

2018.01-2018.03 Q5 Quarterly Summary Report

- 2018 Q1 (January, February, March) Work Progress at a Glance
- Report summary of AIMS progress for the quarter period: January 1 – March 31, 2018
 - Management
 - Publishing
 - Installation
 - Server-side computing
 - Security
 - Machine Learning
 - ESGF Services
 - CDAT
 - VCDAT
 - Hardware/Infrastructure
 - AIMS Publications –
 - Conferences –
- Next Quarter –
 - Management
 - Installation
 - Server-side computing
 - Machine Learning
 - CDAT
 - Publications
 - Hardware
 - ESGF Services

2018 Q1 (January, February, March) Work Progress at a Glance

Color scheme:

In Progress	Completed	Behind Schedule	On Hold
-------------	-----------	-----------------	---------

Task	Q4			Q5			Q6		
	2017-07	2017-08	2017-09	2017-10	2017-11	2017-12	2018-01	2018-02	2018-03
1 Project Management									
1.a Annual ESGF F2F Conference		Agenda & abstracts	Agenda & abstracts		Prepare for ESGF F2F	Successfully Executed ESGF F2F	XC planning next F2F		
1.b Prioritize agency tasks			DOE, NASA, CRIM	On hold: Needs agency clarification					
1.e ESGF proposal	DOE proposal		DOE proposal accepted, R&D 100			Winner of R&D 100 Award			LLNL recognition for R&D 100 Award
1.c Strategic plan	Inter- agency strategic plan	DOE Strategic Plan	Inter- agency strategic plan	On hold: Needs ESGF-XC clarification					
1.d Implementation plan		DOE implementation plan	Inter- agency implementation plan	On hold: Needs ESGF-XC clarification					
2 Data Management (publication, replication, notification, feedback)									
2.a Publishing data to ESGF	Preparing for CMIP6, Obs4Mips, Input4MIPs			Preparing for CMIP6, Obs4Mips, Input4MIPs,		Performance Testing of PrePARE	Input4MIPs publication to produciton	Data challenge CMIP6 test publication	CMIP6, obs4MIPs test publication
2.a Official release of CMOR	CMOR v3.2.5	CMOR v3.2.6	CMOR v3.2.7		CMOR v3.2.8	CMOR v3.3.0		CMOR v3.3.0	CMOR v3.3.1
2.b Standardized distributed reanalysis data									
2.c Integrate heterogeneous data			Hydrology data			Biology test sets			
2.d Translate different archives and data stores			Lightweight data service			Data service server-side			

2.e Automated data replication & publication			Beta testing w/DKRZ, CEDA, IPSL	Ongoing				
2.f Notification and feedback				Ongoing				Development of testing framework and search capability
3 User Interface and Search								
3.a Enhance search			For CMIP6	Preparation for For CMIP6			Waiting on NASA JPL to approve subcontract	
3.b Enhance CoG UI			For CMIP6	Preparation for For CMIP6				
4 Hardware and Network								
4.a New petabyte storage		Purchased 4PB	Integrated 4PB into CSS					Investigating next storage procurement
4.b Network speed up			High-speed data syncing					
4.c Data Transfer Nodes (DTNs)			Performance Tuning for CMIP6					
5 Installation								
5.a Official release ongoing and new operations	ESGF major release v2.5	Minor release v2.5.13	Minor release v2.5.14	Minor release v2.5.17	behind on v.3.0.0 release	Roadmap to release v3.0.0	2.6x devel release testing	2.6.5-7 master releases
5.b New Python-base installation script								v3.0 alpha release (not feature complete)
5.c Installer UI			Prototype	On hold				
6 Server-side computing								
6.a CoG UI integration		Development	Deploy & testing			Testing, Integration		
6.b User metrics tracking								
6.c Docker support						Working toward ESGF docker-based integration		
6.d Backend compute engines			EU's Ophida NASA's EDAS			In place for F2F		
7 Authorization and Authentication								
7.a Manage CA certificates			Manage d for LLNL, NCCS, DIASJP, CMCC, others	Ongoing			Ongoing	
8 Analysis & Visualization Tools								
8.a Official release of Community Data Analysis Tools (CDAT)			CDAT Minor release v2.12			Python 3.0 port and VCDAT first set of requirement completed for F2F		CDAT v8.0 released. Python 3 support, name changed from UV-CDAT to "CommunityData Analysis Tools", CDMS 3.0, cdtime 3.0, genutil 8.0, cdutil 8.0, vcs 8.0, removed Vistrail, added VCDAT 1.0 beta release
8.b Community Data Management System (CDMS)	Reads non-CF compliant files	Integrated ViSUS Reader into CDMS	Port to Python 3	Documentation and final Python 3.0 tweaks				Continue updates to documentation
8.c Visualization Control System (VCS) Speed up				Bug fixes and implement features as needed by E3SM and PMP, feature frozen by F2F.				
8.d VCS additional graphics methods	Streamlines		Error bars					
8.e Color maps			Static color maps for CMIP & E3SM					
8.f Official release of the Community Diagnostics Package (CDP)	CDP major release v1.0	CDP major release v1.1						

8.g CDP parallelism			Large scale parallelism						
8.h CDP input parsing			JSON and CFG						
8.i Visual Community Data Analysis Tools (vCDAT)	Development	Development	Development	Dean First set of requirements attained. New set identified					Released v1.0 beta with CDAT v8.0
8.j vCDAT file and dimension editor									
8.k vCDAT VTK-web									
8.l Animations & movies		List and prioritized animation	Animate hurricanes & human fingerprint						
8.m test suite	Unit tests / Performance test			All repo have CI test suite. vcs-js needs one.					continuous testing with CI
8.n Tutorials and online support		Jupyter Notebook for Taylor Diagram	Jupyter Notebook for CDMS101	<ul style="list-style-type: none"> • CMIP3 results • Polar plots use • Animation in vcs • Using shading and patterns • Portrait Plots 					New Jupyter Notebooks: 1D plots, Text objects
9 Outreach									
9.a Publications	8 publications						ML publications		
9.b Presentations		101 Tech Briefing	DOE ASCR NGNS Pi Meeting			F2F 2017			
10 Machine Learning									
10.a CAM5 tropical cyclone tracking (w/ LBNL)						published in DLPS			
10.b Tracking using LSTM									
10.c Connecting multiple locations									

Report summary of AIMS progress for the quarter period: January 1 – March 31, 2018

ESGF

Members: Dean N. Williams (ESGF Chair and PI); Sasha Ames, Jason Boutte, James Crean, Charles Doutriaux, Cameron Harr, Matthew Harris, William Hill, Tony Hoang, Sookyung Kim, Ji-Woo Lee, Lina Muryanto, Denis Nadeau, Jeffrey Painter, Zeshawn Shaheen, Argonne, JPL, Kitware

Management

- Planning for new hires: developed recruiting postings
- Planning ESGF F2F for 2019

Publishing

- input4MIPs production republication
- obs4MIPs test site
- CMIP6 publication within Data Challenges

Installation

- ESGF 2.6.5-7 releases for Data Challenge needs and security updates [March 2018]
- Converted IDP node component installation to Python [February 2018]
- Release ESGF 3.0-alpha to developer community [March 2018]

Sever-side computing

- Fix bugs for CWT WPS server/End-user API [January-March]
- Migrated aims2 from Docker Swarm to Kubernetes [January]
- Planned 3.0.0 version of WPS server [January]
- Planned 2.1.0 version of WPS server and End-user API [January]
- Planned/Released 2.0.1 version of WPS server [January]
- Created pull request for CWT CoG [February]
- Created/tested/deployed Kubernetes version of CWT WPS server [FebruaryMarch]
- Began 2.1.0 version of End-user API [February-March]
- Deploy automated testing for CWT WPS server [February-March]
- Updated documentation for CWT WPS server [March]
- Updated testing for CWT WPS server [March]
- Plan/Design CWT WPS federated operation (Security, Data management, Cache sharing) [March-]

Security

- Managed the ANL ESGF CA and signed multiple host and intermediate CA certificates for LLNL, NCCS, DIASJP, CMCC, NCI and others

Machine Learning

- Troubleshooting of FC-LSTM model ; determined inadequate for cyclone tracking workload
- Changed model to ConvLSTM (in progress)

ESGF Services

- Develop testing framework for tracking feedback
- Remove node manager from ESGF

CDAT

- Released CDAT 8.0
- Multiple meeting with Zeshawn to try to get xarray/dask started in cdms
- Logos draft for cdms and vcs
- Fixed vcs markers so that size does NOT depend of plot coordinates.
- Re-enabled log axes in VCS plots.
- Notebooks Tutorials: 1D plots, Text objects
- Help EZPlot to make it into release
- Helped user with question about CDAT/VCS/CDSM and install.
- Cleaned up VCS documentation, added nightly doc to website.
- Help spot and figure out cdms issue using esmf regridding with missing values.
- Fixed DV3D. o Hardened the whole testing side of things with help of Lina.
- Finalize CDMS documentations in Read-the-docs. o <http://cdms.readthedocs.io/en/cdmsdocsmerge/>
- Add possibility to use a "map" file in CDMS ESMF regridding.
- Tested with E3SM and compared with NCO regridding with no difference.
- Fix NPY_STRING for ESMF regridding. • Fix mask comparison in CDMS using latest Numpy 1.14

VCDAT

- Released vCDAT to the public as a Beta release.
- Added a feature to clear cells in vCDAT
- Added a feature to edit and delete loaded variables in vCDAT
- Added a component to allow inspecting and edit the values in a cell in vCDAT
- Updated Developer documentation, including the addition of a set of coding standards.

- Reintroduced project boards to streamline the development workflow.
- Created basic user documentation to cover installation, updating, uninstalling, and troubleshooting.
- Released vCDAT as a closed alpha to a select few users for testing.
- Gathered feedback from Alpha users and fixed various platform and environment specific bugs.
- Implemented a variety of fixes and features to improve the user experience per Alpha release feedback.
- Added a user notification system in vCDAT.
- Refactored several components to improve code clarity and reusability as well as to update their user interfaces.
- Added a help button that triggers an interactive tour of the application's basic features.
- Testing and feedback to James
- Created script to test if a new version (in the conda channels used to install VCDAT) is available.
- Installed alpha versions for some users.

Hardware/Infrastructure

- Installed NESSUS agent on all GreenNet servers
- Renewed SSL certificates for all Web servers
- Set up 2 virtual machines for ESGF testing
- Worked on FY18 expansion plan for CSS03 file system
- Updated OS on DTN nodes to latest version

AIMS Publications –

- Zhang et al., CAUSES: Diagnosis of the Summertime Warm Bias in CMIP5 Climate Models at the ARM Southern Great Plains Site, Journal of Geophysical Research, Atmosphere, <https://doi.org/10.1002/2017JD027200>

Conferences –

Next Quarter –

Management

- Planning ESGF F2F for 2019

Installation

- Cut ESGF 3.0-beta [June 2018]

Sever-side computing

- New end-user API for CWT WPS
- Release new version of CWT (June)
- Automated testing for CWT WPS server
- Updated documentation for CWT WPS server
- Design CWT WPS federated operation

Machine Learning

- Convert current model to ConvLSTM based model to capture spatial pattern in image
- Targeting to full paper publication for CIKM 2018 (Deadline is May 22 2018)

CDAT

- Move install to Conda-Forge

- Bug fixes
- Continues testing
- Further updating documentation
- Documentation on CDMS changes with ESMF conservative regridding bug fix and changes to algorithm
- Release VCDAT v1.0
- new release of CDMS, cdtime, genutil, cdutil , vcs
- Jupyter notebooks for animation and native mesh plotting form E3SM

Publications

- Continue testing automated data replication with DKRZ, CEDA, and IPSL
- CMIP6 publication

Hardware

- Work on the critical CPU security bugs (Meltdown & Spectre)
- Finalize FY18 CSS03 file system expansion plan
- Procure CSS03 expansion equipment
- Repurpose CSS01 and 02 file systems for scratch

ESGF Services

- Node manager becomes registry service
- Map search result difference to user db and integrate with feedback
- Publication simulator for testing of tracking