Collaborative Coastal Modeling with SIMULOCEAN

Jian Tao

TEES & HPRC
Texas A&M University

jtao@tamu.edu

Texas A&M Research Computing Week

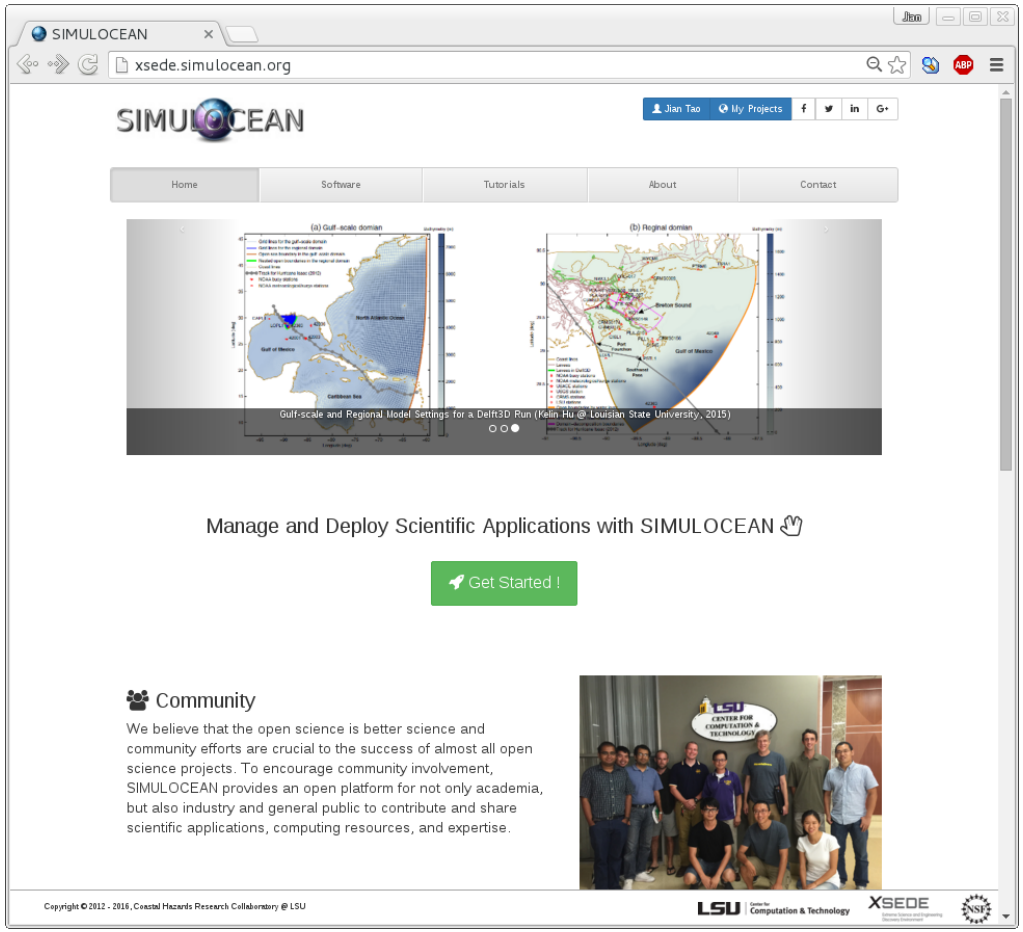
Aug 14, 2017 Urbana–Champaign

Table of Contents

Introduction

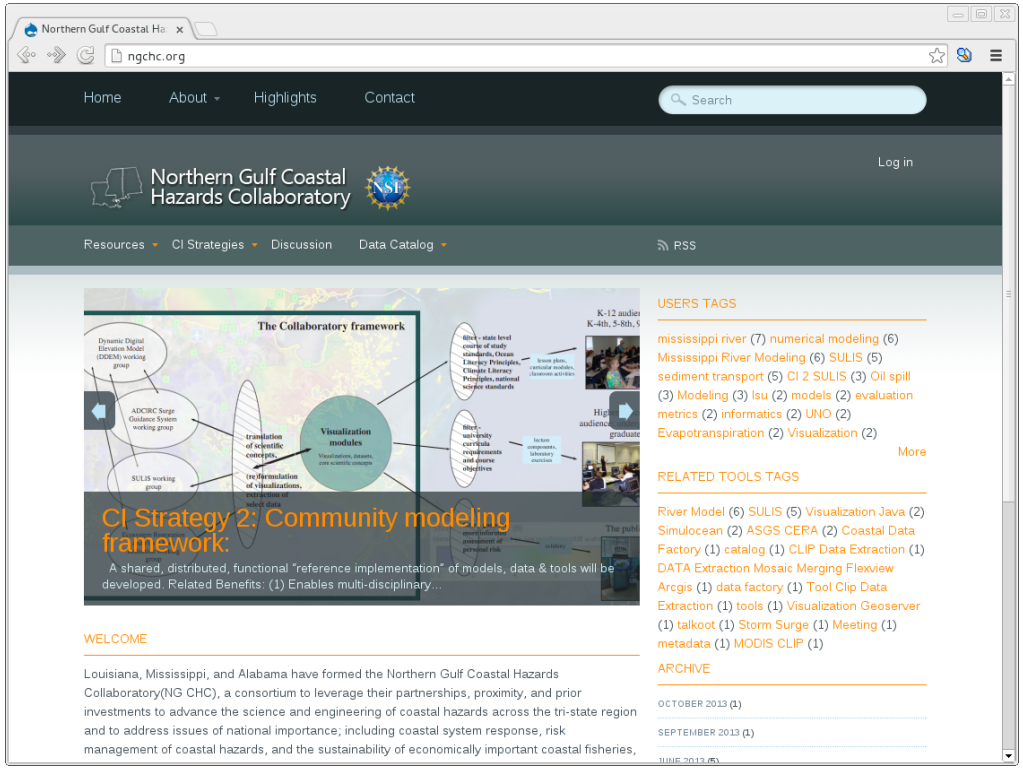
SIMULOCEAN Science Gateway

SIMULOCEAN - http://xsede.simulocean.org



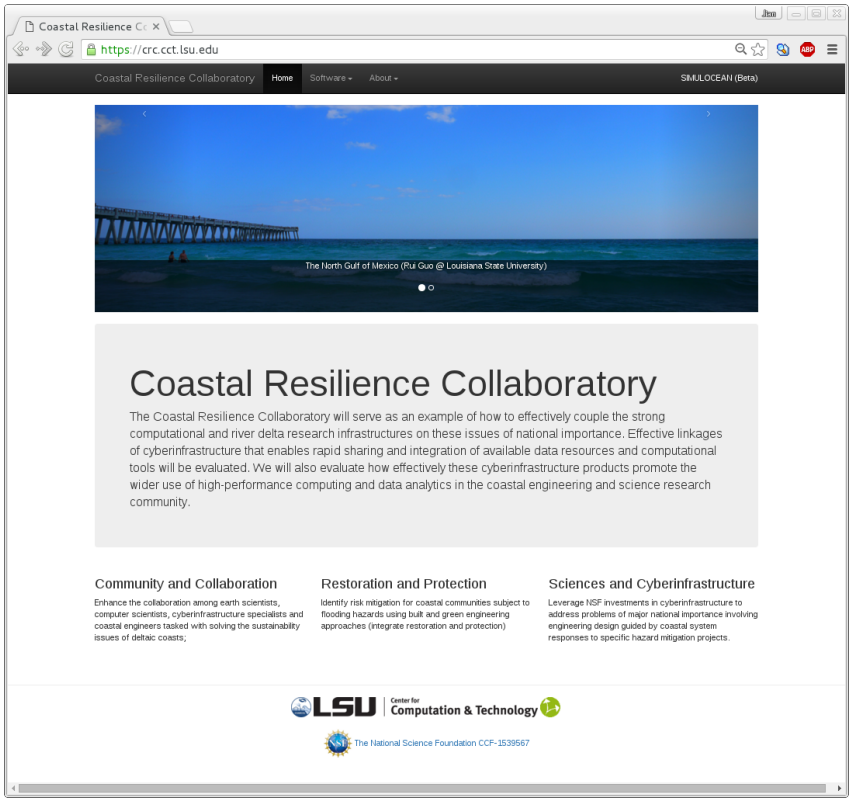
NG-CHC Project - PI: Michael Khonsari

Northern Gulf Coastal Hazards Collaboratory - http://ngchc.org
(NSF Award: EPS-1010640, $2,166,000.00,
10/1/2010-09/13/2013)



CRC Project - PI: Q. Jim Chen

Coastal Resilience Collaboratory - http://crc.cct.lsu.edu
(NSF Award: CCF-1539567, $1,199,154.00,
10/1/2015-9/30/2019)



Our Vision

We envision SIMULOCEAN as

A Computational Platform

We aim to create a computational platform for coastal modeling applications with our competitiveness and expertise on high performance computing technology and coastal applications.

and

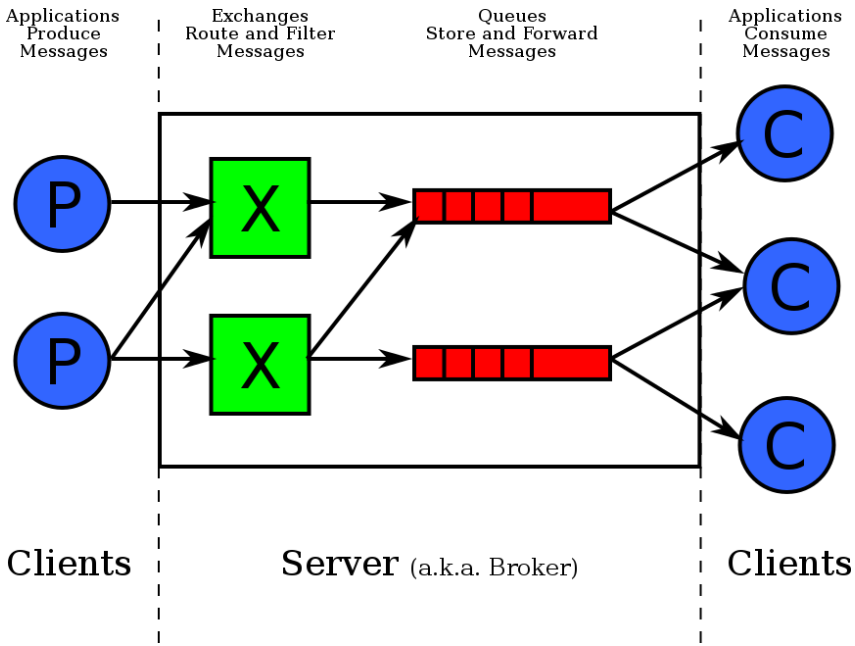
A Collborative and Educational Environment

We aim to advance research, enrich training, inspire collaboration, and inform decision making through highly available innovation-enabling cyberinfrastructure.

Web Programming Technologies

Advanced Message Queuing Protocol

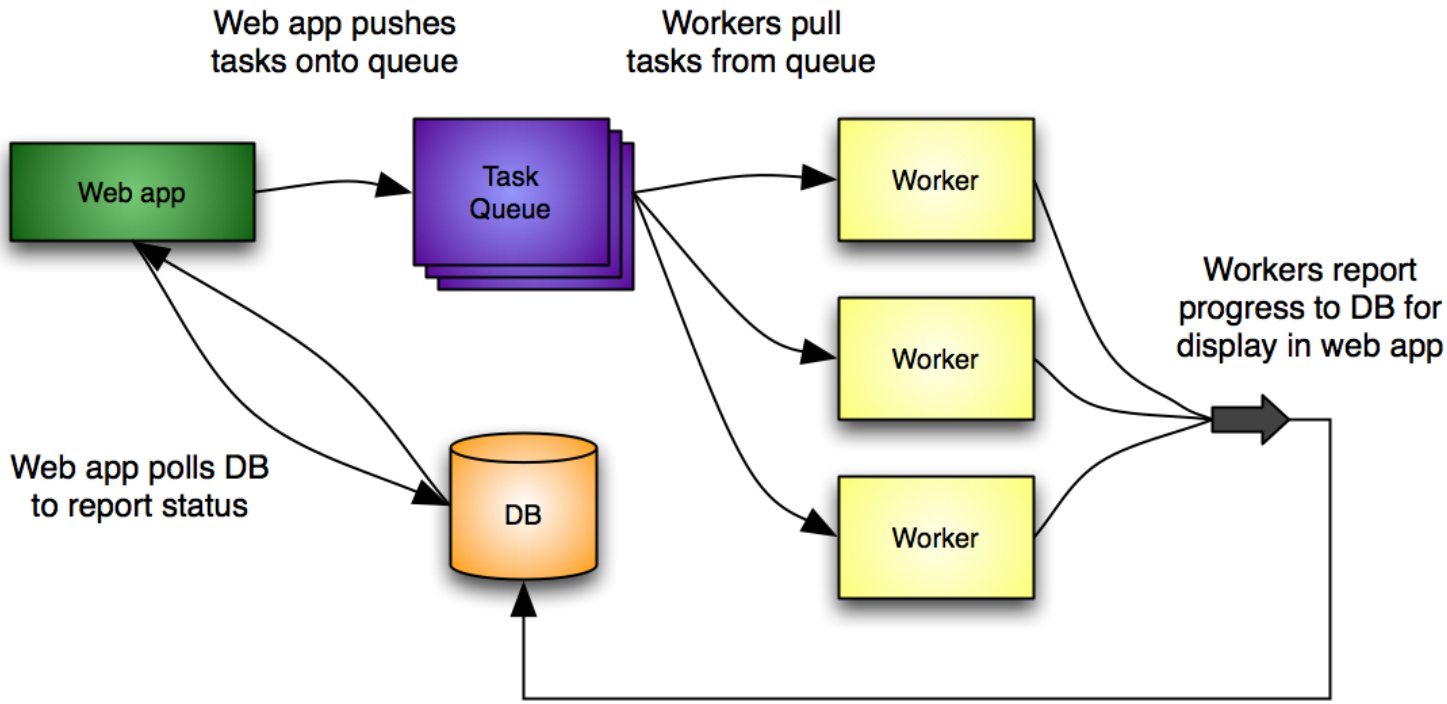
The Advanced Message Queuing Protocol (AMQP) is an open standard application layer protocol for message-oriented middleware (http://www.amqp.org/). **RabbitMQ** is one of several open source message broker software packages that implement AMQP (https://www.rabbitmq.com/).



(image credit: https://www.wikipedia.org/)

Celery - Distributed Task Queue

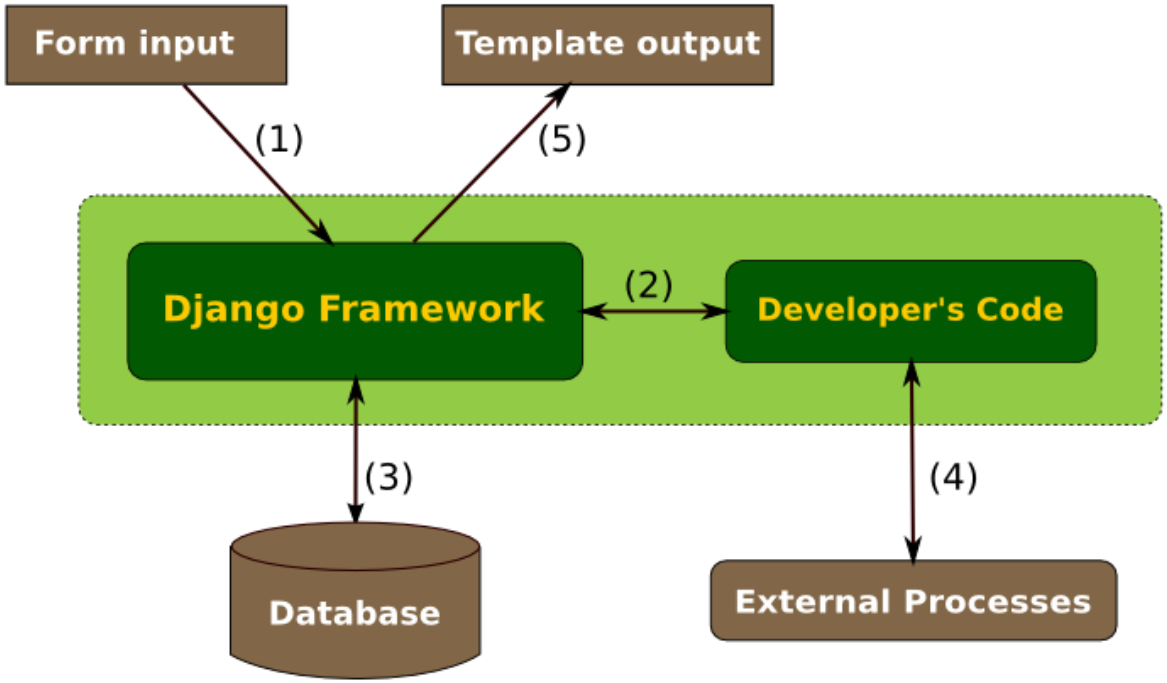
Celery is an asynchronous task/job queue based on distributed message passing (http://www.celeryproject.org/). It supports RabbitMQ and other message brokers.



(image credit: http://digitheadslabnotebook.blogspot.com/)

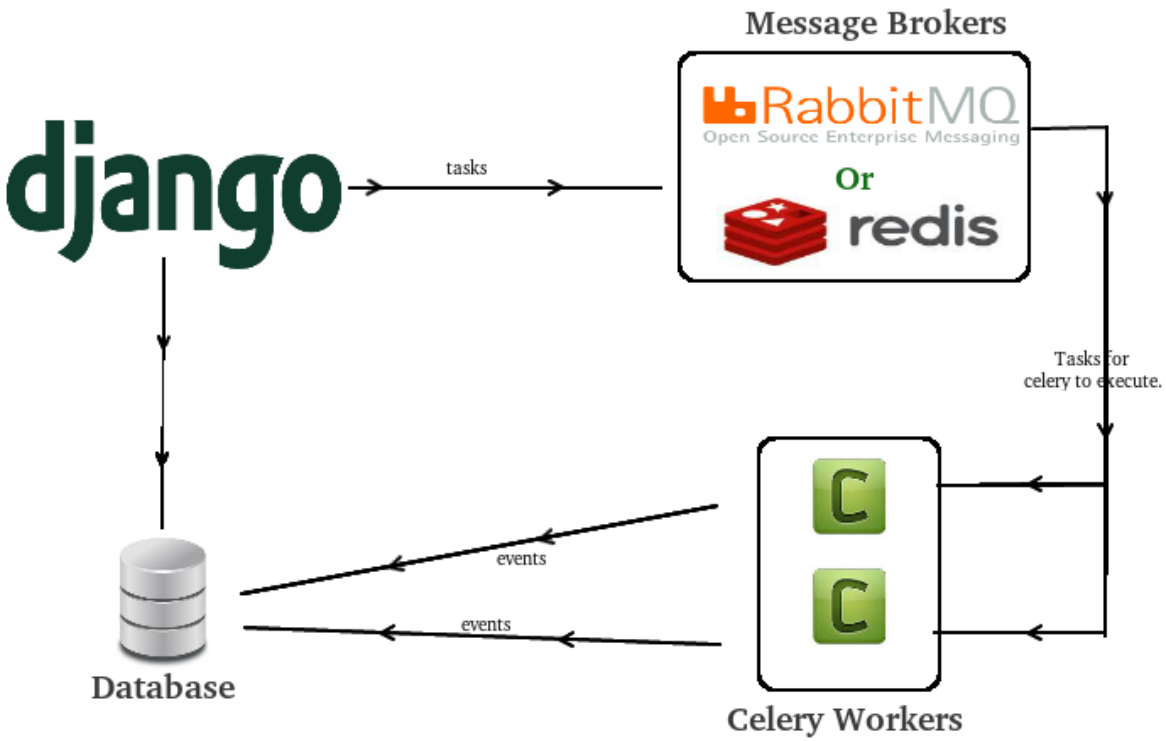
Django Web Framework

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. It supports **MariaDB** and other many other database backends.



(image credit: https://www.djangoproject.com/)

Put All the Blocks Together



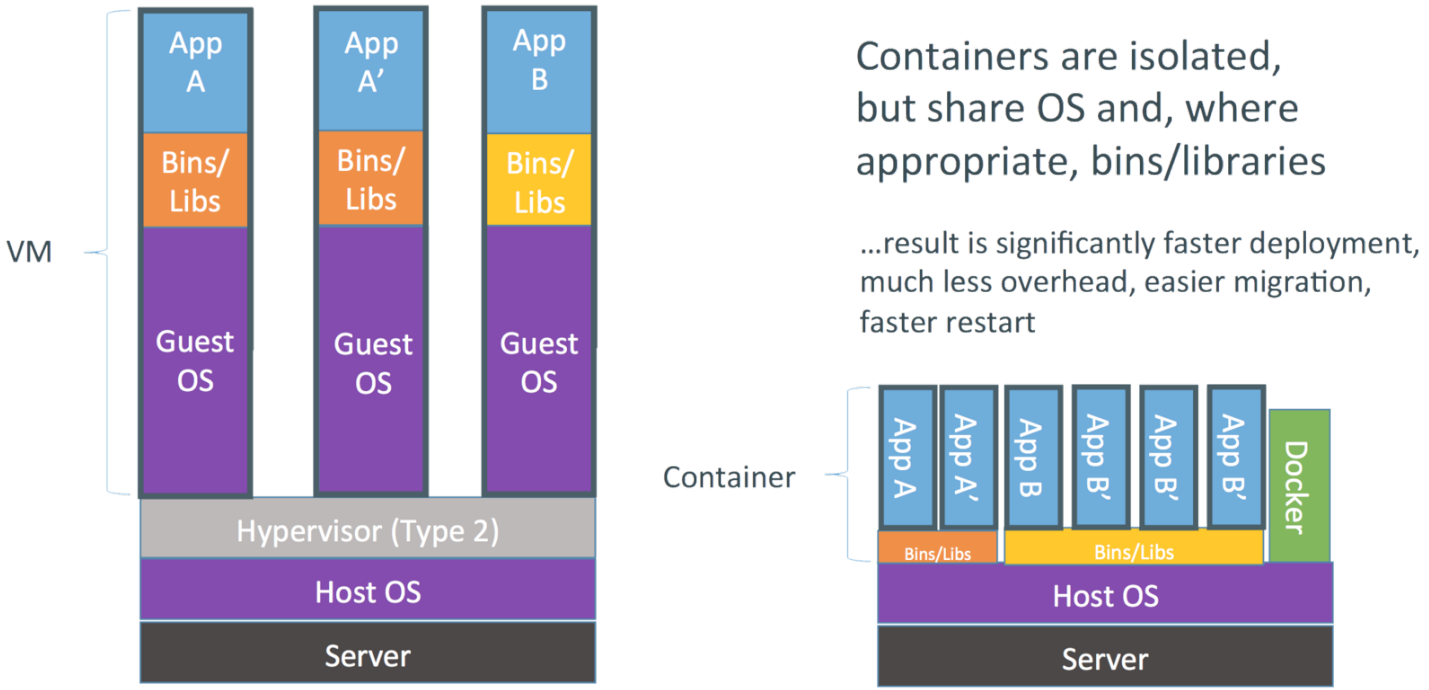
(image credit: http://my-django-python.blogspot.com/)

Containerization with Docker

Container v.s. Virtual Machine

"Containerization is a lightweight alternative to full machine virtualization that involves encapsulating an application in a container with its own operating environment."

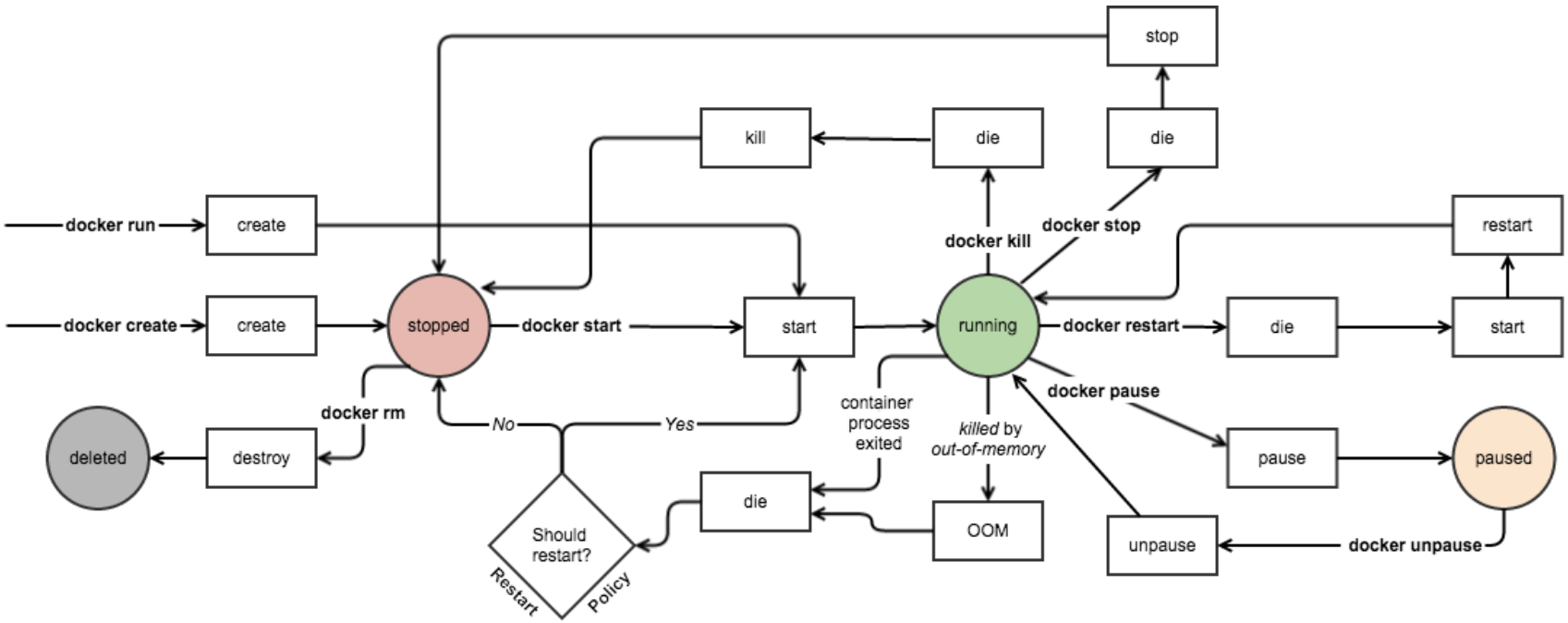
— http://www.webopedia.com/



(image credit: https://www.docker.com/)

Docker

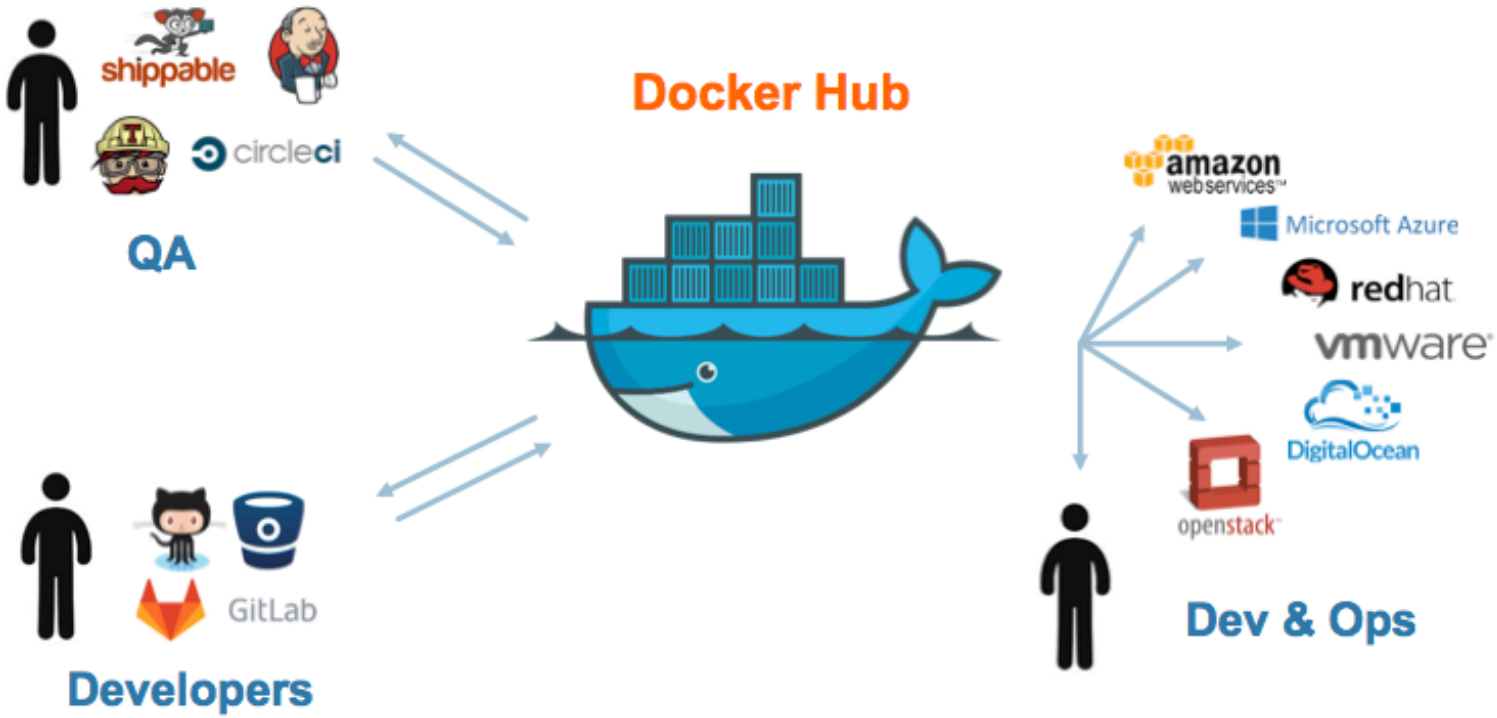
Docker - https://www.docker.com/ is an open platform for distributed applications for developers and sysadmins. It provides an additional layer of abstraction and automation of operating-system-level virtualization on Linux.



(image credit: https://www.docker.com/)

Docker Hub

The Docker Hub - https://hub.docker.com/ is a public registry maintained by Docker, Inc.

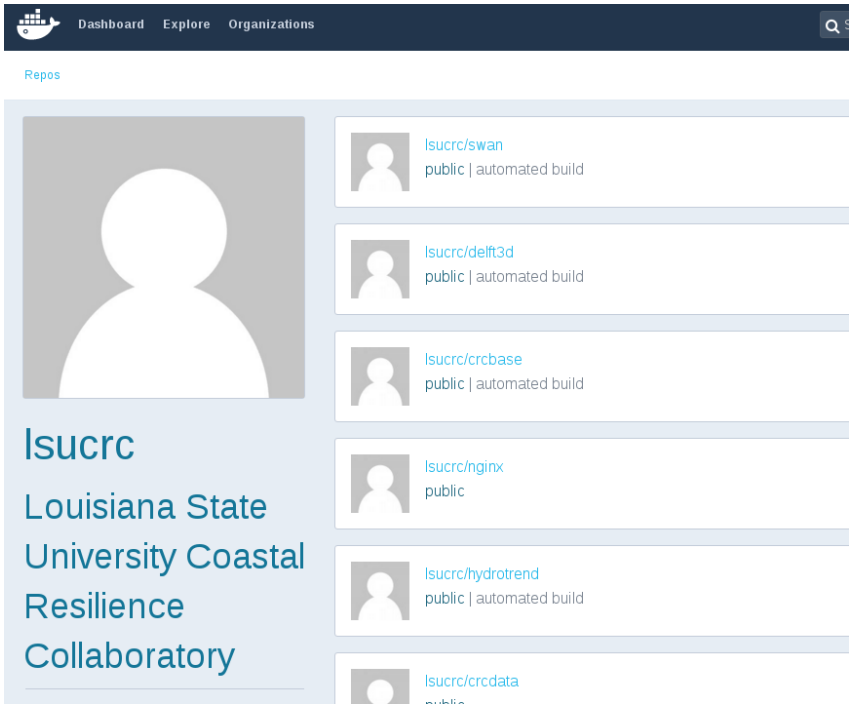


(image credit: https://www.docker.com/)

Coastal Model Repository & SIMULOCEAN

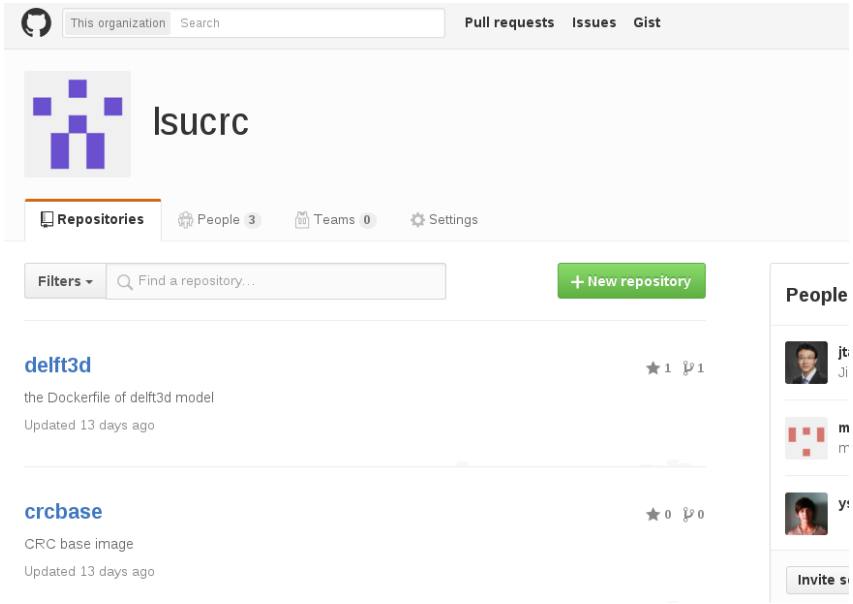
LSU CRC Docker Hub Repo

The Coastal Model Repository is currently hosted on Docker Hub as a public platform for sharing and exchanging open source models. All images hosted at Docker Hub are automatically built and tested.



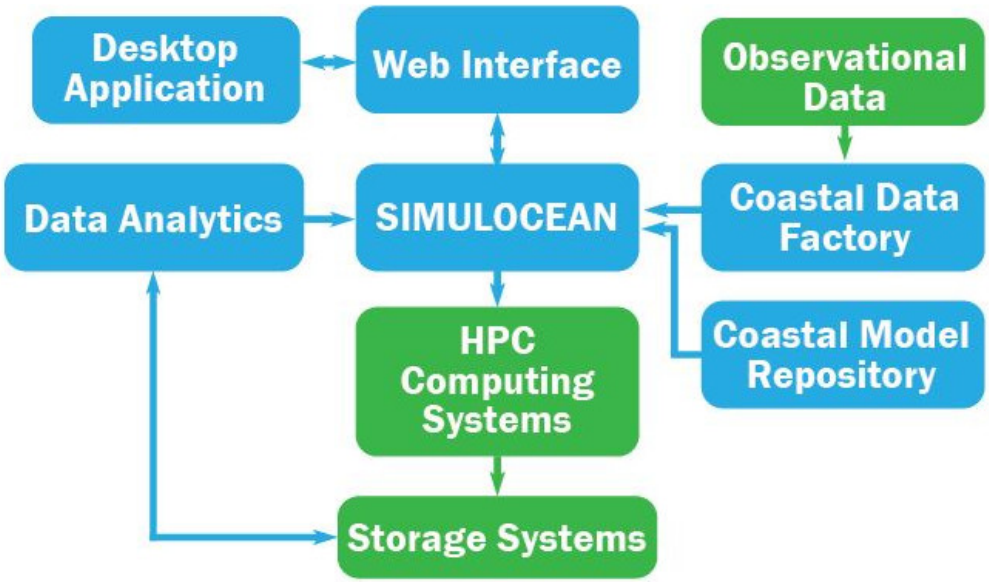
LSU CRC Github Repo

The LSU CRC Github repositories host the Docker files that are linked with the Docker images on Docker Hub.



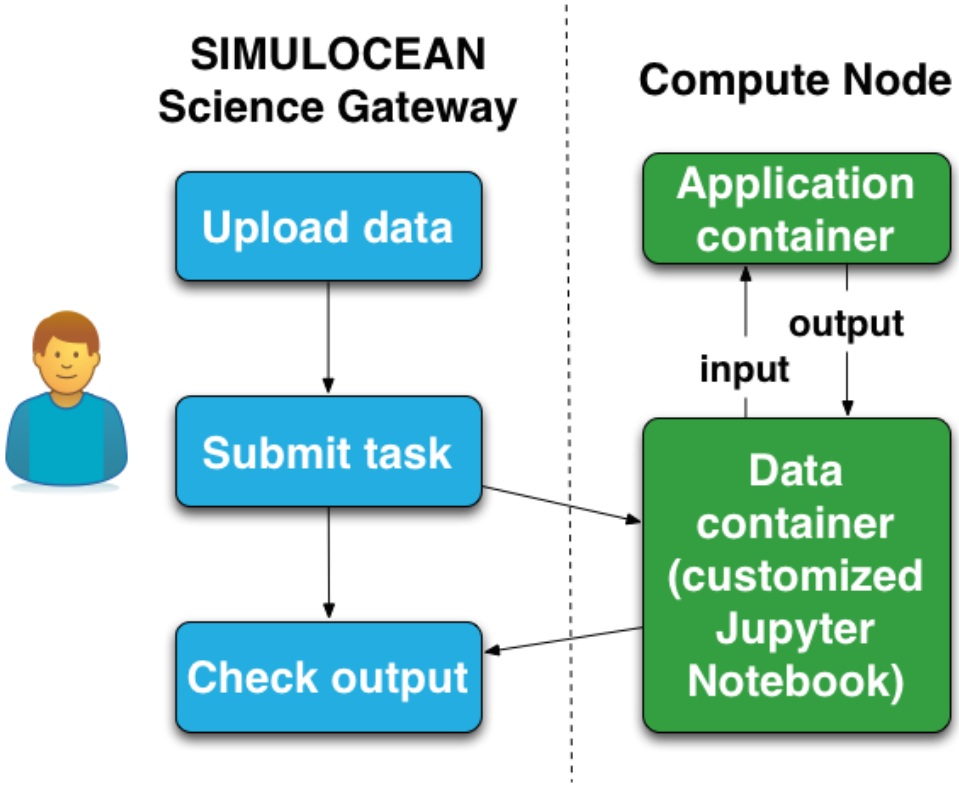
SIMULOCEAN Architecture

A platform for managing and deploying containerized coastal models and other scientific applications on cloud-ready computing systems.



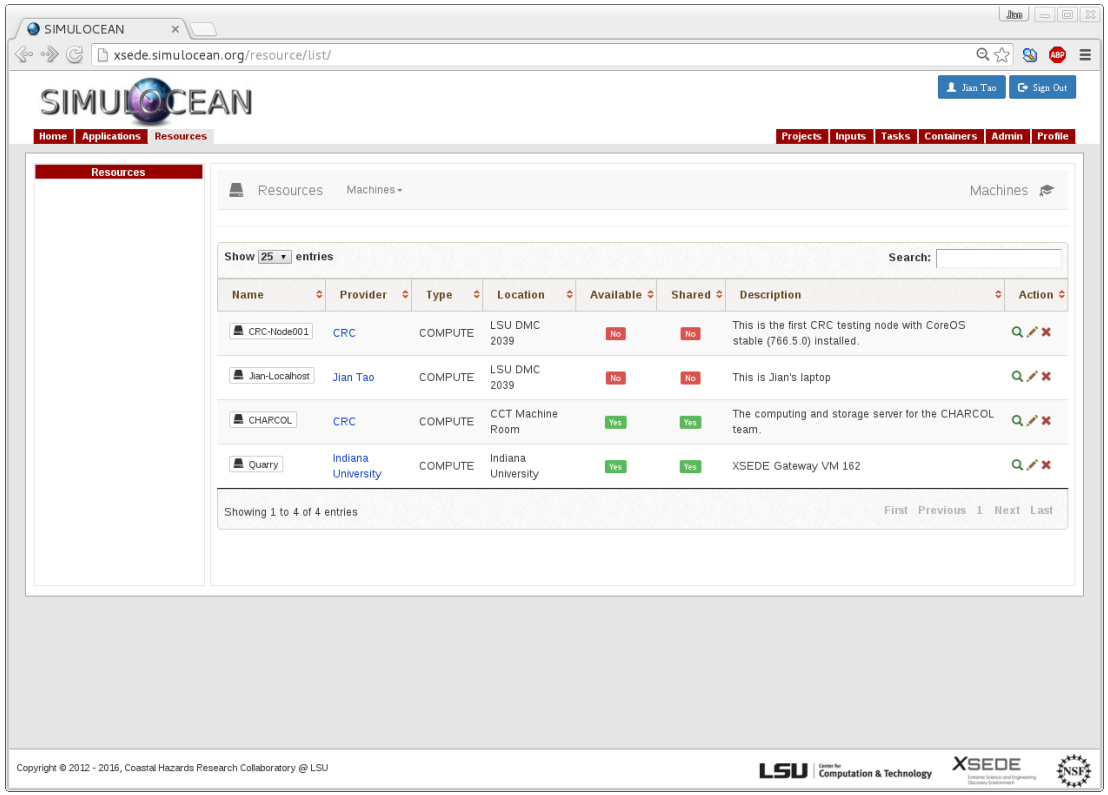
Workflow in SIMULOCEAN

A quick tutorial can be found at
http://xsede.simulocean.org/about/tutorials



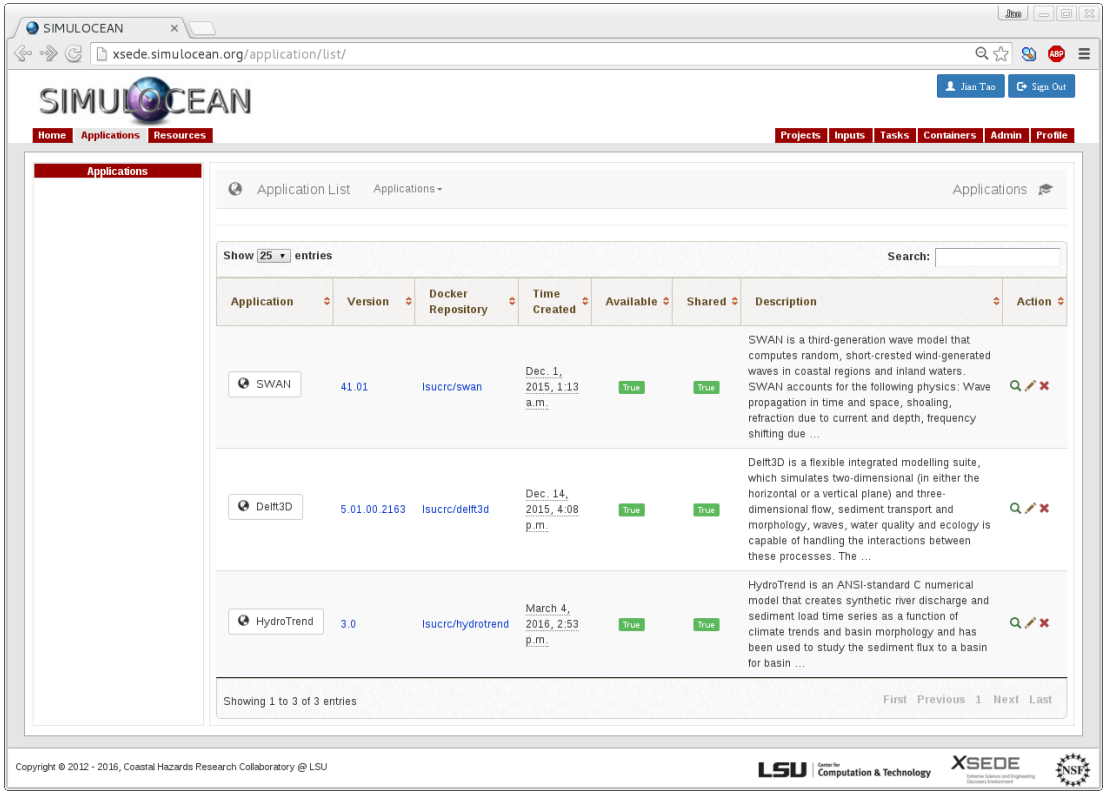
Computing Resources

SIMULOCEAN gets access to computing resources via remote SSH (an encrypted network protocol) execution.



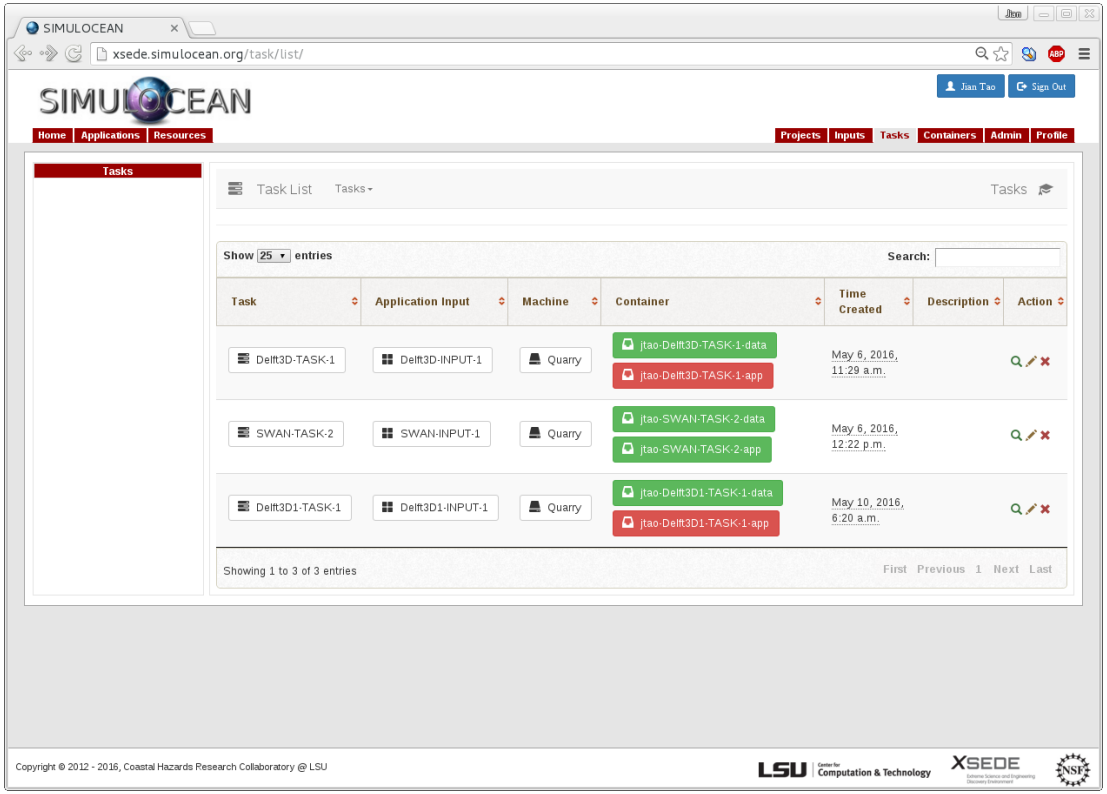
Coastal Models

SIMULOCEAN provides an interface for managing models hosted at Coastal Model Repository.



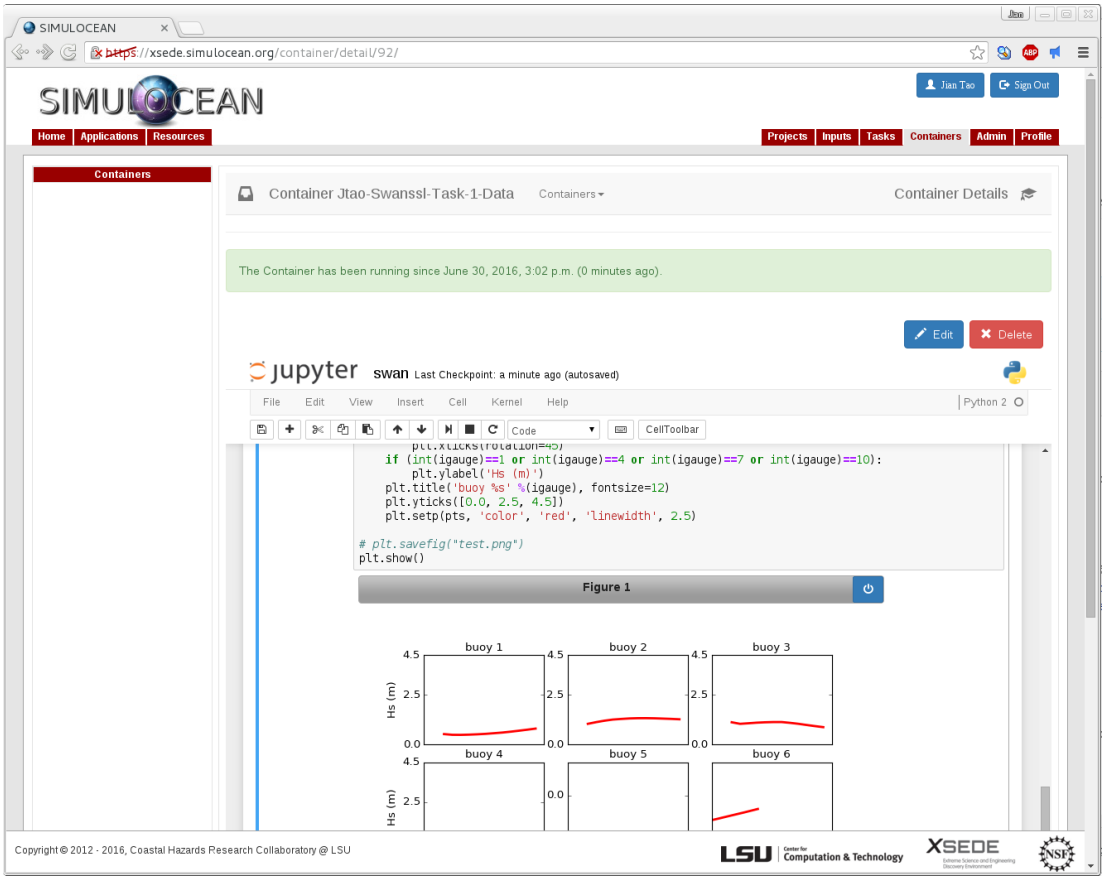
Task/Container Management

SIMULOCEAN launches and monitors Docker containers on computing systems.



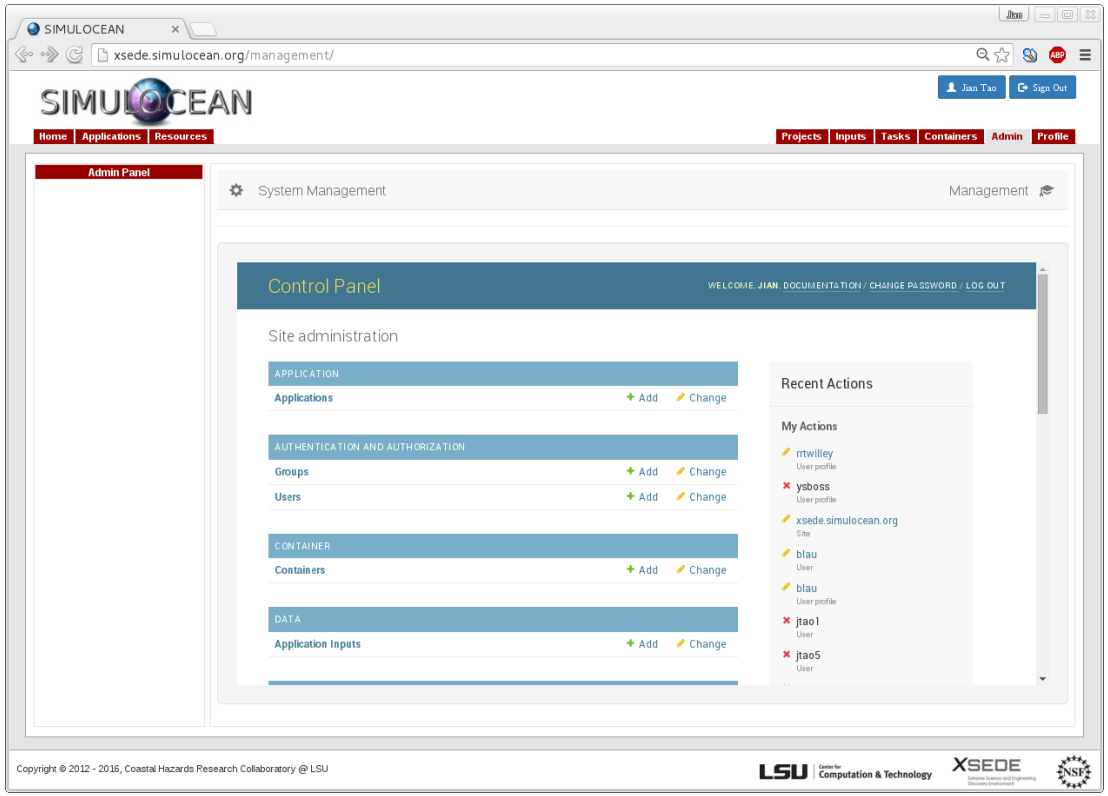
Data Container

For each task, a data container that runs a Jupyter server will be launched to serve data back to users and enable interactive data analysis and visualization.



System Administration

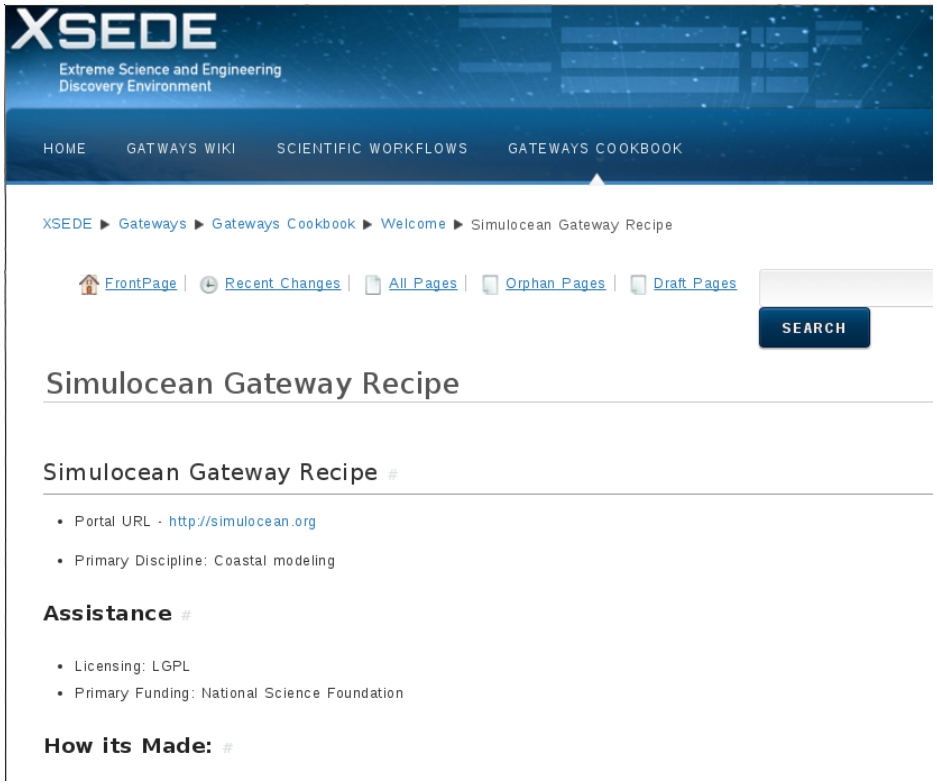
SIMULOCEAN uses Django (https://www.djangoproject.com/) to build the web-based interface and RESTful APIs for serving mobile and desktop applications.



Future Plans

Science Gateway on NSF Cloud

SIMULOCEAN was supported by XSEDE ECSS program. Past ECSS experts include: Stu Martin and Eric Blau (Globus Team), Mona Wong and Andrea Zonca (SDSC Team). We are in the process to test out singularity on Bridges and Comet.



Commercial Cloud Platforms

We will look into Amazon AWS and other commercial cloud platforms to seek opportunities to offer services to more users.

(image credit: http://blog.levvel.io)

Acknowledgments

Acknowledgments

My thanks go to

- SIMULOCEAN team members: Shuai Yuan, Du Jin, Kelin Hu, Q. Jim Chen, and Honggao Liu.
- CHARCOAL group members and our collaborators in NG-CHC and CRC projects
- NSF (Awards EPS-1010640 and CCF-1539567)
- LSU HPC, CCT, LONI, HPRC, TEES, and XSEDE for the computing resources.
- CSDMS Integration Facility and XSEDE Extended Collaborative Support Service (ECSS) program for their support and help