Collaborative Infrastructure to Integrate Local and Cloud Resources and Ease Scientific Collaboration

K.-S. Kuo^{1,2}, R. Ramachandran³, M. Maskey⁴, C. Lynnes², N. Niamsuwan⁵

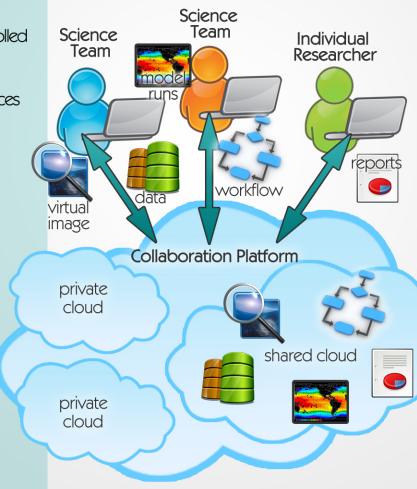
¹Bayesics, LLC; ²NASA Goddard Space Flight Center; ³NASA Marshall Space Flight Center; ⁴University of Alabama-Huntsville; ⁵NASA Jet Propulsion Laboratory

Collaborative Workbench (CWB) to Accelerate Science Algorithm Development

Sharing Knowledge is at the heart of science, yet it is challenging for researchers to effectively share information and tools

Goals

- An architecture for scalable, controlled collaboration
- Selective sharing of science resources
 - among individuals
 - within science teams
 - with the entire science community.
- Software that fits how researchers currently do scientific analysis to promote adoption



Benefits

- Accelerate science algorithm development by
 - distributed science teams
- Reduce redundancy
- Jmprove productivity
- Securely share all science artifacts (data, information, workflow, virtual machines)
- Generalizable to support collaborative science algorithm development for other mission and model enterprises

Significance of Collaborative Workbench (CWB)

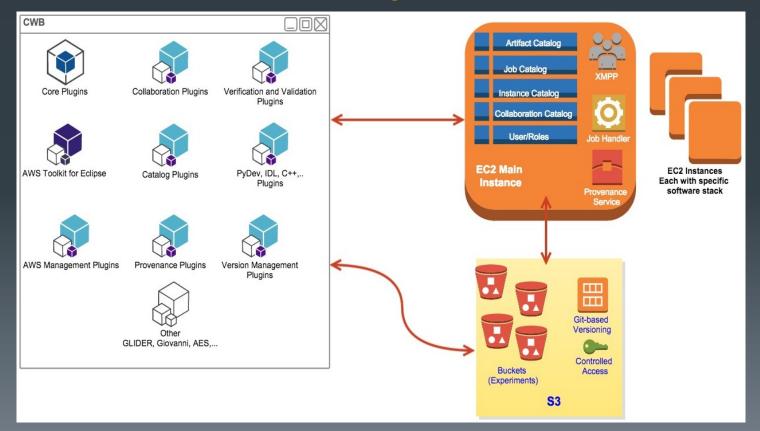
Earth Science is a science for "system of systems."

- Earth scientists are required to collaborate outside of their traditional focus areas.
- Collaborations are increasingly geographically distributed across nations and time zones.
 - GPM (and TRMM) is a prime example.
- We need to accommodate various modes of collaboration, in order for it to be effective.
- Collaboration needs to be made easy so scientists can focus on science, rather than learning collaboration tools.

Outline

- CWB core components
- Logging on to CWB infrastructure.
- Creating experiments and workflows.
- Collaborating asynchronously.
- Collaborating synchronously.
- AES Integration
- NEOS³ Integration Creating NEOS³ jobs
- NEOS³ Integration Obtaining NEOS³ input and output

CWB Core Components



Logging On

- CWB seamlessly integrates your local compute resources with those in Cloud.
 - A CWB server keeps track of collaboration group membership and Cloud resources, so an account (and logging on) is required.
- An experiment is composed of one or more workflows.
- A workflow is a sequence of one or more programs (applications), which are executed in order to accomplish a specific task.
- Three (3) "workspaces"
 - Shared Experiments: Experiments shared by others
 - *My Experiments*: My personal workspace
 - Other Experiments: Experiments created by others but no workflow has been shared.

Logging On

facts.xml figuration	Today, 7:47 PM Today, 11:37 AM Today, 11:37 AM Today, 11:37 AM Today, 11:36 AM Today, 11:37 AM	42 KB 7.4 MB
figuration	Today, 7:47 PM Today, 11:37 AM Today, 11:36 AM	7.4 MB
,	Today, 11:37 AM Today, 11:36 AM	7.4 MB
•	Today, 11:36 AM	
gins	10day, 11:37 AM	
		cwb
1 of 6 selected, 240.7 GB av	ailable	
osh		osh HD + 👔 Users ≻ 🏠 mmaskey + 🔄 Desktop + 📺 user1 + 🎦 cwb + 💟 1 of 6 selected, 240.7 CB available

Creating Experiments and Workflows

- Code and data can be dragged and dropped into CWB personal workspace.
- Code may be executed *either* locally *or* in Cloud.
- Users can choose different Cloud instances (different software configurations, e.g. different IDL versions).

Creating Experiments and Workflows

00	Col	laborative Workbench – /Us	ers/mmaskey/Desktop/user1/cwb/cwb.app/Contents/MacOS/workspace - user1 [user1@abc.com]	R ₂
	× A• 图• 图• 🌣			Q Quick Access
m Search 🖾		- 0		- 0
Search Workflows		Search		
○ Shared Experiments ▶ ■ neos1	23	<u> 2</u> - D		
My Experiments 💱 🕨 🔒 neos1 – Personal		e 🌾 🔻 🗖		
Other Experiment: aogs-demo myexp1 shree-experimen	Copy 9C Paste 9KV Add Experiment Upload Share Execute Workflow Refresh	1# - -	Program List 🖄 Contacts 🥃 EC2 AMI View 🎯 EC2 Instance View 🗟 Console 💈 No consoles to display at this time.	≝ 6 + 6 + ■ ■

Collaborating Asynchronously

 Members of the same group can execute shared experiments and workflows at their own convenience.

Collaborating Asynchronously

$\Theta \Theta \Theta$	Collaborative Workbench - /Us	ers/mmaskey/D	esktop/user1/cwb/cwb.app/Contents/MacOS/workspace - user1 [user1@abc.com]	× ^N
🔛 🕼 🗁 🖟 🗈 🏠 🕱 🖋 🖓 • 💱 • 🌾				Q Quick Access
∰ Search ⊠	- 0	Separate Experiment	mar_bells_surface.pro 🛛	- 0
Search Workflows	Search	PRO mar_bel ; FYI PRINT, 'Ins	ls_surface ide PRO MarBellsSurface'	
C Shared Experiments II ▶ ■ neos1	S 4 9 0	RESTORE, 'm ; Display t s = SURFACE , TITLE ; Overlay t c = CONTOUR , N_LEV , ZVAL , PLANA	TELS = 15 \$ UE \$ R = 0 \$	
My Experiments My Experiments My Experiments And And And And And And And And And		, /BUFF , /OVEF ; Save an i s.Save, 'ma END	PLOT)	
Other Experiments ags-demo myexp1 shree-experiment		Program List	☆ Contacts 🖬 EC2 AMI View 🥸 EC2 Instance View 🖾 Console 🛛	: ⊡• t3• ¤ □

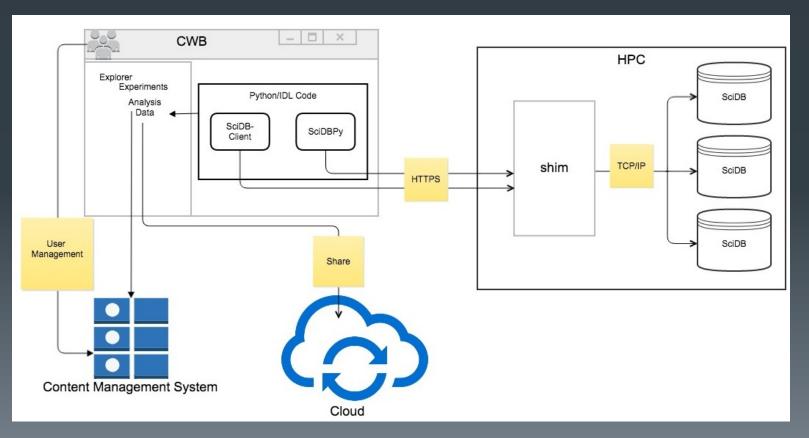
Collaborating Synchronously

- Often asynchronous collaboration is inadequate and synchronous collaboration is required.
- Users need to sign into CWB and connect with an XMPP server (provided by CWB infrastructure).
 - XMPP: eXtensive Messaging and Presence Protocol
- Once connected, multiple users in the same group can chat and share in real time.
- Two (2) members at a time can engage in *collaborative programming*.
 - In the upcoming clip, user1 is on the left and user2 on the right.

Collaborating Synchronously

○ ○ ○ Collaborative Workbench - /						
∰ Search ⊠ □	Experiment	mar_bells_surface.pro 🛿	Q Quick Access			
Search Workflows Search	PRO mar_bell					
Shared Experiments Shared Experiments	; Define the RESTORE, 'mc ; Display th s = SURFACE(, TITLE ; Overlay th c = CONTOUR(e elevation data to use. rbells.dat' rbells.dat' e levation surface in a buffer in e 'Naroon Bells Elevation Data') te contour data on the elevation su clev \$ ULS = 15 \$ E \$				user1
My Experiments 22 S S S S S S S S S S S S S S S S S S	, /BUFFE , /OVERF ; Save an im s.Save, 'mar END	PLOT)				demo-data
Other Experiments 23 aogs-demo myexp1 shree-experiment	No consoles to disp	play at this time.	े २ - ८२ -			
					_	

CWB-AES Integration



Loading edu.uah.tsc.cmac3	<section-header><text></text></section-header>
	Login to CWB
COLLABORATIVE WORKBENCH	Username RxR Password I UAH SciDB Server Server SciDB 1 SciDB 2 Login Cancel New Account Copyright: ITSC, University of Alabama in Huntsville

Starting CWB

Authentication and authorization handled by CWB

Point to the appropriate HPC End Point

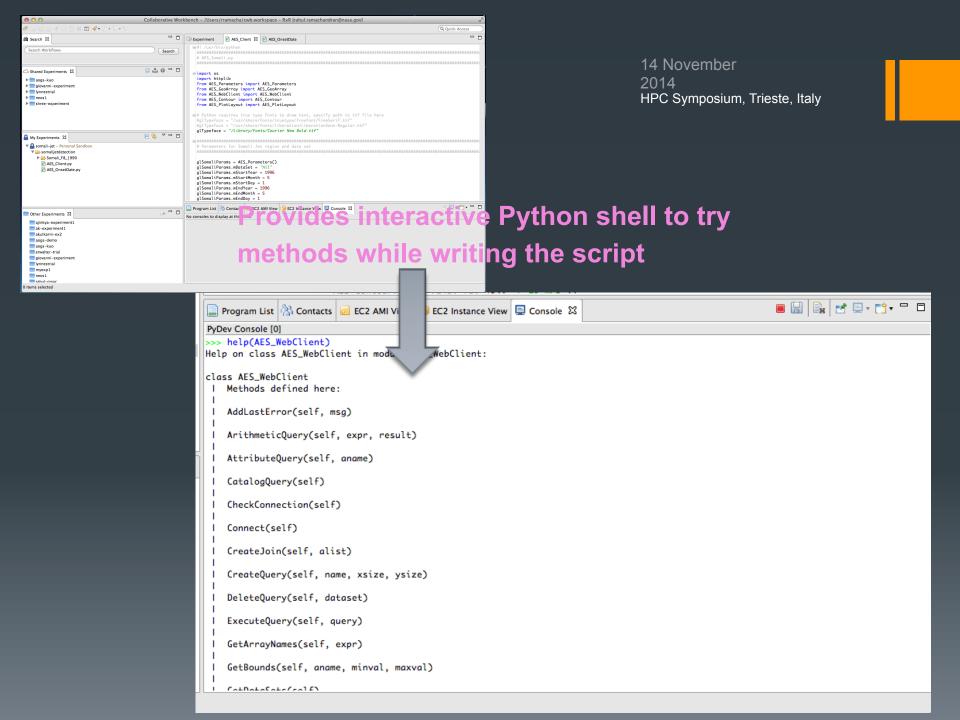
00	Collaborative Worl	kbench – /Users/r	rramacha/cwb.wo	rkspace - RxR [rahul	l.ramachandran@nasa.g	lov]
🧶 🖫 🖻 🐇 🖹 🏦 🗶 🖬 🖋 • 🖢 • 🖓 • 🖓 • 🏷						Q Quick Access
曲 Search 🕱	- 0	Experiment	P AES_Client 🖾	P AES_OnsetDate		
Search Workflows	Search	# AES_Soma1	######################################			,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
Shared Experiments acceleration accelerati		from AES_GG from AES_WG from AES_CG from AES_P1 #glTypeface glTypeface glTypeface #glTypeface	arameters import eoArnay import A ebClient import A lotLayout import equires true type e = "/usr/share/ = "/Library/Fon	ES_GeoArray AES_WebClient S_Contour AES_PlotLayout e fonts to draw te: fonts/truetype/free fonts/liberation/Li ts/Courier New Bold	xt, spec USE da efont/FreeSerif,ttf" iberatio Writer d.ttf" #######CUIERY	rt AES Library and Il the functions to my own script to "data, and execute"
 ▼ Somalijetdetection ▶ Somali_F8_1990 P AES_Client.py P AES_OnsetDate.py 		glSomaliPar glSomaliPar glSomaliPar glSomaliPar glSomaliPar glSomaliPar glSomaliPar	rams = AES_Param rams.mDataSet = rams.mStartYear rams.mStartMonth rams.mStartDay = rams.mEndVear = rams.mEndMonth = rams.mEndMonth =	eters() "nil" = 1996 = 5 1 1996 5	optim algor	
Other Experiments ajinkya-experiment1 ak-experiment1 akulkarni-ex2 aogs-demo aogs-kuo erwalter-trial giovanni-experiment lynnestrial myexp1 neos1 rahul-cmac 0 items selected		No consoles to disp		AMI View 🧐 EC2 Insta	nce View 📮 Console 🔀	

xperiment P AES_Client 🔀 P AES_OnsetDate		14 November	' 🗆	4. Ne
<pre># Create an http connection client = AES_WebClient() #client.Connect("/Users/jrushing/Projects/N client.Connect()</pre>	IASA/Connect/tremor3.txt")	: File	₩Z ₩S	, Italy not 6 6. Le input 7. Fo
<pre># Generate a subset subsetId = GenerateSubset(client, sat, year if (subsetId is None): return False</pre>	, month, day) Open Show		• •	thing 8. Do herd 9. De
<pre># Generate a subset projectId = client.ProjectQuery(subsetId, " if (projectId is None): return False</pre>	Paste	y Context Qualified Name	#X #C #V	pers othe 10. B othe
<pre>print "SomaliJet - Generated Project Array # Get the data array for the subset npArray = client.GetSlice(projectId, 0) if (npArray is None): print "Error: Could not get data for "</pre>	Quick Shift R Shift L	light .eft	₩1	
return False	Debug Run A			a ^r 1 Duthon Dun
<pre># Convert to AES geo array dataArray = AES_GeoArray() dataArray.InitFromNumpyArray(npArray) # Get month string mstr = "May"</pre>	Team Compa Replac PyDev	are With ce With	* * *	 I Python Run Python unit-test Run Configurations
<pre>if (month -= 6): mstr = "June"</pre>	Proper	· ·	₩I	

- Execute optimized data parallel queries on a remote HPC
- Integrate local data processing and analysis seamlessly

Run my script from CWB

My Experiments	
Somali-jet - Personal Sandbox	14 November
Somalijetdetection	2014
Somalijetdetection	
# onset_1990_05_01.png	HPC Symposium, Trieste, Italy Results saved in
<pre># onset_1990_05_02.png # onset_1990_05_03.png</pre>	"My Experiment"
🇰 onset_1990_05_04.png 🗰 onset_1990_05_05.png	workspace
	workspace
<pre>monset_1990_05_06.png</pre>	
<pre>monset_1990_05_07.png monset_1990_05_08.png</pre>	
■ onset 1990 05 08.nnd	Collaborative Workbench – /Users/rramacha/cwb.workspace – RxR [rahul.ramachandran@nasa.gov]
□ 값 스 ☆ 目 含 × 営 ペ・ 約・ 和・ ♡	
a Search 원	Image: Second
(Search Workflows	Search $(\Rightarrow \Rightarrow \Rightarrow B) \oplus (\bigcirc (10, 10))$
	Search 25
 Shared Experiments № aogs-kuo giovanni-experiment lynnestrial neos1 shree-experiment anset_1990_05_22.png onset_1990_05_23.png onset_1990_05_23.png onset_1990_05_25.png onset_1990_05_25.png onset_1990_05_26.png onset_1990_05_28.png onset_1990_05_28.png onset_1990_05_28.png onset_1990_05_30.png Statistics.txt 	20 20 5 C C C C C C C C C C C C C C C C C C C
i statistics.txt	📄 Program List 🚷 Contacts 🥃 EC2 AMI View 🎯 EC2 Instance View 🖳 Console 🕄 📑 👘 🗂
🛑 Other Experiments 🔀	No consoles to display at this time.
ajinkya-experiment1 ak-experiment1 akulkarni-ex2 aogs-demo aogs-kuo erwalter-trial giovanni-experiment lynnestrial myexp1 neos1 ros1	Visualize plots in CWB



NEOS³ Integration Creating NEOS³ Jobs

- NEOS³: NASA Earth Observing System Simulation Suite
 - A project Pl'ed by Simone Tanelli of JPL and funded by NASA Advanced Information System Technology (AIST) program.
 - Web-based integrated simulator for Earth remote sensing applications.
 - Equipped with start-of-the-art modules to enable the realistic simulation of satellite observables.
 - Providing an advanced, sophisticated, and user-friendly simulator package to be used by both *scientists* for research-oriented applications and by *system engineers* for an instrument design purpose.
 - Accessible via a web interface and capable of distributing computationally intensive tasks to remote servers such as those at the NASA Advanced Supercomputing (NAS) Division.

NEOS³ Integration Creating NEOS³ Jobs

00	Collaborative Workbench - /Us	ers/mmaskey/Desktop/user1/cwb/cwb.app/Contents/MacOS/workspace - use	erl [userl@abc.com]
🔡 🕲 🛃 👔 💼 🗶 🛷 - 🔄 - 🖏 - 🏷			Q Quick Access
曲 Search 器	- 0		
Search Workflows	Search		
☐ Shared Experiments 🛛	3 초 - 0		
▶ meos1			
A My Experiments 🕴			
Incos1 - Personal Sandbox			د ۵. ۲
		Program List 🖄 Contacts 🥃 EC2 AMI View 🎯 EC2 Instance View 📮 Console 🛱 No consoles to display at this time.	00.0.0
Other Experiments aggs-demo myexp1 shree-experiment	j∌ ^m B		

NEOS³ Integration Obtaining NEOS³ Input/Output

- A NEOS³ job may take hours.
- NEOS³ module polls the server using Job ID until job is complete.
- Job output can then be retrieved using a function provided in NEOS³ module.

NEOS³ Integration Obtaining NEOS³ Input/Output

00	Collaborative Workbench - /Use	ers/mmaskey/Desktop/user1/cwb/cwb.app/Contents/MacOS/workspace - user1 [user1@abc.com]	R. P.
김 영 쇼 종 自 首 × 🔗 · 친 · 친 · 🌣			Quick Access
曲 Search 🛛	- 0		- 0
marbells	Search		
Shared Experiments 🕺	S 🛨 🗖 🖸		
▼ aogs-kuo ▶ 1 user1 ▶ neoS1			
A My Experiments 🔀			
A aogs-kuo - Personal Sandbox A myexp1 - Personal Sandbox A neos1 - Personal Sandbox			
		📄 Program List 🥃 EC2 AMI View 🎯 EC2 Instance View 🖳 Console 🕱	3 🕒 • 📬 • 🗖 🗖
Other Experiments Other Experiments acqs-demo shree-experiment	(* ⁻		

Epilogue

- CWB implements automatic versioning (in the background) by integrating GitHub.
 - Similar to GoogleDoc, you can always go back to a previous version.
- CWB has integrated
 - Automated Event Service.
 - NASA Earth Observing System Simulator Suite, NEOS³.
 - NASA GES DISC Giovanni service.
- CWB is currently in alpha test.

Acknowledgement

- NASA Computational Modeling Algorithms and Cyberinfrastructure (CMAC) program
- NASA High-End Computing (HEC) program
- NASA Earth Science Technology Office (ESTO)

Adding Workflow(s) in Experiment(s) of Others

Adding Workflow(s) in Experiment(s) of Others

