute0" : "ssh compute-0-0",
        "compute1" : "ssh compute-0-1"
    },
    "resourcedefault" : "local",
    "submitpolicy" : "login"
}
```
Menu Definition
{
  "menu": [
    {
      "id": "tools",
      "label": "Tools",
      "icon": "pings/tools.png",
      "modules": [
        {
          "id": "center",
          "label": "Center",
        },
        {
          "id": "align",
          "label": "Align",
        },
        {
          "id": "data_interpolation",
          "label": "Data Interpolation",
        }
      ]
    },
    {
      "id": "build",
      "label": "Build",
      "icon": "pings/build.png",
      "modules": [
        {
          "id": "build_1",
          "label": "Build 1",
        },
        {
          "id": "build_2",
          "label": "Build 2",
        }
      ]
    }
  ]
}
Menu Definition

```json
{
  "menu": [
    {
      "id": "tools",
      "label": "Tools",
      "icon": "pngs/tools.png",
      "modules": [
        {
          "id": "center",
          "label": "Center",
        },
        {
          "id": "align",
          "label": "Align",
        },
        {
          "id": "data_interpolation",
          "label": "Data Interpolation",
        }
      ]
    },
    {
      "id": "build",
      "label": "Build",
      "icon": "pngs/build.png",
      "modules": [
        {
          "id": "build_1",
          "label": "Build 1",
        },
        {
          "id": "build_2",
          "label": "Build 2",
        }
      ]
    }
  ]
}
```
Center value 1: 0
Minutes to run: 1
Message results: checked
Results interval in seconds: 15
Message text:
Output value 1:
Progress:

genapp environment
HTML5/PHP and Qt/C++ versions side-by-side

\[ N = \frac{D_{\text{MAX}} - D_{\text{MIN}}}{\pi} \]
2D Plot
HTML5/PHP: Flot
Qt/C++: Qwtplot

Molecular Viewer
HTML5: JSmol
Qt/C++: Multiple options (RASMOL, Jmol, others)
## Module fields

<table>
<thead>
<tr>
<th>Field</th>
<th>Input support</th>
<th>Output support</th>
<th>Append (output-only primarily for messaging)</th>
<th>Default (input-only)</th>
<th>Min (input-only)</th>
<th>Max (input-only)</th>
<th>Step (input-only)</th>
<th>Required (input-only)</th>
<th>Repeater (input-only)</th>
<th>Help (example)</th>
<th>Notes</th>
<th>Genapptest example</th>
</tr>
</thead>
<tbody>
<tr>
<td>integer</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>html5-only</td>
<td>yes</td>
<td>html5-only</td>
<td>html5-only</td>
<td>tools: center</td>
<td>repeat example interact: interact_1 step example tools: center build: build_1</td>
</tr>
<tr>
<td>float</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>html5-only</td>
<td>yes</td>
<td>no</td>
<td>html5-only</td>
<td>tools: center</td>
<td>build: build_2</td>
</tr>
<tr>
<td>text</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>html5-only</td>
<td>build: build_1</td>
<td></td>
</tr>
<tr>
<td>textarea</td>
<td>yes</td>
<td>yes</td>
<td>html5-only</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>html5-only</td>
<td>multiline default input not yet supported build: build_2</td>
<td></td>
</tr>
<tr>
<td>checkbox</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>html5-only</td>
<td>html5-only</td>
<td>additional 'checked' attribute build: build_2 repeat example calculate: calculate_1</td>
<td></td>
</tr>
<tr>
<td>radio</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>html5-only</td>
<td>build: build_2</td>
<td></td>
</tr>
<tr>
<td>listbox</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>html5-only</td>
<td>html5-only</td>
<td>no required attribute, but it will always have a value row spacing minor layout issue on repeater build: build_2 repeat example calculate: calculate_2</td>
<td></td>
</tr>
<tr>
<td>plot2d</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>html5-only</td>
<td>tools: data_interpolation</td>
<td></td>
</tr>
<tr>
<td>atomicstructure</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>html5-only</td>
<td>tools: data_interpolation more info</td>
<td></td>
</tr>
<tr>
<td>file</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>html5-only</td>
<td>additional multiple attribute (input-only) simulate: simulate_1</td>
<td></td>
</tr>
<tr>
<td>label</td>
<td>html5-only</td>
<td>html5-only</td>
<td>no</td>
<td>yes*</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>html5-only</td>
<td>*can be header 1, 2, 3, 4 or blank interact: interact_1</td>
<td></td>
</tr>
<tr>
<td>hidden</td>
<td>html5-only</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>this is primarily to pass a value in json to the executable analysis: pm</td>
<td></td>
</tr>
<tr>
<td>html</td>
<td>no</td>
<td>html5-only</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>allows html output none currently, but it's a typical field layout tools: center</td>
<td></td>
</tr>
<tr>
<td>progress</td>
<td>no</td>
<td>html5-only</td>
<td>no</td>
<td>no</td>
<td>yes (for output)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>progress bar default ranges is 0 to 1.0, setting max to a positive value will make the range 0-max tools: center</td>
<td></td>
</tr>
</tbody>
</table>
Module definition

```
"moduleid": "center",
"label": "Center",
"executable": "center",
"submitpolicy": "all",
"fields": [
  {
    "role": "input",
    "id": "input1",
    "label": "Center value 1",
    "type": "float",
    "default": 0.0,
    "min": 0.0,
    "max": 1500,
    "step": 0.01,
    "required": "true"
  },
  {
    "role": "input",
    "id": "runminutes",
    "label": "Minutes to run",
    "type": "float",
    "default": 1,
    "min": 0,
    "max": 10000,
    "required": "true"
  },
  {
    "role": "input",
    "id": "usemessage",
    "label": "Message results",
    "type": "checkbox",
    "checked": "true",
    "repeater": "true"
  },
  {
    "role": "input",
    "id": "messagesecs",
    "label": "Results interval in seconds",
    "type": "integer",
    "default": 15,
    "min": 1,
    "max": 1000,
    "repeat": "usemessage"
  }
]
```

Executable Input

```
{
  "input1": "0",
  "runminutes": "1",
  "usemessage": "on",
  "messagesecs": "15",
  "_logon": "emre",
  "_project": "",
  "_window": "d67fa871-46ea-4af3-d384-72bf0a82de3b",
  "_uuid": "fde73ea0-a2b4-11e4-ae42-5189a0917617",
  "_base_directory": "/share/apps/genaapp"
}
```

Executable Output

```
{
  "udphost": "160.36.200.43",
  "udpport": 37779,
  "resource:default": "local"
}
```

Outline

- CCP-SAS
- GenApp Framework
  - Framework user perspective
  - **Execution model & messaging**
  - File management
  - Job management
  - Brief demo (if time & network allows)
- Future tasks
\[ N = \frac{D_{\text{MAX}}(q_{\text{MAX}} - q_{\text{MIN}})}{\pi} \]
Enhanced execution model HTML5/PHP

Web Browser

Input JSON
AJAX
JSON ACK

WebSocket
JSON Message

Input JSON
"start module"

Monitor daemon

Input JSON

Module's Executable

JSON Output

AJAX
JSON Output

"get results"

Enhanced execution model HTML5/PHP with Apache Airavata
Nadeem Anjum
Google Summer of Code / Mentor: Suresh Marru

Web Browser

Input
JSON
AJAX
JSON
ACK

WebSocket
JSON Message

PHP "start module"

Launch Job
Input JSON

Monitor daemon

Check Status

AIRAVATA

Module's Executable

Input
JSON

JSON Output

Time

Outline

- CCP-SAS
- GenApp Framework
  - Framework user perspective
  - Execution model & messaging
  - File management
  - Job management
  - Brief demo (if time & network allows)
- Future tasks
The preferred mechanism for data input and output within the framework is formatted in JSON.

When a module execution is submitted, a unique run directory is created for its execution.

Files associated with the submission are placed in this directory and made available to the executing module.

SASSIE code has cases where outputs from a previously executed module are expected to be available to a subsequent one.

Projects

- Persistent directories
- Project locking
- File management
Outline

- CCP-SAS
- GenApp Framework
  - Framework user perspective
  - Execution model & messaging
  - File management
  - Job management
  - Brief demo (if time & network allows)
- Future tasks
<table>
<thead>
<tr>
<th>Module</th>
<th>Project</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>tools/center</td>
<td>testproject1</td>
<td>2015 Jan 19 11:25:56 UTC</td>
<td>93.3%</td>
</tr>
<tr>
<td>tools/data_interpolation</td>
<td>testproject1</td>
<td>2015 Jan 19 11:07:42 UTC</td>
<td>2015 Jan 19 11:07:42 UTC</td>
</tr>
</tbody>
</table>

Refresh • Cancel • Clear Lock • Remove Job • Reattach •

Submit

Messages
Outline

- CCP-SAS
- GenApp Framework
  - Framework user perspective
  - Execution model & messaging
  - File management
  - Job management
  - Brief demo (if time & network allows)
- Future tasks
Resources

- Current infrastructure
  - University of Tennessee Knoxville
  - “Entropy” server 128 core, 256 GB ram, 8 Tesla K20m GPUs, Rocks OS
  - Running HTML5/PHP
- Indiana University Quarry nodes
  - Testing HTML5/PHP
  - Trac wiki with integrated subversion repository
- Multiple students working at NIST
- Successful first workshop utilizing the infrastructure for SASSIE code on “entropy”
  - American Conference on Neutron Scattering 1 June 2014
- XSEDE TG-MCB140255 Computational support for small angle scattering for advanced analyses of structural data in chemical biology and soft condensed matter
- ORNL Titan 10% (pending)
Outline

- CCP-SAS
- GenApp Framework
  - Framework user perspective
  - Execution model & messaging
  - File management
  - Job management
  - Brief demo (if time & network allows)
- **Future tasks**
Future tasks

- New target languages (GSoC)
  - JAVA
  - Qt5 → iOS & Android
- Wrapping US-SOMO code into the framework (post-Doc)
- Integration with latest Airavata (GsoC)
- Advances in job management
- Apache Incubator
Thanks for listening

Questions: ask now or email me at emre@biochem.uthscsa.edu

Acknowledgments

- To E. Brookes
  - NIH NIGMS K25GM090154 / PI
  - NSF CHE-1265817 / PI
- NSF XSEDE TG-MCB140255
- UTHSCSA BCF
- To N. Anjum
  - Google Summer of Code

Funded by a grant from the National Institute of General Medical Sciences of the National Institutes of Health