# HUBzero: A Platform For Scientific Collaboration

**NITEESH KUMAR AKURATI**[1,*]

[1] *School of Informatics and Computing, Bloomington, IN 47408, U.S.A.*
[*] *Corresponding authors: akuratin@indiana.edu*

*paper2, April 12, 2017*

**HUBzero is a cyberinfrastructure in a box that is used to create dynamic websites for educational activities as well as scientific research. HUBzero provides an platform where researchers can publish and share their research software and related material for educational purposes on the web. Other researchers will be able to access the tools as well as material using a Web browser and can also launch simulation runs on the national Grid infrastructure without having to download any code.**



**Keywords:** cyberinfrastructure, BigData, Simulation Tools, Collaboration

https://github.com/cloudmesh/sp17-i524/blob/master/paper2/S17-IR-2001/report.pdf

## 1. INTRODUCTION

HUBzero is a platform created to support nanoHUB.org, an online community for the Network for Computational Nanatechnology(NCN), which is funded by the U.S.National Science Foundation since 2002 to connect the theorists across the academia who develop simulation tools with the educators and the experimentalists who might use them. HUBzero has been expanded to more than 30hubs since 2002 and the hubs are growing every year into various fields ranging from development of assistive technology to volcanlogy[1].

"HUBzero is now supported by a consortium including Purdue, Indiana, Clemson and Wisconsin. Researchers at Rice, the State University of New York system, the University of Connecticut and Notre Dame use hubs. Purdue offers a hub-building and -hosting service and the consortium also supports an open source release, allowing people to build and host their own"[**?** ].

A hub web 2.0 functionality with an unique middleware, thereby providing a platform that is more powerful than an ordinary website. In a hub users will able to create, publish, share information, network and also have access to interactive visualization tools. The users of the hub ultimately collaborate to develop the educational community and make it a powerful resource. A hub can direct jobs to national resources such as TeraGrid, Purdue's DiaGrid as well as other cloud systems[2].

## 2. MIDDLEWARE

HUBzero mix of social networking and simulation is only possible with the help of unique middleware. HUBzero's middleware hosts the live session tools and makes it easy to connect the cloud computing infrastructure and supercomputing clusters to solve the large computational problems. The simulation tools runs on clusters of executions hosts hosted near the Web server and will be accessed by user's browser via VNC(Virtual Network Connection)[1].

Each user has a home directory which is unique to that particular user with access control, quota limitations and conventional ownership privileges. Each tool runs on a lightweight restricted virtual environment, that controls access to the networking, file systems and server processes. Tools are run with the privileges of a particular user[1].

Tools running on the hub will have a XII Window System environment. Each container runs a special X server which also acts as a VNC server is used to create the graphical session. The middleware of HUBzero controls the tool container's network operations. The middleware also monitors the start time as well as the duration of each connection. The connections that are not being used for a considerable period after starting is terminated and marked as inactive[3].

## 3. FEATURES

HUBzero unique blend of social networking and simulation power made it popular among research and scientific communities. HUBzero's platforms are sophisticated because of the various features it offers. HUBzero sites provides a lot of features for collaboration, development and deployment of tools, interactive user interfaces for simulation tools, online community groups, Wikis, Blogs, Provision for Feedback Mechanisms, Metrics for Usage, content management resources and everything that would be needed for a collaborative platform

### 3.1. Development and Deployment of Tools

Each hub comes with a companion site for developing source code based on open source package called Trac for project management. Uploading a tool is a little complicated. Tools must be uploaded, compiled, tested, fixed, compiled again, and tested

again many times before being published. Each tool will have its own project area within the site, with a repository for source code control, a ticketing system for bug tracking as well as a wiki area for project documentation.

A team member from development get access to workspace which is a Linux desktop running in a secured execution environment, accessed via a Web browser. The tools that are available on the hub dont come from the core development team rather they are developed by various collaborators across the world. Developers can use HUBzero's rapture toolkit to create a GUI with little effort[4].

### 3.2. Online Presentations

HUBzero facilitates Online Presentations along with the tools. The PowerPoint slides are combined with voice and animation. The online presentations provide user a standard seminar experience. HUBzero uses MacroMedia Breeze to deliver the online presentations in a compact format using Flash, which is present on 98 percent of the world desktops. The online presentations provided by HUBzero can also be distributed as podcasts which be accessed by the users on-the-go via their personal device assistants[4].

### 3.3. Interactive Simulation Tools

The underlying premise of HUBzero is its signature service to deliver live graphical simulation tools that are interactive and can be accessed using a normal web browser. The tools in a hub are interactive; you can zoom a molecule, analyze 3D volume interactively and need not even wait for the webpage to be refreshed. A user can visualize results in real-time and need for wait for processing. One can deploy new tools without having to write any special code for deploying on the web[4].

HUBzero's Rappture Toolkit helps to create GUI with little effort. The HUBzero infrastructure provides tool execution as well as delivery mechanism based on Virtual Network Computing(VNC). Using the architecture provided by HUBzero, the first ever developed hub, nanoHUB has brought about 50 simulation tools live in span of 2 years.

### 3.4. Collaboration and Support

HUBzero is a popular because it not only provides simulation tools, wikis etc., but also provides sophisticated mechanism for colleagues to collaborate among themselves and work together. HUBzero middleware hosts tool sessions, where a single session can be shared with among many number of people. Each session provides a textbox entry beneath it to allow the users to enter the colleagues names they wish to collaborate with. A group of people can see the same session at the same time and discuss ideas over the phone or instant messaging [4].

HUBzero also provides community forum modeled after Amazon.com's Askville where users can post questions. Users can also tag the posts with the tool names and concepts so that users can find questions matching their interests and post answers accordingly. Users also can file trouble reports if something goes wrong which are sent directly to tool development team. Developers view the help requests as tickets and investigate and resolve the issues. Users also can have access to wishlist where they can request new features for a particular tool. The developers can determine a wish's relative priority and judges which are important and requires little effort are ordered accordingly. There is points and bounty system which enhances involvement of users in the hub by providing premium access to the resources based on the levels of bounties and points.

### 3.5. Content Tagging, Wikis abd Blogs

The resources available on a hub are categorized by a series of tags. Each tag has an associated page on the hub where its resources are listed and meaning is defined. Tags can be defined by anyone like the contributor, users of the page or the hub adminstrator.

Hubs support topic pages. Each topic is created using a standard wiki syntax by a specified list of authors. Users can comment on the content of the wiki or even suggest changes. The original moderators of the topic are notified about the suggestions which they will consider if apt. The page can also be given a ownership. Topic pages act as lightweight articles that help to describe various resources on the hub[4].

## 4. APPLICATIONS

HUBzero platform has been very popular in science and engineering fields. Many hubs have come into existence in various fields ranging from cancer to nanotechnology. More than 30 hubs have been spawned since HUBzero platform had come into existence. All these hubs serve more than 8,50,000 visitors from 172 countries worldwide during year 2012 alone.

### 4.1. NanoHUB

nanoHUB is the reason behind the HUBzero platform. HUBzero was initially created to support nanoHUB.org, an online community for the Network for Computational Nanotechnology, which is funded by US National Science Foundation in 2002 to connect theorists who develop simulation tools with the experimenters and engineers.

nanoHUB is an online portal for nanotechnology where researchers, students and instructors collaborate to share scientific tools, simulation tools, educational material, research etc., It uses cyberinfrastructure to provide access to this tools and also the instructional materials. The users can run experiments, download lectures as well as review research[5].

### 4.2. Big Data Analysis in Social Science Using HUBzero

Datasets from social media are infact large and can grow beyond a individual or institution analytical capability of common software tools. Institutions such as Non-STEM or undergraduate schools lack the compute infrastructure and personnel needed to allows researchers, students, instructors to create and analyze large datasets.

The social science students are expected to be competent users. Therefore, in order to provide infrastructure necessary to expose undergraduates social science students to data intensive computing, State University of New York (SUNY) Oneonta teamed with University at Buffalo's Center for Computational Research (CCR) to establish a collaborative virtual community focusing on data intensive computing education[6].

## 5. LICENSE

HUBzero has been released as Open Source under the LGPL-3.0 license. HUBzero is community code unlike the GNU General Public License it doesn't "infect" your code. Under HUBzero's license, one can treat HUBzero as a "library," create your own unique components within our framework, and license your derivative works any way you like.

## 6. CONCLUSION

HUBzero has been a unique platform with social networking and simulation power has been able to resonate science and engineering communities. As hub keeps to grow the capabilities continues to grow, more tools and related content will be added. A hub lets various researchers, engineers collaborate to solve larger problems by connecting a series of models from independent authors.

## REFERENCES

[1] M. McLennan and R. Kennell, "Hubzero: a platform for dissemination and collaboration in computational science and engineering," *Computing in Science & Engineering*, vol. 12, no. 2, 2010.

[2] M. McLennan and G. Kline, "Hubzero paving the way for the third pillar of science," webpage, 2011. [Online]. Available: https://www.hpcwire.com/2011/02/28/hubzero_paving_the_way_for_the_third_pillar_of_science/

[3] "Information technology at purdue research compting(rcac)," webpage. [Online]. Available: https://www.rcac.purdue.edu/services/hubzero/

[4] "hubzero:platform for scientific collaborations," webpage. [Online]. Available: https://hubzero.org/tour/features

[5] G. Klimeck, M. McLennan, S. P. Brophy, G. B. Adams III, and M. S. Lundstrom, "nanohub. org: Advancing education and research in nanotechnology," *Computing in Science & Engineering*, vol. 10, no. 5, pp. 17–23, 2008.

[6] J. M. Sperhac, S. Gallo, J. B. Greenberg, B. Lowe, B. Wilkerson, G. Fulkerson, and B. Heindl, "Teaching big data analysis in the social sciences using a hubzero-based platform," Oct 2014. [Online]. Available: https://hubzero.org/resources/1235