



# CoESRA: From Virtual Desktop to Science Gateway

Hoang Anh Nguyen, David Abramson, Siddesware  
Guru, Yi Sun, Chris Hines & Jason Rigby

GCE16 Presentation  
by Hoang Anh Nguyen



# Outlines



- CoESRA
- Desktop-as-a-Service Challenges
- CoESRA 2.0: a Science Gateway
- Conclusions

# CoESRA



- **Collaborative Environment for Ecosystem Science Research and Analysis**
  - Initiated by TERN in mid 2014
  - Environment for synthesis and analysis of scientific data for the ecosystem science community
- **Aims**
  - Environment for repeatability, reproducibility
  - Lower barriers to start computational experiments

# CoESRA



- Use common research infrastructure in Australia
  - Australian Access Federation (AAF) for authentication - federated identity
  - Cloud-based research infrastructure
    - National eResearch Collaboration Tools and Resources (NeCTAR) Cloud for computing
    - Research Data Storage (RDS) for storage
- Desktop-as-a-Service
  - Ready-to-use Linux desktops via Web (Guacamole)
  - Kepler scientific workflow
  - Nimrod/G for job distribution

# CoESRA

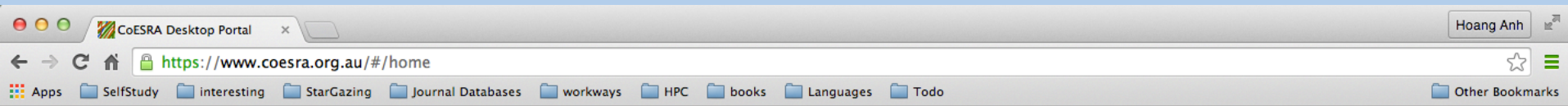


- CoESRA 1.0
  - Each desktop is a Cloud node.
  - Each node is launched and destroyed automatically based on system load.
    - Puppetize at launch
  - Each session lasts 2 days
    - Manual extension
  - In production since early 2015
  - APIs are not exposed

<https://www.coesra.org.au>



# CoESRA 1.0



*Collaborative Environment for Ecosystem Science Research  
and Analysis*

[Home](#) [About ▾](#) [Features](#) [FAQ](#) [Policy](#) [Logout](#)

[Go to your Desktop](#)



**NCRIS**  
National Research  
Infrastructure for Australia  
An Australian Government Initiative



TERN and ANDS are supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS)

Copyright © 2015 TERN (University of Queensland). All Rights Reserved.

# CoESRA 1.0



CoESRA Desktop Portal

Hoang Anh

<https://www.coesra.org.au/#/home>

vm428.coesra.org.au

<https://vm428.coesra.org.au/guacamole/#/client/c/vm428.coesra.org.au>

Applications Menu

Unnamed1

Applications Menu

Home

File System

Unnamed1

File Edit View Workflow Tools Window Help

Tag workflow: select or type tag and press enter

View: Workflow

Components Data Outline

Search Components

Advanced Sea... Sources Cancel

All Ontologies and Folders

- Components
- Projects
- Statistics
- Demos
- Actors
- Coesra-term
- Dataturbine
- Directors
- Job
- MyWorkflows
- NimrodK
- NimrodK-microscopy
- NimrodK-neuroimaging
- Opendap
- Outreach
- Provenance
- R

User Instructions

- Click
- User
- Save
- If you
- The v
- sent
- No p
- All us
- Plea

CoESRA

Client for Ecosystem and Analysis

gout

ail will be

# CoESRA 1.0



- Limitations
  - Not flexible enough to serve different configurations of desktops
  - Complicated to support multiple 'flavors'
- CoESRA 1.1



# CoESRA 1.1



- CoESRA 1.1
  - Virtual Cluster on Cloud.
  - Based on Strudel-web \*
  - Users specify node information as if submitting a batch job
  - APIs are OAuth2-compliant
- In production since March 2016.  
<https://portal.coesra.com/strudel-web>

<https://www.massive.org.au/userguide/cluster-instructions/strudel-web>

# CoESRA 1.1



Program | ScienceGateways.org x ScienceGateways.org x Strudel Web x

https://portal.worqways.com/strudel-web/#/desktop-manager

Apps SelfStudy interesting StarGazing Journal Databases workways HPC books Languages Todo Apache Spark course User-friendly langu... Other Bookmarks

Collaborative Environment for Ecosystem Science Research Home About Features FAQ Policy Desktop Logout

h.nguyen30@uq.edu.au

CoESRA

Ecosystem Science Analysis and Synthesis

COESRA



## Launch a desktop

Nodes	Processors per node	Memory (gb)	Time (hours)	
1	2	8	48	Project ▾

LAUNCH

## Running desktops



#	Time remaining	State	
105	1-23:59:50	✓	<a href="#">SHOW DESKTOP</a> <a href="#">STOP DESKTOP</a>

## Server messages

Any messages from the server will be displayed here

NCRIS  
National Research  
Infrastructure for Australia  
An Australian Government Initiative



# Up until now...



- Desktop-as-a-Service model
  - Web-based access Linux desktop
  - Kepler as scientific workflow
  - Nimrod/G as distributed middleware
- Data management
  - File based
  - Private & Public folders
- Sharing is possible
  - But not easy

# Challenges



- Learning curve
  - Linux
  - Kepler
  - Nimrod/G
- Troublesome to perform large-scale experiments
- Workflow sharing
- Data management

# CoESRA 2.0

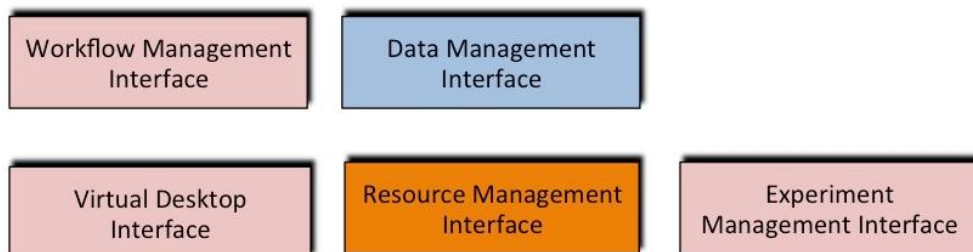


- Aims to address (most of) limitations
- Extends Strudel-web API
  - Workflow management
  - Execution management
  - Data management
  - Resource management

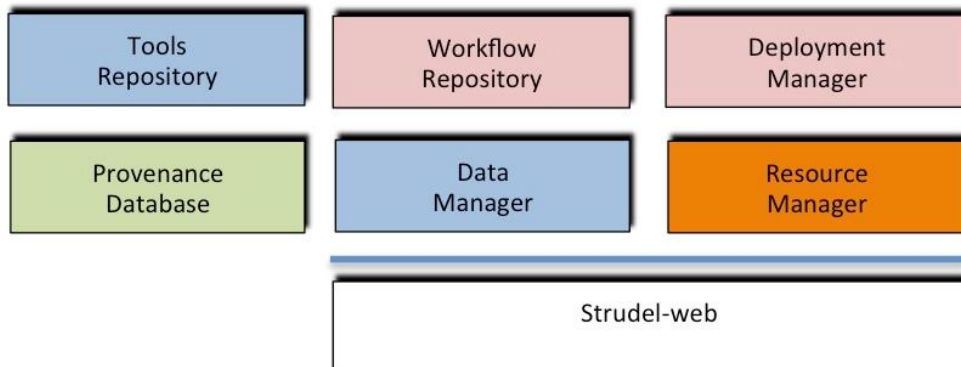
# CoESRA 2.0



## Web components



## Web Services



## Resource Access





# CoESRA 2.0



- Workflow Repository
  - Central repo for workflows
  - Easier for sharing, versioning
- Deployment management
  - Execute workflows from the Web
    - Workflow execution wrapped as SLURMM job
  - Virtual desktop is still available
    - Advanced users
    - Workflow developers

# CoESRA 2.0



- Data management
  - Input/output data are in NeCTAR object store
  - Data catalogues based on executions
- Resource management
  - Seamless integration with existing HPCs
  - Usage reports
  - Work-in-progress
- Expected to roll out early 2017

# Conclusions



- CoESRA
  - Version 1.0
  - Version 1.1
- Desktop-as-a-Service Limitations
- CoESRA version 2.0

# Acknowledgement



## Project Sponsors



## Collaborators



# Acknowledgement



- Siddesware Guru - project manager
- Yi Sun - Software developer
- Chris Hines & Jason Rigby - Strudel-web developers



# QUESTIONS