

Orchestrating Containerized Scientific Applications with SIMULOCEAN

Jian Tao

Center for Computation & Technology
Louisiana State University
jtao@cct.lsu.edu

DASPOS Workshop 2016

May 20th, 2016 Notre Dame

Table of Contents

Introduction

SIMULOCEAN Science Gateway

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

SIMULOCEAN

Home Software Tutorials About Contact

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 © 2012 - 2016, Grand Hyatt Research Collaboratory @ LSU

LSU XSEDE NSF

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

Search

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) SULIS (5) sediment transport (5) CI 2 SULIS (3) Oil spill (2) Modeling (2) Iau (2) models (2) evaluation metrics (2) informatics (2) UNO (2) Evapotranspiration (2) Visualization (2)

RELATED TOOLS TAGS

River Model (6) SULIS (5) Visualization Java (2) Simulocan (2) ASOS CERA (2) Coastal Data Factory (1) catalog (1) CLIP Data Extraction (1) DATA Extraction Mosaic Merging Preview Argo (1) data factory (1) Tool Clip Data Extraction (1) tools (1) Visualization Geoserver (1) talkoot (1) Storm Surge (1) Meeting (1) metadata (1) MOOSE CLIP (1)

ARCHIVE

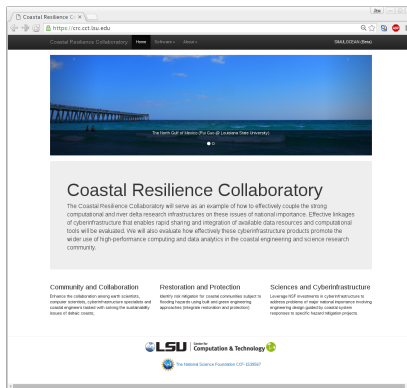
OCTOBER 2012 (3)
SEPTEMBER 2012 (3)
AUGUST 2012 (3)

WELCOME

Louisiana, Mississippi, and Alabama have formed the Northern Gulf Coastal Hazards Collaboratory (NGCHC), 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)



Our Vision

We envision SIMULOCEAN as

A Computational Platform

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

and

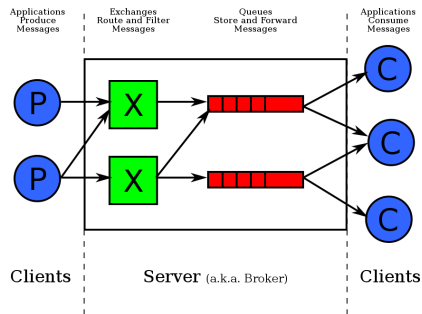
A Collaborative and Educational Environment

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

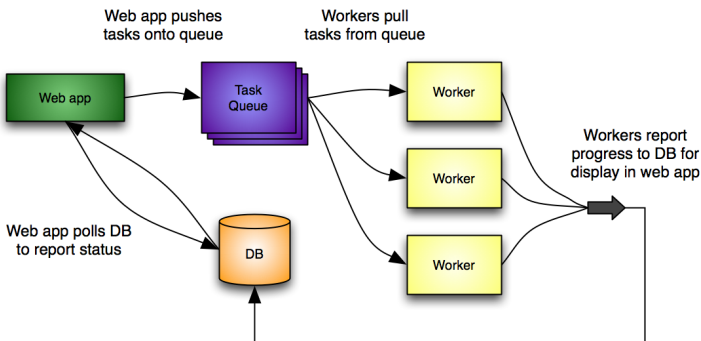
The Advanced Message Queuing Protocol (AMQP) is an open standard application layer protocol for message-oriented middleware (<http://www.amqp.org/>). **RabbitMQ** is one of several open source message broker software packages that implement AMQP (<https://www.rabbitmq.com/>).



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

Celery - Distributed Task Queue

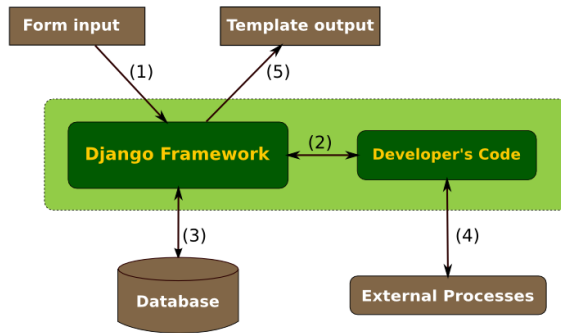
Celery is an asynchronous task/job queue based on distributed message passing (<http://www.celeryproject.org/>). It supports RabbitMQ and other message brokers.



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

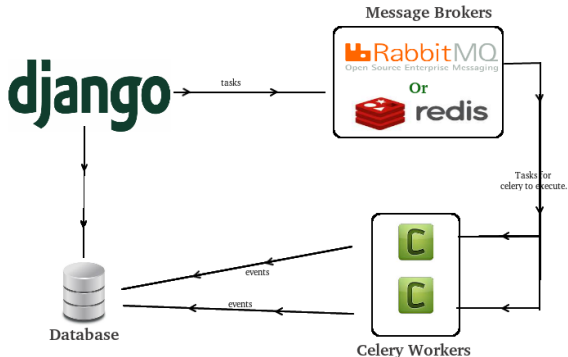
Django Web Framework

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



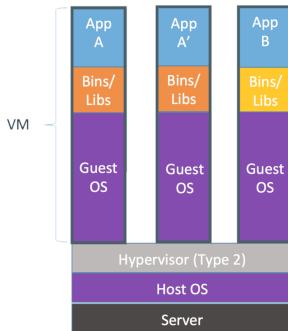
(image credit: <http://my-django-python.blogspot.com/>)

Containerization with Docker

Container v.s. Virtual Machine

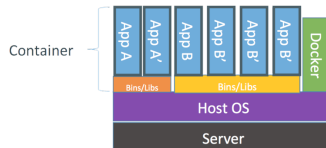
“Containerization is a lightweight alternative to full machine virtualization that involves encapsulating an application in a container with its own operating environment.”

— <http://www.webopedia.com/>



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

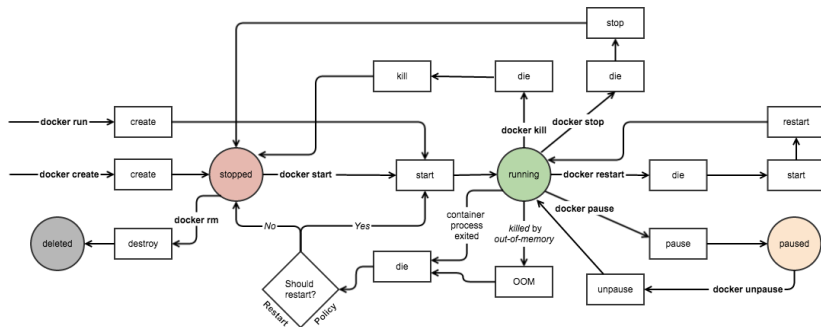
...result is significantly faster deployment,
much less overhead, easier migration,
faster restart



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

Docker

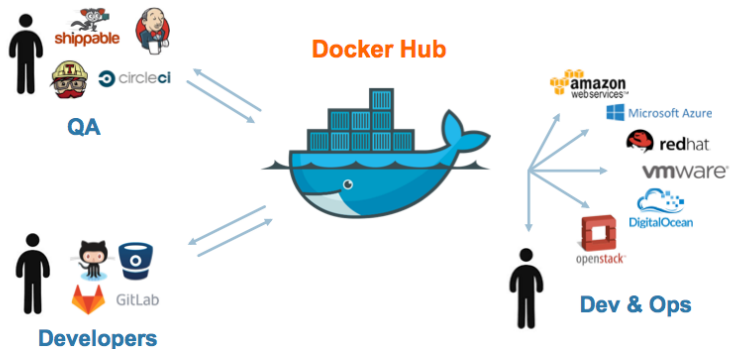
Docker - <https://www.docker.com/> is an open platform for distributed applications for developers and sysadmins. It provides an additional layer of abstraction and automation of operating-system-level virtualization on Linux.



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

Docker Hub

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

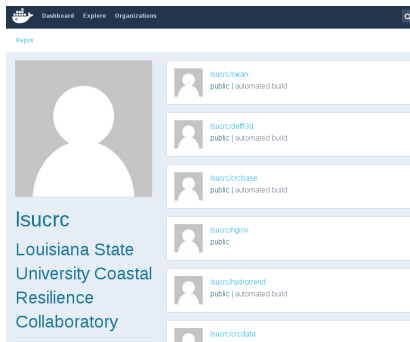


(image credit: <https://www.docker.com/>)

Scientific Application Repository & SIMULOCEAN

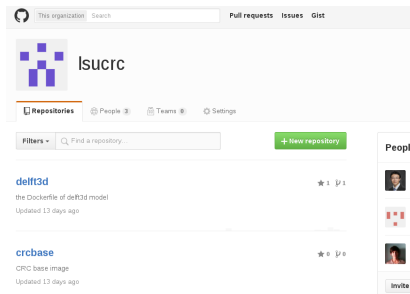
LSU CRC Docker Hub Repo

The Science Application 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.



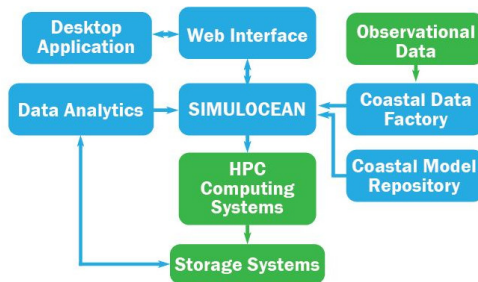
LSU CRC Github Repo

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



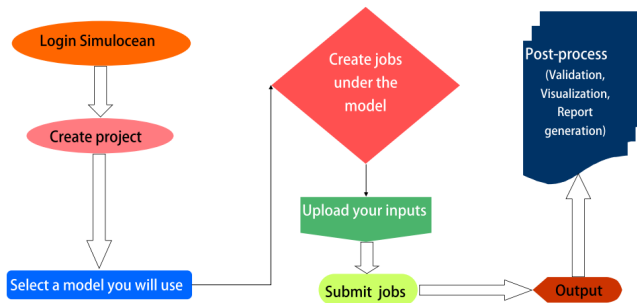
SIMULOCEAN Architecture

A platform for managing and deploying containerized 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 shows the SIMULOCEAN web interface. The browser address bar displays 'xseeds.simulocean.org/resource/list/'. The page has a navigation bar with links: Home, Applications, Resources, Projects, Intranet, Tools, Comments, Admin, Profile. The 'Resources' section is active, showing a list of resources under the 'Resources' tab. The list includes columns for Name, Provider, Type, Location, Available, Shared, Description, and Action. There are four entries listed:

Name	Provider	Type	Location	Available	Shared	Description	Action
CRC-head01	CRC	COMPUTE	LSU DMC 2009	Yes	Yes	This is the first CRC testing node with CentOS stable (766.5.6) installed.	Q / X
Jian-Localhost	Jian Tao	COMPUTE	LSU DMC 2009	Yes	Yes	This is Jian's laptop.	Q / X
CHARCOL	CRC	COMPUTE	CCT Machine Room	Yes	Yes	The computing and storage server for the CHARCOL team.	Q / X
Query	Indiana University	COMPUTE	Indiana University	Yes	Yes	XSEDE Gateway VM 182	Q / X

Showing 1 to 4 of 4 entries. Navigation links: First, Previous, Next, Last.

Copyright © 2012 - 2016, Coastal Hazards Research Collaboratory @ LSU. Logos for LSU, XSEDE, and NSF are visible at the bottom.

Scientific Applications

SIMULOCEAN provides an interface for managing scientific applications hosted.

The screenshot displays the SIMULOCEAN web application interface. The header includes the SIMULOCEAN logo and navigation tabs: Home, Applications, Resources, Projects, Ingers, Tools, Containers, Admin, Profile. The main content area is titled 'Application List' and shows a table of applications. The table has columns for Application, Version, Docker Repository, Time Created, Available, Shared, Description, and Action. Three applications are listed: SWAN, Del3D, and Hydr3D. Each application entry includes a brief description and a 'Show' button. The footer of the interface includes copyright information (Copyright © 2012 - 2016, Coastal Hazards Research Collaboratory @ LSU) and logos for LSU, XSEDE, and NSF.

Application	Version	Docker Repository	Time Created	Available	Shared	Description	Action
SWAN	41.81	lsu/swan	Dec. 31, 2015, 3:13 A.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
Del3D	5.81.00.2183	lsu/del3d	Dec. 14, 2015, 4:38 P.M.	Yes	Yes	Del3D is a flexible integrated 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
Hydr3D	3.0	lsu/hydr3d	March 4, 2016, 2:53 P.M.	Yes	Yes	Hydr3D is an ANSYS 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

First Previous 1 Next Last

Task/Container Management

SIMULOCEAN launches and monitors Docker containers on computing systems.

The screenshot displays the SIMULOCEAN web application interface. The top navigation bar includes links for Home, Applications, Resources, Projects, Inputs, Tasks, Containers, Admin, and Profile. The main content area is titled 'Task List' and shows a table of tasks. The table has columns for Task, Application Input, Machine, Container, Time Created, Description, and Action. Three tasks are listed: 'Der3D-TASK-1', 'SWAN-TASK-2', and 'Der3D-TASK-1'. Each task row shows the application input, machine, container name, creation time, and a description. The 'Action' column contains icons for search and delete. The footer of the interface includes copyright information and logos for LSU, XSEDE, and NSF.

Task	Application Input	Machine	Container	Time Created	Description	Action
Der3D-TASK-1	Der3D-INPUT-1	Query	ipao-Der3D-TASK-1-4403 ipao-Der3D-TASK-1-4403	May 6, 2016, 12:29 a.m.		🔍 ✖
SWAN-TASK-2	SWAN-INPUT-1	Query	ipao-SWAN-TASK-2-0003 ipao-SWAN-TASK-2-0003	May 6, 2016, 12:22 p.m.		🔍 ✖
Der3D-TASK-1	Der3D-INPUT-1	Query	ipao-Der3D-TASK-1-4403 ipao-Der3D-TASK-1-4403	May 10, 2016, 6:20 a.m.		🔍 ✖

Showing 1 to 3 of 3 entries

First Previous 1 Next Last

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

LSU Center for Computation & Technology XSEDE NSF

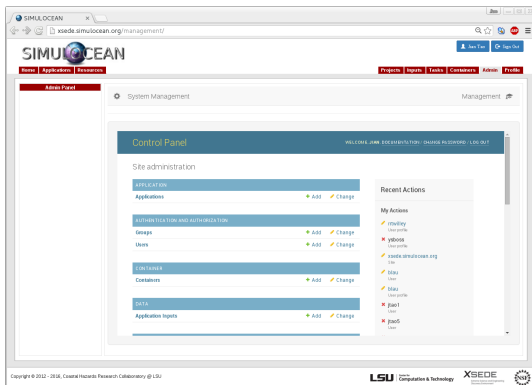
Data Container

For each task, a data container that runs a light-weight web server will be launched to serve data back to users.

The screenshot shows the SIMULOCLEAN web interface. The browser address bar displays `xsede.simuloclean.org/container/data/U41/`. The page has a navigation bar with links: Home, Applications, Resources, Projects, Systems, Tools, Containers, Admin, Profile. The main content area is titled "Containers" and shows a container named "Container Jiao-DeR3D-Task-1-Data". A green status bar indicates: "The Container has been running since May 6, 2016, 11:28 a.m. (3 days, 23 hours ago)". Below this, there is a section titled "Index of /" showing a directory listing of files: `conf112_files.tar.gz` (241 KB), `conf12_files.tar.gz` (533 KB), and `conf13_files.tar.gz` (206 KB). A table below the index shows properties for the container: `id` (41), `Name` (jao-DeR3D-TASK-1-Data), `User` (1), `Folder` (users/jao-DeR3D/TASK-DeR3D-TASK-1), and `Time Created` (2016-05-06 11:29:49+00:00). The footer contains copyright information (© 2012-2016, Central Hazard Research Collaboration @ LSU), logos for LSU, XSEDE, and NSF.

System Administration

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



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

LSU Center for Computation & Technology

XSEDE

NSF

LSU Center for Computation & Technology

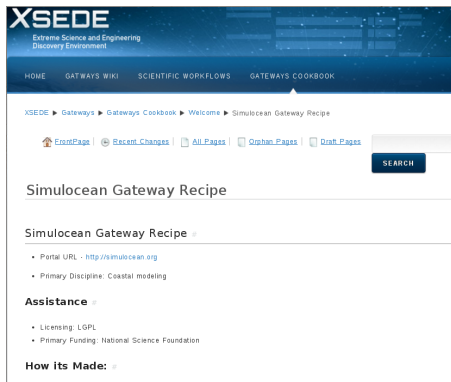
XSEDE
Extreme Science and Engineering
Discovery Environment



Future Plans

Science Gateway on NSF Cloud

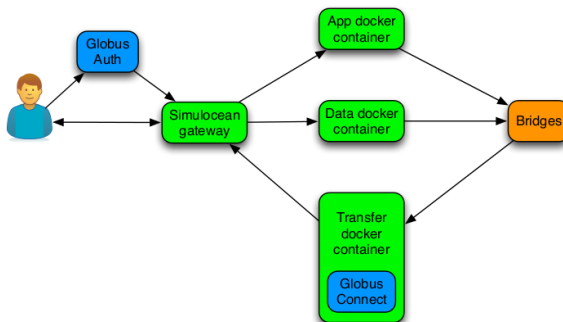
SIMULOCEAN is now also supported by XSEDE ECSS program. Current ECSS Experts include: Stu Martin and Eric Blau (Globus Team), Mona Wong (SDSC)



The screenshot shows the XSEDE website interface. The header features the XSEDE logo and the tagline "Extreme Science and Engineering Discovery Environment". Below the header is a navigation bar with links: HOME, GATEWAYS WIKI, SCIENTIFIC WORKFLOWS, and GATEWAYS COOKBOOK. The main content area displays the breadcrumb trail: XSEDE > Gateways > Gateways Cookbook > Welcome > Simulocean Gateway Recipe. Below this is a search bar and a "SEARCH" button. The title "Simulocean Gateway Recipe" is prominently displayed. Underneath, there is a section for "Simulocean Gateway Recipe" with a list of items: Portal URL - <http://simulocean.org> and Primary Discipline: Coastal modeling. Further down, there is an "Assistance" section with items: Licensing: LGPL and Primary Funding: National Science Foundation. At the bottom, there is a "How its Made:" section.

Authentication with Globus Connect

We work with the Globus (<https://www.globus.org/>) team to provide authentication services with Globus Connect for SIMULOCEAN and potentially other XSEDE Science Gateways.



(image credit: Mona Wong from SDSC)

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