Dean N. Williams
Supported Projects: AIMS, ESGF, UV-CDAT, E3SM, PCMDI, DREAM, WIP, and ASCAC

Quarterly Report for July 1, 2017 – September 30, 2017

Quarter Accomplishments:

AIMS:

- As a member of the *LLNL Climate Leadership* team, I led and participated in daily, weekly, monthly, and quarterly meetings and teleconferences for the following LLNL funded projects and activities: ESGF, E3SM, CDAT, PCMDI, DREAM, WIP and ASCAC. **[July – September 2017]**
- As a committee member of the *DOE Advanced Scientific Computing Advisory Committee (ASCAC),* I participated in the 2017 Fall Meeting held in Washington D.C. (September 26). The meeting reviewed the DOE's Computational needs and covered views on science and science technologies, as they relate to the Office of Advanced Scientific Computing Research (ASCR); the 40th anniversary Computational Sciences Graduate Fellowship (CSGF); Small Business Innovation Research (SBIR); machine learning; and DOE supported computing technologies that made a difference (e.g., TCP/IP, MPI, AMR, ADIOS, PETsc, FatBit, HPSS. **[September 2017]**
- As the lead for AIMS, I resumed the search for a visualization post-doc for fiscal year 2018. This is due to the fact that the DREAM project received favorable funding from the DOE Office of Advance Scientific Computing Research (ASCR). Also, due to the fact that we lost two AIMS software engineers (i.e., Sam Fries and Jim McEnerney), I hired James Crean, a former AIMS intern. With Denis Nadeau going 50% time on the ESGF project, I am working to hire Ed Brown, a former intern from CSU, Chico to work on ESGF and PCMDI efforts. **[June 2017 - September]**
- Led the Computation 101 Technical Briefing on the Analytics and Informatics Management Systems (AIMS) collection of projects. The presentation was divided into four sections, where I presented the overall view and synergy of the many AIMS projects and goals, Renata McCoy presented the overview of the Energy Exascale Earth System Model (E3SM) project, Sasha Ames presented the overview of the Earth System Grid Federation project, and Charles Doutriaux presented the overview of the Community Data Analysis Tools (CDAT). The presentation was recorded by LLNL's TID team. The presentations can be seen on the AIMS website (http://aims.llnl.gov). **[August 2017]**
- Working with Tom Landry, from Canada, AIMS and its many projects connected with the International Coalition of Scientific Gateways (ICSG). This group provides a large registry of science services and platforms to the Canadian science community. We are mainly looking to integrate ESGF and CDAT into the strategic plans of this coalition. **[August – September 2017]**
- Provided input and assisted in the writing of all AIMS performance appraisals for group members (and some none group members). This also includes sitting on the CASC GL selection committee. **[August – September]**
- As the LLNL Climate Software lead, I initiated the efforts to produce a steady stream of animations / movies for the LLNL Climate Group. From our monthly discussions, we have five candidate climate animations to start the movie productions.

ESGF and CDAT:

- As the Chair of ESGF, I led and participated in weekly and monthly teleconferences pertaining to ESGF development and operations. This includes leading written proposals, reports, articles, letters of support and collaborations posters, and presentations to various organizations. **[July – September 2017]**
- Last quarter, I led the defense of the ESGF Scientific Focus Area (SFA) Triennial proposal and in person two-day review. **[July 2017]** The presentations can be found online at: https://esgf.llnl.gov/2017-DOE-Review.html. This quarter we were informed that our proposal has been **accepted!**
  - o The statement from the DOE BER headquarters, "After careful consideration of the review panel comments, subsequent internal discussions with other DOE program managers, and recognizing that the concerns are relatively minor, BER has decided to **accept** the proposed work plan."
  - o This continues ESGF and some CDAT funding for an additional three-years (i.e., FY18, FY19, FY20) at $4M/year. **[July 2017]**
- As the Principal Investigator and Chair of ESGF, I addressed the six questions from the SFA panel reviewers in a formal 14-page document. The questions cover ESGF's contingency plan, usability, documentation, tutorials, CDAT, testing mechanisms, and metrics. **[September 2017]**
- In anticipation of the ESGF Face-to-Face (F2F) Conference, I have completed the initial ESGF F2F Conference agenda and have collected all the ESGF abstracts covering the designated plenary talks, posters, and demonstrations. The draft agenda and abstracts can be found online at the ESGF public facing website: http://esgf.llnl.gov.
- The ESGF Installation Working Team (ESGF–IWT) released the new ESGF software stack (v2.5.x) in preparation for CMIP6 data. This release includes the minimum set of software requirements for CMIP6 and E3SM. It also includes the start of the integration of better petabyte storage and server-side hardware for data reduction. **[July - September 2017]**
- In July, I wrote the 2017 R&D 100 Awards entry for the ESGF. In August, I was informed that ESGF is a finalist at this year's R&D 100 Awards! The year's R&D 100 winners will be announced at the annual black-tie awards dinner on November 17th at the Walt Disney World Swan Resort in Orlando, Florida.
- CDAT team released version 2.12.x to the community in **September**. The release corrected many aspects of the visualization output and performance enhancements, especially for E3SM diagnostics efforts. The technical team is led by AIMS' Deputy Charles Doutriaux who is also responsible for the Anaconda release. Other features include new visualization output such as streamlines (led by Kitware) and scientific Jupyter Notebook examples (led by Ji-Woo Lee and associates).

Energy Exascale Earth System Model (E3SM):

- As the Energy Exascale Earth System Model (E3SM) Workflow co-group lead, I worked with the team leaders to complete the assignments associated with the 12-Month Roadmap for the time period January 1, 2017 through December 31, 2017. The roadmap includes components, such as the Process Flow, E3SM Workbench, E3SM Diagnostics, UV-CDAT, Data Management, Provenance (i.e., ProvEn), and Hardware. This also includes tracking and monitoring software releases pertinent to the success of E3SM and to the community at large. **[July - September 2017].**
- As the E3SM Workflow co-group lead, I also led and participated in weekly and monthly teleconferences pertaining to E3SM development and operations. This

includes leading written reports, posters, letters of collaborations, and presentations to various organizations. **[July – September 2017].**

- As a member of the E3SM council and the E3SM Workflow and Data Management Group lead, I led the input for the E3SM SFA workflow team. The input included relevant program objectives for the E3SM Workflow Group and stated clear milestones and metrics for scientific progress, including process flow and workbench, data management, analysis and visualization, diagnostics, provenance capture, infrastructure support, and user support. **[July 2017]**

DREAM:

- As the DREAM PI, I led and participated in weekly and monthly teleconferences pertaining to DREAM organization, development, and operations. This includes leading writing reports and ESGF and CDAT integration. **[July – September 2017].**
  - The team (i.e., Jason Boutte and Luca Cinquini) presented the CoG user interface integrated with server-side capabilities at the **August** NGNS PI meeting.
  - The team (i.e., Cameron Christensen from the University of Utah) presented the "ViSUS Streaming Visualization" at the **September** monthly meeting. He also showed how the streaming visualization effort will be integrated into the ESGF CoG interface.
- Established an ESGF Docker working group within the ESGF community. This group will share in the development of Docker. The group will be funded out of the DREAM and European Union (EU) Copernicus projects and include the U.S., Italian, French, and Germany countries. The ESGF Working Team will be led by the DREAM co-PI, Luca Cinquini. **[September 2017]**
- Attended the 2017 Annual Next Generation Networks for Science (NGNS) Principal Investigator's (PI) Meeting in Rockville, MD. At the meeting, I presented the current state of the DREAM project and Cameron Christensen demonstrated ESGF and the integration of ViSUS' visualization streaming.

Papers:

- The Journal "Scientific Reports" released the article, "Tropospheric Warming Over the Past Two Decades", where AIMS member **Jeff Painter** was listed as a co-author. **[July 2017]**
- **Sookyung Kim**, **Sasha Ames**, **Ji-Woo Lee**, **Chengzhu Zhang**, Aaron C. Wilson, and **Dean Williams**", have an accepted paper to Climate Informatics, entitled "Massive Scale Deep Learning for Detecting Extreme Climate Events", NCAR/TN536+PROC. **[July 2017]**
- Added input to LLNL's Science and Technology Review (S&TR) article, entitled "The Atmosphere around Climate Models". The article covers the decades of history involving LLNL's development of predictive climate models (i.e., from Chuck Leith's first ever color animated climate model in 1960 to present day E3SM efforts). The article is in rough draft form and is expected out in the spring of 2018. **[August - September 2017]**
- The article: "Use Cases of Computational Reproducibility for Scientific Workflows at Exascale" is a collaborative effort with PNNL, LLNL, and BNL laboratories. **Sterling Baldwin** is the second author on this provenance paper. **[August - September 2017]**

- **Sookyung Kim**, **Sasha Ames**, **Ji-Woo Lee**, **Chengzhu Zhang**, Aaron C. Wilson and **Dean Williams**", have an accepted paper to IEEE ICDM, entitled "Framework for Detection and Localization of Extreme Climate Event with Pixel Recursive Super Resolution". It will be presented at the Seventh Data Mining Workshop on Earth System Science. **[September 2017]**
- **Lee, J.-W.,** K. Sperber, P. Gleckler, C. Bonfils, and K. Taylor, 2017: Quantifying the Agreement Between Observed and Simulated Extratropical Modes of Interannual Variability. Climate Dynamics (in review) **[September 2017]**

Next Quarter's Roadmap

- I have two main ESGF focus for the next quarter roadmap:
    1. Prepare for the 7th Annual ESGF F2F Conference, which will be held December 4 – 8 at the Sheraton Fisherman's Hotel in San Francisco, California. The conference will focus on infrastructure to house tens of petabytes of data generated in support of projects such as the DOE Energy Exascale Earth System Model (E3SM) and the international Coupled Model Intercomparison Project, phase 6 (CMIP6). **[October – December 2017]**
    2. As the Chair of ESGF, I will focus on the setting up the ESGF priority list for DOE, NASA, EU, Australia, and Canada projects. This will be in preparation for revising the ESGF **strategic and implementation plans** for the next three years. **[October - December 2017]**
- As the Chair and PI for the ESGF, I will be attending two-day R&D 100 Conference, November 16-17, featuring an impressive lineup of educational sessions and keynote speakers. I will also be attending the annual black-tie awards dinner on November 17th, in hopes that ESGF is indeed a R&D 100 Awards winner! **[November 2017]**
- As a member of the E3SM council and co-lead for the Workflow Group, I will attend the E3SM SFA Review Meeting, to be held New Jersey. This meeting will discuss the critical elements within the overall scientific charge of the E3SM project. **[October 2017]**
- I will continue to lead ESGF, CDAT, E3SM Workflow, DREAM, and other AIMS *software development efforts and* continue to lead and/or participate in several da*ily, weekly, monthly, and quarterly meetings and teleconferences* for the ESGF, E3SM, CDAT, PCMDI, WIP, DREAM, and ASCAC projects. This include pushing for the release of provenance in the E3SM Process Flow and the release of CDAT v3.0 which includes the visual user interface (vCDAT). These must be in place for the 7th Annual ESGF F2F Conference. **[October - December 2017]**
- I will continue chair and participate on several review and *advisory committees and boards.* **[October - December 2017]**
- I will continue to push the *AIMS Leadership* team to develop *strategic and implementation roadmaps with JIRA tasks* for stakeholders and the community at large. **[October - December 2017]**
- Continue climate animations / movies by prioritizing the animations accordingly and helping Timo and the visualization team with what they may need in order for success animation productions. These animations will be used for recruiting efforts and promotion of LLNL's climate and visualization capabilities. For example, Ben Santer is looking to use the visualizations at his Bankhead theater presentation in March f2018 to help explain climate detection. **[October - December 2017]**

Resources Required to Achieve Goals

- Purchased four petabytes of disk storage for the upcoming CMIP6 archive. The hardware should be in place in early FY18.
- For server-side data reduction (i.e., subsetting, averaging, and regridding), a compute cluster will be needed for the CMIP6 community. Sometime in FY18, we will need to investigate the size of the needed compute cluster and purchase it for anticipated CMIP6 use.
- Work on recruiting Ed Brown to my group as a replacement for departing AIMS members.