Best Practices in CI Development: The Science Gateways Community Institute perspective

Compiled on behalf of the SGCI leadership team: Michael Zentner (mzentner@purdue.edu, Incubator), Marlon Pierce (marpierc@indiana.edu, Extended Developer Support), Maytal Dahan (maytal@tacc.utexas.edu, Scientific Software Collaborative), Katherine Lawrence (kathla@umich.edu, Community Engagement and Exchange), Linda Hayden (haydenl@mindspring.com, Workforce Development), Nancy Wilkins-Diehr (wilkinsn@sdsc.edu, Management)

BACKGROUND

In 2004, the NSF-funded TeraGrid project launched a program called science gateways, designed to increase accessibility of supercomputers to all scientists. In the years following, the term has acquired a much broader meaning; it now serves as a descriptor for all advanced web interfaces used for research, education and scholarship. These tools have profoundly changed the productivity of the research community, from students to authors publishing in leading journals.

Launched in 2016 as one of the NSF Office of Advanced Cyberinfrastructure's first two software institutes, the Science Gateways Community Institute (SGCI) [Ref. 1] helps researchers build and discover better, more functional gateways by developing, documenting, and disseminating best practices.

SGCI's design was informed by a series of studies from an EArly-concept Grant for Exploratory Research (EAGER) in 2009 to a conceptualization phase institute award in 2014. We are organized in six functional areas. The leads of each area form SGCI's leadership team. This position paper incorporates material presented in an article published in Computing in Science and Engineering in 2018 [Ref. 2].

INCUBATOR

Early studies identified the need for the variety of expertise needed to build effective gateways. Gateway developers often have a need for specialized expertise, but cannot afford to hire full time staff with that expertise. For example usability, graphic design, cybersecurity and quality assurance were all areas where projects wished they had expertise. SGCI offers short term consulting on these and other topics. Because our consultants advise many projects, they are able to bring those best practices developed over many years to our clients.

In a novel delivery of content, Michael Zentner, lead for the Incubator launched the SGCI Bootcamp in 2017. Bootcamp brings together gateway stakeholders (developers, PIs and others) for an intensive one week program where participants explore entrepreneurial topics including articulating their idea for a gateway, who benefits, where the gateway fits in the market, how to make it happen and how to sell the idea. Cohorts of bootcamp participants have formed over time and a new Brain Trust idea will bring together multiple people in rapid-fire problem-solving.

EXTENDED DEVELOPER SUPPORT

Modeled after the successful XSEDE Extended Collaborative Support Services (ECSS) program, SGCI's Extended Developer Support (EDS) program provides multi-month, in depth support to those looking to develop or enhance science gateways. While the ECSS program can provide support to researchers using XSEDE resources, SGCI's program can provide support to anyone developing a gateway, whether it uses local computing infrastructure, commercial clouds. We can also assist gateways that don't have a large computing footprint at all, but instead provide access to scientific instruments, streaming sensor data or citizen science projects.

SGCI developers have expertise with many different gateway frameworks and are able to use that expertise when evaluating projects for new clients. For example gateways with a significant compute component might be best served by a particular framework, those with an emphasis on collaborations by a different framework. This kind of best practices knowledge is critical in helping clients make important technology decisions and getting projects off on the right foot quickly.

SCIENTIFIC SOFTWARE COLLABORATIVE

SGCI has established a gateway catalog with nearly 600 entries that includes links to production gateways as well the software used to build them. The catalog is searchable by keyword (including tags indicating what was used to build the gateway) and organized by domain areas. It contributes to software reuse and the sharing of best practices in several ways. Those thinking of building a gateway can perhaps find

an existing one with the desired functionality. Or developers can find a gateway in another domain area where they like the look and feel and can contact the developer to use the same approach for their own work. 13 SGCI Affiliates agree to provide support for those interested in building gateways with their software, listed in the catalog.

COMMUNITY ENGAGEMENT AND EXCHANGE

Community Engagement and Exchange (CEE) connects the community through an annual Gateways conference, monthly webinars and a regular newsletter including project highlights and blog posts about topics of interest to the gateway community. The workshop series that evolved into the annual Gateways conference began in 2005. Through the years, developers from across domains have been very interested in sharing experiences with others developing science gateways. The conference also includes publishing opportunities so that these experiences can be shared more broadly. These events have spawned spinoff workshops in other regions, specifically the International Workshop on Science Gateways held in Europe since 2009. The groups publish annually in a joint journal special issue, where they review one another's work, furthering the exchange of best practices.

SGCI is also helping to change the environment in which these gateways are built on campuses, to promote the exchange of best practices there. Typically, developers are hired by principal investigators, in specific departments, to work on specific projects where they work in isolation. With this type of distribution, it is difficult to assemble a team with diverse expertise who are then available for assignment to a variety of projects throughout the organization. Even more important, talented developers feel a lack of job security, since funded projects in a department may come and go. A more centralized model, proven effective at universities such as Notre Dame, Indiana and Purdue University, allows much greater scalability and impact. SGCI is promoting a step-by-step approach to formulating such groups on campuses.

WORKFORCE DEVELOPMENT

SGCI's student programs provide summer opportunities for students to work side-by-side with SGCI staff on real world projects for clients. Exposure of students to staff members well-versed in best practices provides great training for future job opportunities. Besides the technical learnings, students observe how staff members approach projects and manage their time.

REFERENCES

[1] Lawrence, Katherine A., Michael Zentner, Nancy Wilkins-Diehr, Julie A. Wernert, Marlon Pierce, Suresh Marru, and Scott Michael. "Science gateways today and tomorrow: positive perspectives of nearly 5000 members of the research community," Concurrency and Computation: Practice and Experience 27, No. 16 (2015): 4252-4268.

[2] N. Wilkins-Diehr and T. Daniel Crawford, "NSF's Inaugural Software Institutes: The Science Gateways Community Institute and the Molecular Sciences Software Institute," Computing in Science & Engineering, vol. 20, no. 5, 2018, pp. 26-38. doi:10.1109/MCSE.2018.05329813.