OneSciencePlace: Sustainability Measures for a Science Gateway Framework

Sandra Gesing[1], David R. Benham[2] and the HUBzero© team
[1]University of Notre Dame, sandra.gesing@nd.edu
[2]Purdue University, dbenham@purdue.edu

The use of science gateways for research has majorly increased in the last decade, reflected in usage by research infrastructures such as XSEDE [1] as well as in the usage of science gateways frameworks such as HUBzero© [2]. One major concern for science gateways is their sustainability beyond an initial funding cycle [3]. HUBzero© is a well-used science gateway framework actively developed for over a decade. The needs of its community are one of the primary driving forces guiding the team behind HUBzero©. For example, requirements in the community led to the integration of JupyterHub [4], RStudio [5] and the provision of Docker containers [6] for the tool submission environment. Besides its community-driven requirements, the team behind HUBzero© continuously analyzes the existing science gateway landscape, the usage of instances of HUBzero© as well as trends in the usage of computational platforms in general to keep HUBzero© robust, scalable, and sustainable.

HUBzero© has begun development of a science gateway platform called OneSciencePlace (OSP). OSP is a SaaS-based (Software-as-a-Service) solution to give researchers a way to execute, share, and archive their research in a publicly accessible venue. HUBzero© hubs are often financed as part of a large grant. Researchers also write HUBzero© into their proposals and the HUBzero© team often provides boilerplate content for researchers to use in their proposals that indicate they have chosen HUBzero© as their cyber infrastructure platform. However, once funding for a project has ceased, all the work for the project is archived. Most often, without additional funding, the project archive has no external accessibility. The tools, discussions, wiki pages, and any other expertise housed on the platform ends up inaccessible to the community and the hub not maintained.

The HUBzero© team has seen this happen often and overcoming this loss of valuable information was the initial goal of OneSciencePlace (OSP). OSP was initially proposed as a long-term repository for HUBzero© hubs, so that their work could continue in some format going forward. Funding OSP was going to be primarily through surcharges on hubs interested in being archived at the end of their funding. As the project matured, the goal changed to OSP to become the primary platform on which HUBzero© offers its services in the future.

Recognizing the importance that sustaining single hubs contribute to the sustainability of HUBzero© as science gateway framework, OSP is a broad effort to build a cloud-based scientific tool sharing Software-as-a-Service (SaaS) platform addressing sustainability from different angles. We believe that following actions will majorly contribute to sustainability.

Part of OSP is envisioned as a **long-term community-hosting concept**. It offers researchers a community that lasts well beyond what is normally the traditional end of a research project. Currently, HUBzero© provides a single hosting approach for all communities that will co-exist with the community-hosting concept. The single hosting can be used as place to enable

cooperative efforts and cross-disciplinary research, giving researchers the confidence their work will not exist in an isolated silo, but in a robust flexible environment.

OSP will use **flexible financing models** to enable long-term sources of revenue to fund the infrastructure required to deploy and maintain research for long periods of time. Part of the solution are volunteer computing models similar to ones established by the BOINC to provide a free tier of access to OSP users as well as enabling community sponsored resource offerings. Paid services will also include necessary surcharges to provide for the longevity of the platform. Offering cyber infrastructure to small and midsize research efforts enables OSP to engage research efforts across the financial spectrum.

Extensive **resource usage tracking** will enable to track costs of consumed resources and careful analysis of usage patterns will enable OSP efficiently pool resources to control costs across the platform.

Cross-site embeddable tools on OSP will ensure tools developed on OSP are not trapped there. Once a tool is developed, a simple mechanism to embed the tool in other websites will enable users outside OSP to utilize tools and data on the platform.

Scalability in the OSP infrastructure is essential. The platform will utilize microservices and container technologies to help minimize costs when necessary while providing resources for large scale efforts.

Interoperability of OSP with existing toolsets is an essential way to offer the tools researchers need. OSP will take a develop-only-when-necessary approach, leveraging existing tool sets whenever possible and focus more of our development efforts on innovation.

References

[1] https://www.xsede.org/ecosystem/science-gateways

[2] McLennan, Michael, and Rick Kennell. HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering. Computing in Science & Engineering 12, no. 2 (2010): 48-53.

[3] Gesing, S., Zentner, M., Casavan, J., Hillery, B., Vorvoreanu, M., Heiland, R., Marru, S., Pierce, M., Mullinix, N., Maron, N. Science Gateways Incubator: Software Sustainability Meets Community Needs. 2017 IEEE 13th International Conference on e-Science (e-Science), Auckland, 2017, pp. 477-485. doi:10.1109/eScience.2017.77

[4] Fernando Pérez, Brian E. Granger, IPython: A System for Interactive Scientific Computing, Computing in Science and Engineering, vol. 9, no. 3, pp. 21-29, May/June 2007, doi:10.1109/MCSE.2007.53. URL: https://ipython.org

[5] https://www.rstudio.com/

[6] Dirk Merkel. 2014. Docker: lightweight Linux containers for consistent development and deployment. Linux J. 2014, 239, pages.