

# Collaborative Research with SIMULOCEAN Science Gateway

Jian Tao

TEES & HPRC  
Texas A&M University

[jtao@tamu.edu](mailto:jtao@tamu.edu)

Texas A&M Research Computing Week

June 9, 2017 College Station



TEXAS A&M ENGINEERING  
EXPERIMENT STATION



DIVISION OF RESEARCH  
TEXAS A&M UNIVERSITY  
HIGH PERFORMANCE RESEARCH COMPUTING



Center for  
Computation & Technology



# Table of Contents

- 1 Introduction
- 2 Web Programming Technologies
- 3 Containerization with Docker
- 4 Coastal Model Repository & SIMULOCEAN
- 5 Future Plans
- 6 Acknowledgments

# Introduction

# SIMULOCEAN Science Gateway

SIMULOCEAN - <http://xsede.simulocean.org>

SIMULOCEAN

Home Software Tutorials About Contact

US Gulf coast states

US Inland states

Manage and Deploy Scientific Applications with SIMULOCEAN

Get Started!

**Community**

We believe that the open science is better science and community efforts are crucial to the success of almost all open science projects. To encourage community involvement, SIMULOCEAN provides an open platform for not only academia, but also industry and general public to contribute and share scientific applications, computing resources, and expertise

Copyright © 2011 - 2016, Grand Hyacks Research Collaboratory @ LSU

LSU Computer & Technology XSEDE

# 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)

Home About Highlights Contact

Northern Gulf Coastal Hazards Collaboratory

Log in

Resources Strategies Discussion Data Catalog RSS

### The Collaboratory framework

**CI Strategy 2: Community modeling framework**

A shared, distributed, functional "reference implementation" of models, data & tools will be developed. Related Benefits: (1) Enables multi-disciplinary

#### USERS TAGS

- mississippi river (7) numerical modeling (6)
- Mississippi River Modeling (6) SILLUS (5)
- sediment transport (5) CI 2 SILLUS (2) CI 1a (1)
- Modeling (7) Iu (2) models (2) evaluation metrics (2) informatics (2) UNO (2)
- Evapotranspiration (2) Visualization (2)

More

#### RELATED TOOLS TAGS

- River Model (6) SILLUS (6) Visualization Java (2)
- SimulOcean (2) ASOS CERA (2) Coastal Data Factory (1) catalog (1) CLIP Data Extraction (1)
- DATA Extraction Mosaic Merging Preview
- Angle (1) data factory (1) Tool Clip Data Extraction (1) tools (1) Visualization Converter (1) talkoot (1) Storm Surge (1) Meeting (1) metadata (1) MOOS CLIP (1)

#### ARCHIVE

- oct10ee 2012 (g)
- sep10ee 2012 (g)
- aug 2012 (g)

WELCOME

Louisiana, Mississippi, and Alabama have formed the Northern Gulf Coastal Hazards Collaboratory(NG CHC), a consortium to leverage their partnerships, proximity, and prior investments to advance the science and engineering of coastal hazards across the tri-state region and to address issues of national importance, including coastal system response, risk management of coastal hazards, and the sustainability of economically important coastal fisheries.

# 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)

Coastal Resilience Collaboratory

Home Services About

Coastal Resilience Collaboratory

The Storm Out of Season (Pia Car) @ Louisiana State University

## Coastal Resilience Collaboratory

The Coastal Resilience Collaboratory will serve as an example of how to effectively couple the strong computational and/or data research infrastructures on these issues of national importance. Effective linkages of cyberinfrastructure that enables rapid sharing and integration of available data resources and computational tools will be evaluated. We will also evaluate how effectively these cyberinfrastructure products promote the wider use of high-performance computing and data analytics in the coastal engineering and science research community.

**Community and Collaboration**  
 Enhance the collaboration among earth scientists, engineers, scientists, cyberinfrastructure specialists and students to learn with solving the sustainability issues of public issues.

**Restoration and Protection**  
 Identify the mitigation for coastal communities subject to flooding from the coastal, full and green engineering approaches (program restoration and protection).

**Sciences and Cyberinfrastructure**  
 Leverage NSF investments in cyberinfrastructure to address problems of major national importance involving engineering design (predicts coastal system responses to specific hazard mitigation projects).

LSU Center for Computation & Technology

The Moore Science Foundation (CFP-020507)

# 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

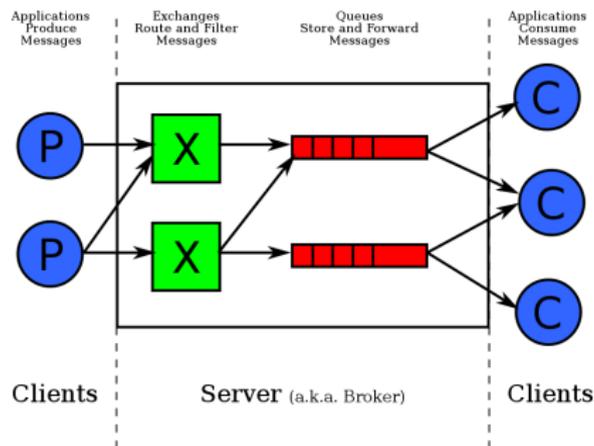
## A Collaborative and Educational Environment

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

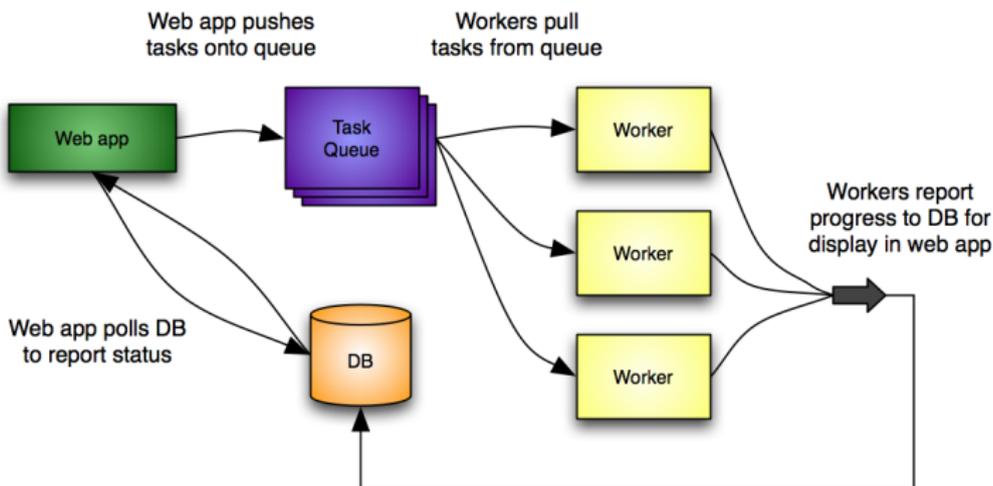
The Advanced Message Queuing Protocol (AMQP) is an open standard application layer protocol for message-oriented middleware (<http://www.amqp.org/>). **RabbitMQ** is one of several open source message broker software packages that implement AMQP (<https://www.rabbitmq.com/>).



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

# Celery - Distributed Task Queue

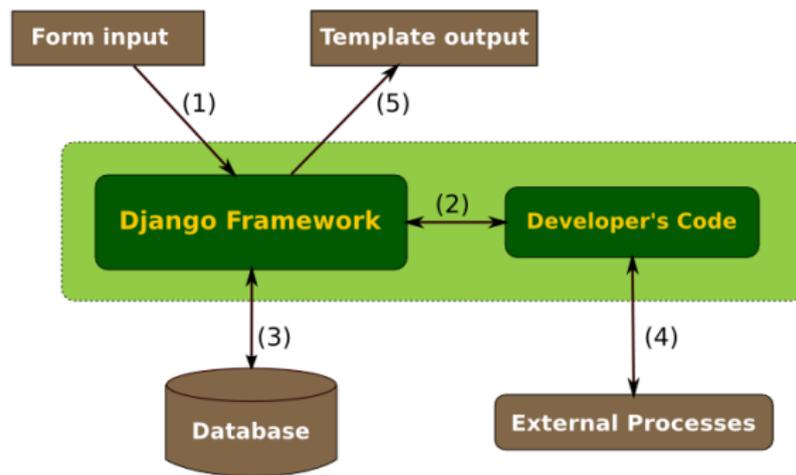
Celery is an asynchronous task/job queue based on distributed message passing (<http://www.celeryproject.org/>). It supports RabbitMQ and other message brokers.



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

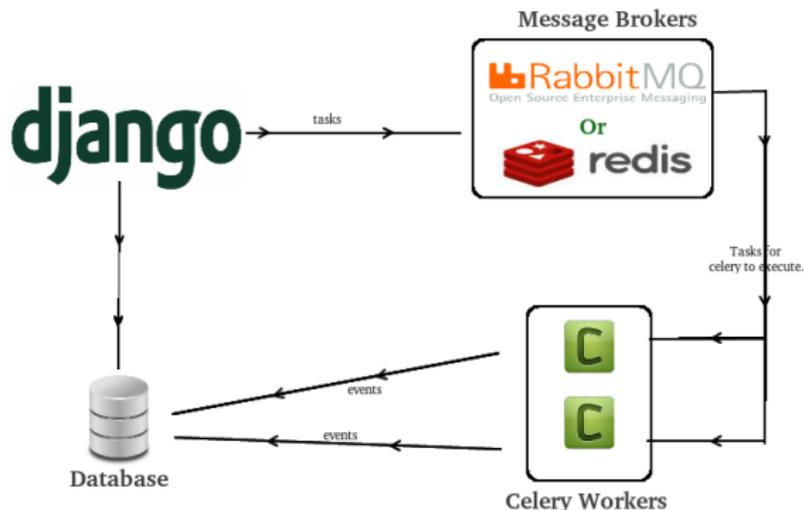
# Django Web Framework

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



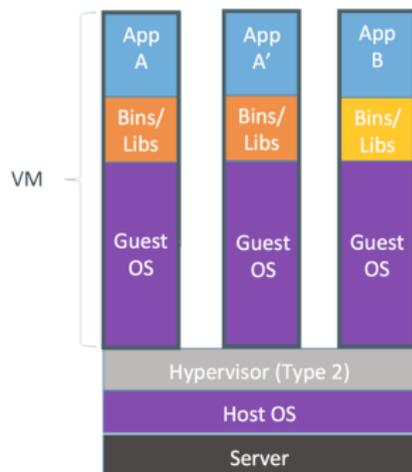
(image credit: <http://my-django-python.blogspot.com/>)

# Containerization with Docker

# Container v.s. Virtual Machine

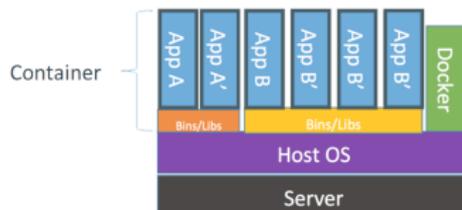
“Containerization is a lightweight alternative to full machine virtualization that involves encapsulating an application in a container with its own operating environment.”

— <http://www.webopedia.com/>



Containers are isolated, but share OS and, where appropriate, bins/libraries

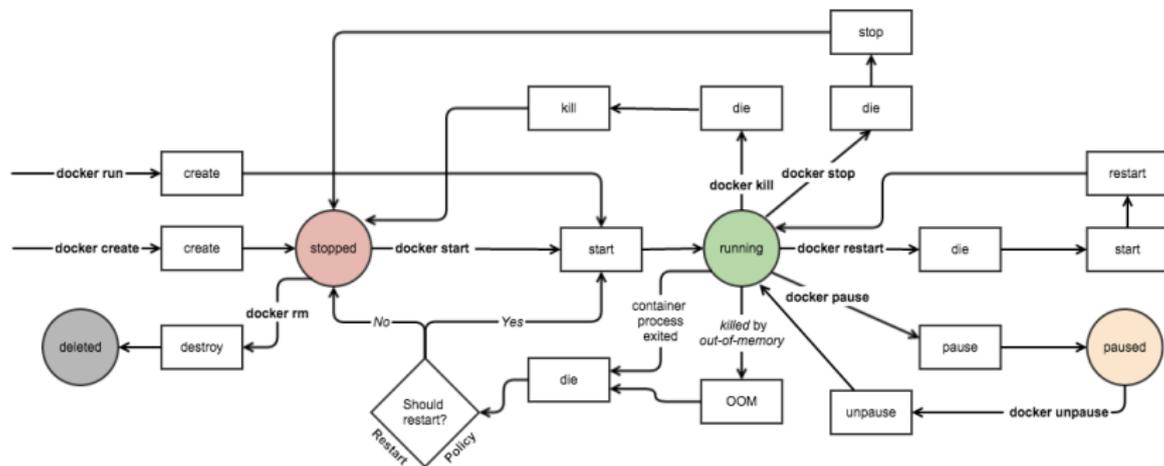
...result is significantly faster deployment, much less overhead, easier migration, faster restart



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

# Docker

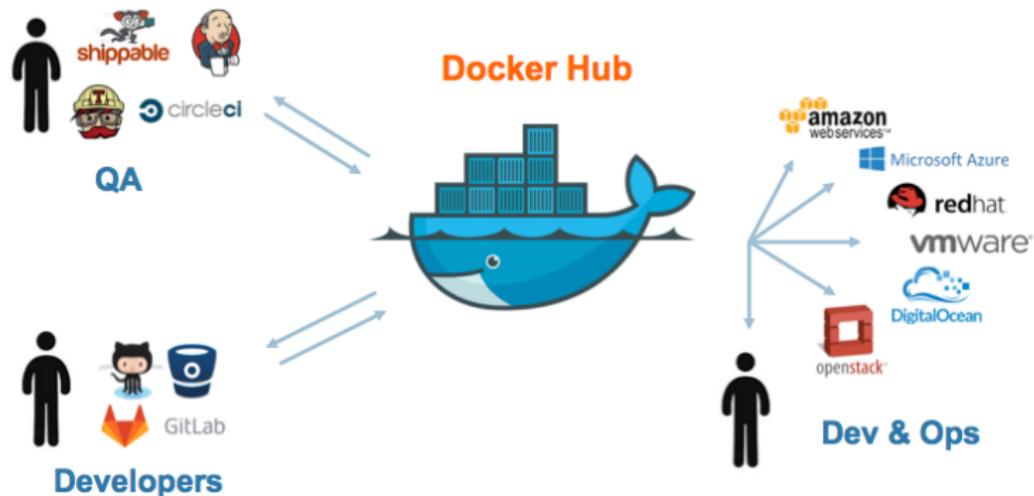
Docker - <https://www.docker.com/> is an open platform for distributed applications for developers and sysadmins. It provides an additional layer of abstraction and automation of operating-system-level virtualization on Linux.



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

# Docker Hub

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

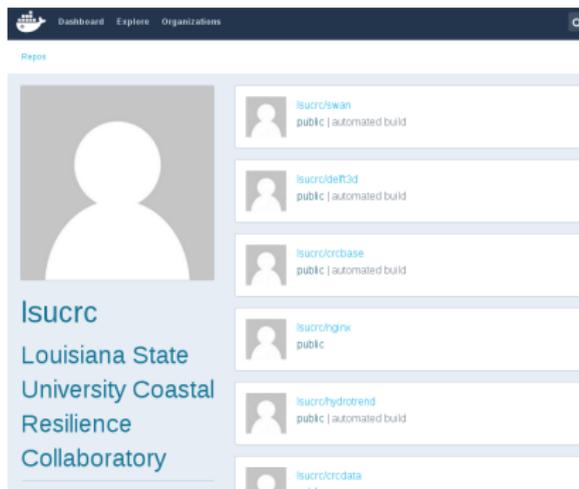


(image credit: <https://www.docker.com/>)

# 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.



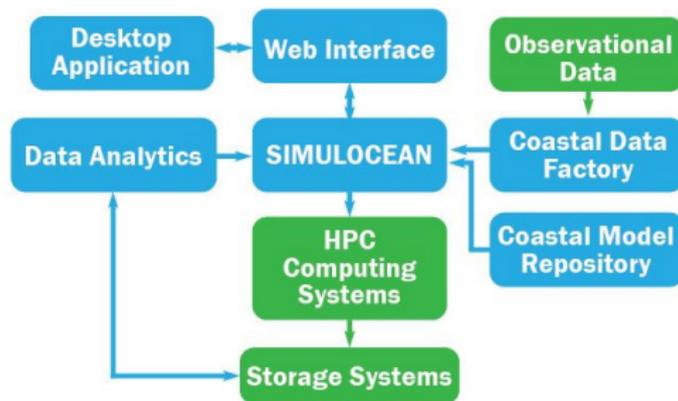
The screenshot displays the GitHub organization page for 'lsucrc'. At the top, there is a search bar and navigation links for 'Pull requests', 'Issues', and 'Gist'. The organization's profile picture and name 'lsucrc' are visible. Below this, there are tabs for 'Repositories', 'People', 'Teams', and 'Settings'. A search bar for repositories is present, along with a '+ New repository' button. The main content area lists two repositories:

- delft3d**: the Dockerfile of delft3d model, Updated 13 days ago, with 1 star and 1 pull request.
- crcbase**: CRC base image, Updated 13 days ago, with 0 stars and 0 pull requests.

On the right side, there is a 'People' section showing profile pictures and names of users associated with the organization, along with an 'invite' button.

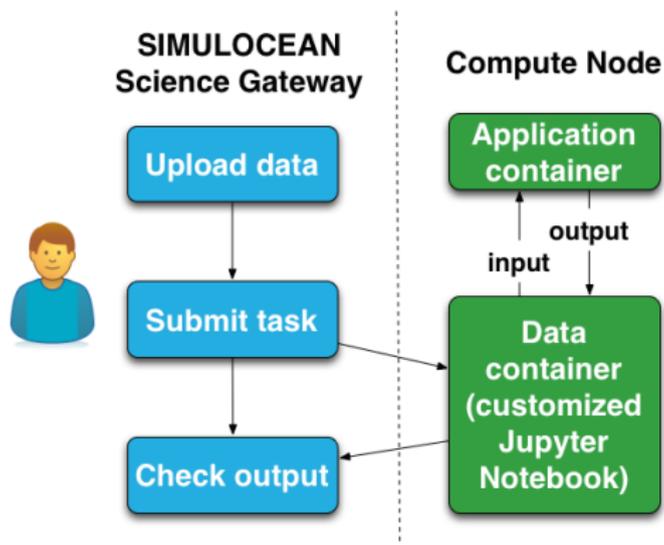
# SIMULOCEAN Architecture

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



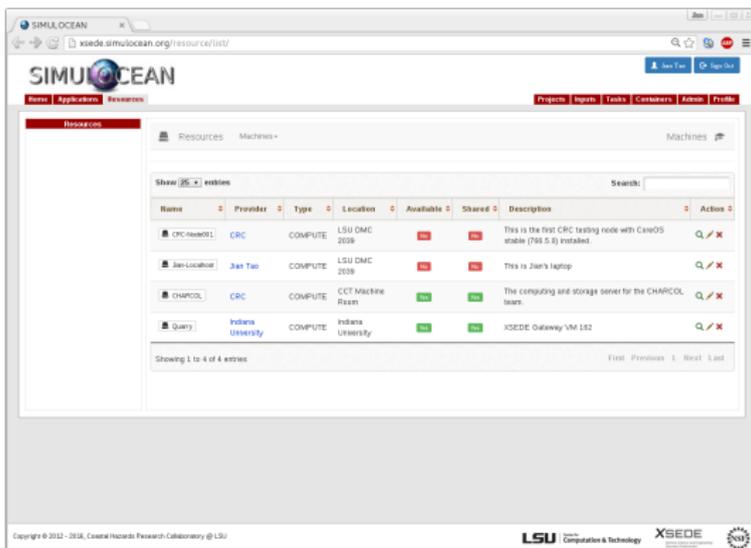
# Workflow in SIMULOCEAN

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.



The screenshot displays the SIMULOCEAN web application interface. The page title is "SIMULOCEAN" and the URL is "xseede.simulocean.org/resource/list". The interface includes a navigation menu with "Home", "Applications", "Resources", "Projects", "Logins", "Tasks", "Comments", "Admin", and "Profile". The main content area is titled "Resources" and "Machines" and shows a list of 4 resources. The table below summarizes the data shown in the screenshot:

Name	Provider	Type	Location	Available	Shared	Description	Active
CRC-aws01	CRC	COMPUTE	LSU DMC 2039	Yes	Yes	This is the first CRC testing node with CentOS stable (766.5.8) installed.	Yes
See-Location	Jian Tao	COMPUTE	LSU DMC 2039	Yes	Yes	This is Jian's laptop	Yes
CHARCOL	CRC	COMPUTE	CCT Machine Room	Yes	Yes	The computing and storage server for the CHARCOL team.	Yes
Quary	Indiana University	COMPUTE	Indiana University	Yes	Yes	XSEDE Gateway VM 182	Yes

Showing 1 to 4 of 4 entries

Copyright © 2012 - 2016, Coastal Hazards Research Collaboratory @ LSU

LSU | Center for Computational & Technology | XSEDE | 2017

# Coastal Models

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

The screenshot displays the SIMULOCEAN web application interface. The page title is "SIMULOCEAN" and the URL is "vseeds.simulocean.org/application/stz". The interface includes a navigation menu with "Home", "Applications", "Resources", "Projects", "Pages", "Tools", "Comments", "Admin", and "Profile". The main content area is titled "Applications" and shows a table of model applications.

Application	Version	Coastal Model Repository	Time Created	Available	Shared	Description	Action
SWAN	41.01	vscocean	Dec. 1, 2015, 1:13 P.M.	Yes	Yes	SWAN is a third generation wave model that computes random, short-crested wind generated waves in coastal regions and inland waters. SWAN accounts for the following physics: Wave propagation in time and space, refraction, refraction due to current and depth, frequency shifting due to ...	Q X
DeR3D	5.01 (0.2183)	vscocean@lsu	Dec. 14, 2015, 4:39 P.M.	Yes	Yes	DeR3D is a flexible stratogrid modeling suite, which simulates two-dimensional (in either the horizontal or a vertical plane) and three-dimensional flow, sediment transport and morphology, waves, water quality and ecology in capable of handling the interactions between these processes. The ...	Q X
HydroTrend	2.0	vscohydrotrend	March 4, 2016, 2:53 P.M.	Yes	Yes	HydroTrend is an ANSI standard C numerical model that creates synthetic river discharge and sediment load time series as a function of climate trends and basin morphology and has been used to study the sediment flux to a basin for basin ...	Q X

Showing 1 to 3 of 3 entries

Copyright © 2012 - 2016, Coastal Hazards Research Collaboratory @ LSU

LSU Louisiana State University  
Computation & Technology

XSEDE  
Extreme Scale Data Environment

NSF  
National Science Foundation

# Task/Container Management

SIMULOCEAN launches and monitors Docker containers on computing systems.

The screenshot displays the SIMULOCEAN web application interface. The main content area shows a 'Task List' with a table of tasks. The table has columns for Task, Application Input, Machine, Container, Time Created, Description, and Action. There are three task entries visible, each with a green 'Running' status and a red 'Stopped' status.

Task	Application Input	Machine	Container	Time Created	Description	Action
DeR3D-TASK-1	DeR3D-INPUT-1	Quary	<span>ipod-DeR3D-TASK-1-6493</span> <span>ipod-DeR3D-TASK-1-499</span>	May 6, 2016, 11:29 a.m.		🔍 ✕
SWAN-TASK-2	SWAN-INPUT-3	Quary	<span>ipod-SWAN-TASK-2-6023</span> <span>ipod-SWAN-TASK-2-499</span>	May 6, 2016, 12:22 p.m.		🔍 ✕
DeR3D-TASK-1	DeR3D-INPUT-1	Quary	<span>ipod-DeR3D-TASK-1-6493</span> <span>ipod-DeR3D-TASK-1-499</span>	May 10, 2016, 8:20 a.m.		🔍 ✕

Showing 1 to 3 of 3 entries

Copyright © 2012 - 2016, Coastal Hazards Research Collaboratory @ LSU

LSU Louisiana State University  
 XSEDE Extreme Scale Computing & Technology

# 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.

The screenshot displays the SIMULOCEAN web interface. At the top, there is a navigation bar with links for Home, Applications, Resources, Projects, News, Tutorials, Containers, Admin, and Profile. The main content area shows a container named 'Container Jian-Tao-Swainst-Task-1-Data' with a status message: 'The Container has been running since June 30, 2016, 2:02 p.m. (0 minutes ago)'. Below this, a Jupyter notebook is open, showing a code editor with the following Python code:

```

import numpy as np
import matplotlib.pyplot as plt

if int(ligauge1) or int(ligauge2) or int(ligauge3) or int(ligauge4) or int(ligauge5) or int(ligauge6) or int(ligauge7) or int(ligauge8) or int(ligauge9) or int(ligauge10):
    plt.ylabel("Hs (m)")
    plt.title("Busy has %d ligauge", fontSize=12)
    plt.xticks([0.0, 2.5, 4.5])
    plt.setp(plt.gca().get_lines(), color='red', linewidth=2.5)

# plt.savefig("test.png")
plt.show()

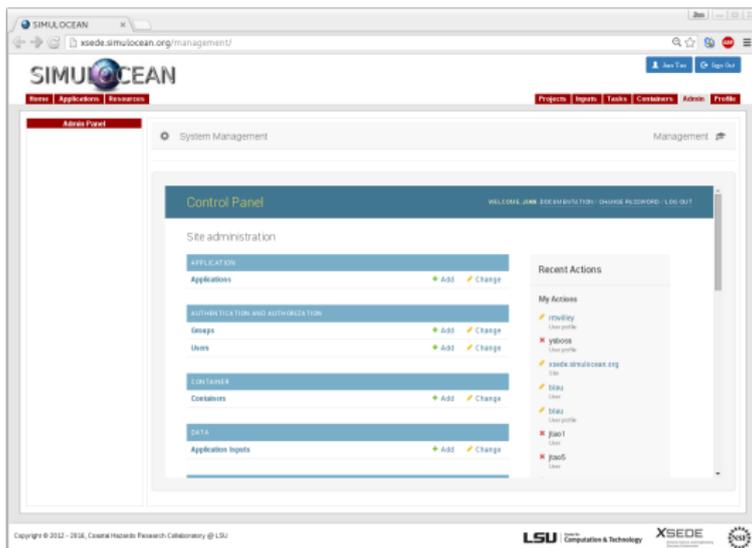
```

Below the code editor, a figure titled 'Figure 1' is displayed, showing six subplots arranged in a 2x3 grid. Each subplot is titled 'busy 1' through 'busy 6'. The y-axis for each subplot is labeled 'Hs (m)' and ranges from 0.0 to 4.5. The subplots show red lines representing wave heights. The first three subplots (busy 1, 2, 3) show relatively flat lines, while the last three (busy 4, 5, 6) show more dynamic behavior, with busy 6 showing a sharp increase.

At the bottom of the interface, there is a footer with copyright information: 'Copyright © 2012 - 2016, Coastal Hazards Research Collaboratory @ LSU'. Logos for LSU (Louisiana State University) and XSEDE (Extreme Scale Data Environment) are also present.

# System Administration

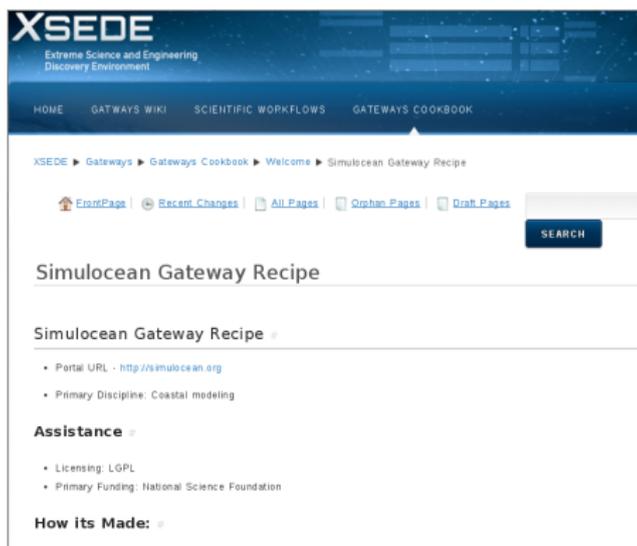
SIMULOCEAN uses Django (<https://www.djangoproject.com/>) to build the web-based interface and RESTful APIs for serving mobile and desktop applications.



# Future Plans

# Science Gateway on NSF Cloud

SIMULOCEAN has also been supported by XSEDE ECSS program. Past and current ECSS experts include: Stu Martin and Eric Blau (Globus Team), Mona Wong and Andrea Zonca(SDSC Team).



The screenshot shows the XSEDE website interface. The header includes the XSEDE logo and navigation links: HOME, GATEWAYS WIKI, SCIENTIFIC WORKFLOWS, and GATEWAYS COOKBOOK. The breadcrumb trail is: XSEDE > Gateways > Gateways Cookbook > Welcome > Simulocean Gateway Recipe. Below the breadcrumb is a utility bar with links for FrontPage, Recent Changes, All Pages, Orphan Pages, and Draft Pages, along with a SEARCH button. The main content area is titled "Simulocean Gateway Recipe" and contains the following information:

- Portal URL - <http://simulocean.org>
- Primary Discipline: Coastal modeling

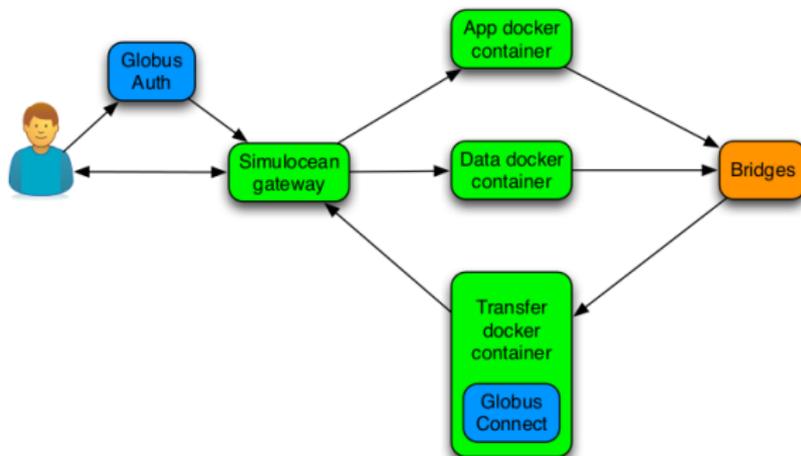
**Assistance**

- Licensing: LGPL
- Primary Funding: National Science Foundation

**How its Made:**

# Security and Authentication

We are investigating various authentication procedures for SIMULOCEAN on XSEDE and other computing platforms.



(image credit: Mona Wong from SDSC)

# Commercial Cloud Platforms

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



# Acknowledgments

# Acknowledgments

My thanks go to

- SIMULOCEAN team members: Shuai Yuan, 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