ARISTOTLE
CLOUD FEDERATION

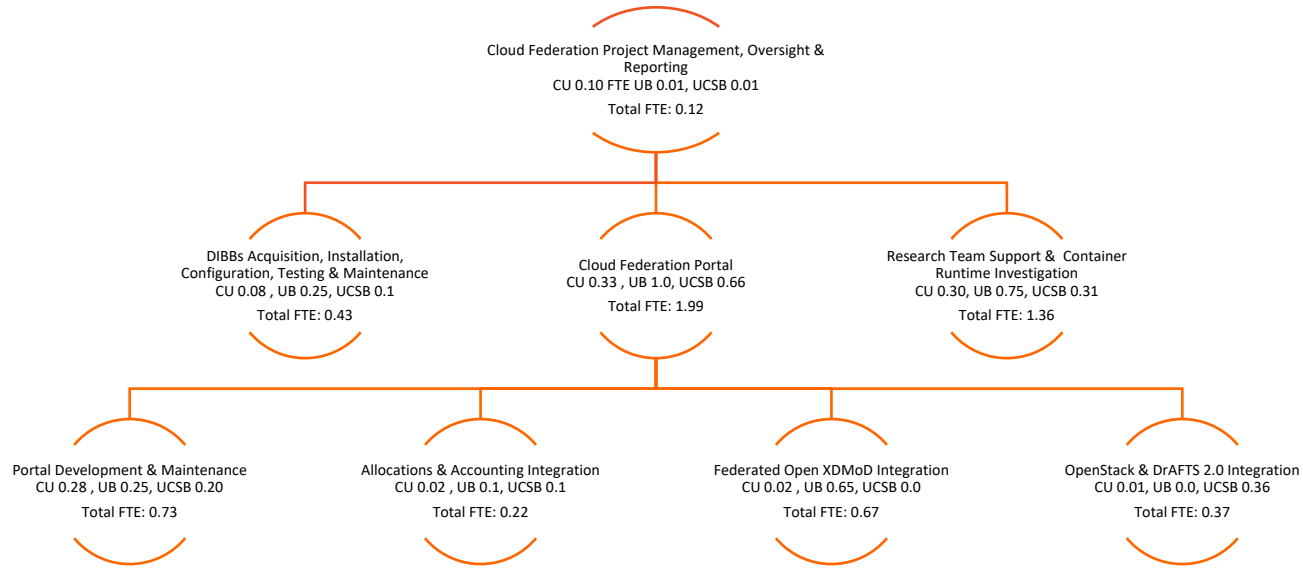
CC*DNI DIBBs: Data Analysis and Management Building Blocks for Multi-Campus Cyberinfrastructure through Cloud Federation

Program Year 6: Quarterly Report 2
(No-Cost Extension Year Ending 9/30/2021)

Submitted by David Lifka (PI)
lifka@cornell.edu

3/29/2021

This is the Program Year 6: Quarterly Report 2 (PY6 Q2) of the Aristotle Cloud Federation team. This report is part of a one-year No-Cost Extension. We report on plans and activities for each area of the project Work Breakdown Structure (WBS).



#1541215

Contents

#1541215

1.0 Cloud Federation Project Management, Oversight & Reporting

1.1 Subcontracts
All subcontracts are in place. Nothing new to report.

1.2 Project Change Request
No new project change requests.

1.3 Project Execution Plan
The Project Execution Plan (PEP) was approved by NSF on 12/18/2015. We are operating as planned.

1.4 PI/Partner Meetings
- The Aristotle project has added a new partner—VEXXHOST—to help us perform a one-time deployment, migration, and support of a new OpenStack private cloud at the University at Buffalo (UB) and to help migrate current workloads to it. UB plans to move to a non-vendor specific version of OpenStack so that they are in a better position to support it on their own without vendor lock-in. This approach will also provide services to provision Kubernetes and increase the use of open source tools. A private cloud is available from VEXXHOST for additional flexibility. Staffing shortages and Covid-19 restrictions have made it increasingly difficult to bring in new staff, particularly someone with this skill set. UB's experience with this cloud stack will be shared with Cornell, UCSB, and other interested campuses. See https://vexxhost.com/.

1.5 Project Planning Meetings

12/8/2020 status call:
- Our goal is to run a WRF science application on multiple-platforms to document the effort required to run on XSEDE and cloud platforms using a variety of deployment methods. The potential WRF and/or benchmarking runs are:
  - Stampede2 with a Slurm script
  - Stampede2 with a Slurm script and Singularity
  - Google Cloud with Terraform/Kubernetes in a Docker container
  - Google Cloud or Red Cloud with X-Containers
  - Red Cloud Ansible/Virtual Cluster scripts in a Singularity container
- Knepper is setting up a meeting with atmospheric scientist and wind turbine expert Sara Pryor to determine which WRF application to use for our multi-platform runs.
- In preparation, Vaillancourt is working with Myers on new Google Cloud deployment experiments with Terraform/Kubernetes and is testing basic WRF runs on Stampede2.

1/5/2021 status call:
- Wineholt is working on the Singularity and WRF containers and run scripts. Scripts for the bare metal run and Singularity run will need to be similar. Vaillancourt will do the test runs.

1/8/2021 status call:
- Myers is developing the K8s (Kubernetes) cluster Slurm script. The configuration steps are taking a long time; we will resolve this issue with assistance from Lee. The rest of the build went fine.

#1541215

- Trumbore has a meeting with the Exotanium team (http://exotanium.io/) to discuss the X-Container testing plan.
- Sprouse is completing the documentation for running an elastic Slurm cluster in a Red Cloud image (https://github.com/CornellCAC/CRI_CAC_Red_Cloud_Slurm_Cluster). Knepper would like to understand the constraints and performance you can get with our virtualized cluster on Red Cloud. That information would be very interesting to our cloud users. For example, Matt Marx, would like to use virtual clusters with Slurm to search large-scale, patent data sets to reduce barriers to the commercialization of science and technology (https://zenodo.org/record/4235193#.YE-pRZ1KiUm).

1/19/2021 status call:

- Sara Pryor identified the WRF application she'd like us to test. It's very sensitive to the compiler and to which optimizations are initialized. We want to regenerate the science in multiple places. "Retargetability" with ease of use is the goal. While some users want to use the highest performance machine no matter what and are willing to wait for it, other users want the answer in the shortest turnaround time and don't care where it runs as long as it gets the same answer.

2/2/2021 status call:

- We met with Tristan Shepherd from Pryor's team, built a container, and compiled WRF in Red Cloud for the offshore wind turbine use case. We're currently working through an error. There will be a public WRF container and a private container which will not be shared because it contains custom physics. The memory requirements are heavy. We may explore rootless Docker work in multi-node. The main goal, however, is to investigate whether real science can be done using XSEDE and public clouds interchangeably in an automated way with the right answer and right scaling.
- The next steps are science code runs with Singularity on Stampede2, Docker/K8s runs on a public cloud (likely Google), and bare metal runs on Stampede2.

2/5/2021 status call:

- Wineholt continued to work with Shepherd on the WRF use case. We plan to write a white paper on the effort required and performance of the application on multiple platforms using different deployment methods.
- Good news: the WRF science application is up and running on Red Cloud with X-Containers. We wrote up instructions on how to deploy it. We need someone on the Aristotle team to try to reproduce the steps to get it running on Red Cloud.
- Reynolds will use our Slurm script to run containers with the HPL benchmark on virtual clusters.

2/16/2021 status call:

- Pryor needs all runs completed by August 31st to meet sponsor deadlines. We are targeting April 18th for some preliminary results. Hurricane Sandy will be used as an X-Containers/Singularity functionality test to assure reproducibility.

3/2/2021 status call:

- The container team is meeting weekly with Tristan Shepherd to get the WRF containers working. We have a public Docker container available now on Docker Hub and GitHub. We're working out some minor build and run issues with the science code; we can run the WRF job with a certain physics parameter turned off but the researchers want it turned on. WRF is sensitive to where it's

#1541215

compiled. How many cores you have to run WRF on depends on where you compiled WRF. We recommend compiling where you are running. We believe we have a plan that will set up our container with a reproducible build and it will be even easier to recompile. We will include this capability in our documentation. We'll use the Docker containers to run on Stampede2.

3/5/2021 status call

- Due to version updates, some pieces of our Terraform with Kubernetes platform that we created during the summer are not working now. Once we update the versions, we'll be back in sync.
- We'll be doing Docker container runs using the X-Container runtime. The next step is to run a project container on multiple nodes and compare outputs.
- Sprouse ran an HPL multi-node job with MPI on a cluster and will likely scale it to 4 nodes x 28 cores with HPL data next. We want users to be able to stand up their own virtual clusters quickly. Weill Cornell Medicine-Qatar are among the many research groups with a strong interest in this capability. We are flushing out the documentation for this and are discussing building a virtual cluster in the cloud service by this fall rather than offering only a build-your-own capability. We would tie that service directly into our accounting system to account for cloud usage. This would involve standing up a multi-user virtual cluster head node.
- Wineholt is continuing to work with Shepherd on a public and private Docker WRF image and is doing test runs on Google Cloud and Stampede2. He is currently running on 8 cores or less and needs to correct a compilation error to run on more than 8.
- The virtual cluster with Slurm is running in Red Cloud and we are working on completing scripts and Ansible playbooks for creating a virtual cluster in Red Cloud's OpenStack environment.

3/9/2021 status call:

- UB is in the middle of revamping their OpenStack cloud and will be using open source Ceph. The new cloud will be built while their production cloud remains in operation. They will have a new OpenStack installation when it is completed which will be way more self-supporting than relying on Red Hat support whose ongoing costs appear high. The build will be Ubuntu. They will be using some funds to update SSD drives in their Ceph cluster as well.
- VEXXHOST is helping with their move to open source. They are showing UB their installation process as they go so the Aristotle team will have good knowledge transfer. A lot of AWS start-ups use Ubuntu; very few use CentOS. UCSB said that they have experienced network management challenges with Ubuntu. They believe it is fine for getting instances up. Users love it because if you're downloading an image it will have the latest of everything. However, if you're managing a site with Ubuntu, it's a bit more challenging, but getting better. UB's running Panasas for parallel storage.
- UCSB asked NSF to repurpose $50K in hardware funds to purchase hardware (racks, switches, power) for the science team building infrastructure for the Citrus Under Protective Screening (CUPS) project and the remaining funds would be spent on summer salaries. The switch was requested because the person working on the project got accepted into their PhD program and they have not been able to replace her. UCSB will have complete access to the CUPS infrastructure while they replace the screen.

3/16/2021

- We believe that the flexibility to run on multiple resources can reduce time to science. Many in the HPC community do not believe public clouds are useful for real science. We are investigating whether a complicated science code can get utility out of both XSEDE and public clouds. Public

#1541215

clouds and campus clouds may not do it any faster, but may be useful if they reduce turnaround time. The price/performance ratio and reduction in turnaround time are important areas of investigation. Once you remove the barrier to entry to public cloud in a large-scale science effort, there are use cases where it makes sense to use public cloud. 95% of batch queues use 16 nodes or less (lots small jobs) and then users run their one big job. Running small jobs on the cloud and big runs on national resources like Stampede2 may make sense. TACC would be happy because they would have more big runs and public/campus cloud providers would be happy because they're being used in a small, light scale way. This is our hypothesis. We believe it may be true but no one has been able to show it definitively. We will run WRF on a public cloud and Stampede2 to investigate. No one has developed a fully automated way to run back and forth between XSEDE and public clouds. Our longer-term goal is to develop a tool to do that and offer it as a service. We want to bring ease of use to the party and facilitate immediate turnaround of lower performance, small scale jobs that you know can then move into a large-scale production machine for the big run.

2.0 Container Runtime Investigation Summary

The Cornell Aristotle team has built initial containers with Docker, Singularity, and X-Containers including basic benchmark software and has started to build WRF containers for testing science cases. The next step is to build multi-node container examples that can run in parallel for larger problems. Testing has begun on the Stampede2 system as well as on Red Cloud, utilizing a virtual cluster toolkit developed by XSEDE and improved by the Cornell Aristotle team. Collaboration with the Sara C. Pryor research team on WRF specifics has developed science-supporting test cases that will inform the analysis of each of the container runtime systems.

3.0 DIBBs Acquisition, Installation, Configuration, Testing & Maintenance Report

3.1 Hardware Acquisition

- Cornell, UB, and UCSB had no hardware acquisitions this quarter.
- Cornell is busy assessing what network equipment is required for expanding the existing Red Cloud and installing a new Red Cloud in parallel.

3.2 Installation, Configuration, and Testing

- UB started the installation of their new OpenStack cloud with the VEXXHOST team. VEXXHOST set up an Ubuntu 20.02 MaaS provisioning environment and enrolled all of the equipment UB had purchased for the Red Hat Director install. UB is working on freeing up a few Ceph OSDs and GPU hypervisors to be used in the new deployment. Once the OSDs are ready, VEXXHOST will enroll them in MaaS and then provision the Kubernetes Orchestration layer. They will then deploy OpenStack and test to see that the features UB requested are working. From there, UB will migrate the workloads from the current OpenStack to the new deployment. As instances are migrated, the hardware will move to the new deployment. Eventually all the instances will be migrated and the remaining resources, including the other half of the Ceph OSDs, will be added to the new deployment. UB is early in this process, but they are really impressed with VEXXHOST and are confident this will be a smooth deployment and transition.

#1541215

3.3 Federated Identity Management

Researchers use single sign-on at any member site.

3.4 Cloud Status by Site

The chart below shows each site's production cloud status.

	Cornell	Buffalo	UCSB
Cloud URL	https://redcloud.cac.cornell.edu	https://lakeeffect.ccr.buffalo.edu/ (access only to federation)	https://openstack.aristotle.ucsb.edu/
Status	Production	Production	Production
Software Stack	OpenStack	OpenStack	OpenStack
Hardware Vendors	Dell	Dell, Ace	Dell, HPE, DXC
DIBBs Purchased Cores	*616	**792	***740
RAM/Core	8GB	up to 8GB	9GB Dell, 10GB HPE
Storage	Ceph (1.6PB)	Ceph (768TB)	Ceph (720TB)
10gb Interconnect	Yes	Yes	Yes
Largest Instance Type	28core/192GB RAM	24core/192GB RAM	48core/119GB RAM
Globus File Transfer/End Points	Yes/CAC Home Directories	Yes/CCR Home and Projects Directories	Yes/POSIX UCSB Aristotle
Globus OAuth 2.0	Yes	Yes	Yes
Total Cores (DIBBs purchased cores + existing cores) = 2776	* 616 additional cores augmenting the existing Red Cloud (1316 total cores).	** 792 total cores (UB Lake Effect Cloud and CCR cloud will be one pool after upgrade).	***740 cores in UCSB Aristotle cloud (956 total cores, Aristotle is separate from UCSB campus cloud)

4.0 Cloud Federation Portal Report

Content updates to the project portal are ongoing (https://federatedcloud.org).

4.1 Software Requirements & Portal Platform

No software changes were made to the portal platform this quarter.

4.2 Integrating DrAFTS into the Portal

The *Aristotle AWS Pricing Tool*, based on UCSB's original DrAFTS spot market prediction technology, is now available. It helps users compare Aristotle resources to the various AWS alternatives based on performance, cost, and price-performance. Visit https://federatedcloud.org/using/drafts.php to learn more or go directly to the *Aristotle AWS Pricing Tool* at http://169.231.235.92:5000/.

#1541215

4.3 Integrating Open XDMoD into the Portal

4.3.1 Application Kernels (AK) Containerization in the Cloud

AK containers are used in all Aristotle OpenStack instances within the XDMoD performance monitoring module.

4.3.2 XDMoD Cloud Integration

All 3 sites are running Federated Open XDMoD 9.0.
.

4.4 Allocations & Accounting

The database schema is available to the broader community via GitHub.

5.0 Research Team Support

5.1 Science Use Case Team Updates

Use Case 1: A Cloud-Based Framework for Visualization & Analysis of Big Geospatial Data

The OUTSTEP Community Platform (https://outsteps.org/) is hosted on Aristotle and now has over 100 members including partnering universities, government agencies, and non-profit organizations. The supported network consists of the leaders listed here: https://outsteps.org/Leaders/. Working with Professor Alfonso Mejia of Penn State, we are currently studying the network interactions between the different members using the data collected on the digital platform. An NSF proposal focusing on developing a community for the Lower Great Lakes sustainability research was recently submitted; it makes use of our initial analyses. We are currently working on a paper that describes the digital platform and its capabilities.

Use Case 2: Global Market Efficiency Impact

Dominik Roesch and University of Utah collaborators Jonathan Brogaard and Matthew Ringgenberg used Aristotle and the financial framework to investigate whether human traders still matter at a time in which trading is dominated by computers. For that, they exploited the exogenous closure of the NYSE Floor due to Covid-19. In January, they finished a new version of their paper "Does Floor Trading Matter" (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3609007) that will be presented at the *8th Annual Conference on Financial Market Regulation*. Other projects include a LASSO regression predicting short-term stock returns using the whole cross-section of international stocks and an investigation of CumEx trading (called the "Biggest Tax Heist" in history by the *NY Times*).

Use Case 3: Application of the Weather Research and Forecasting (WRF) Model for Climate-Relevant Simulations on the Cloud

New Sara C. Pryor team simulations: Currently our effort is focused on implementing a multi-node WRF simulation capability. Partial success has been achieved.

#1541215

Analysis of Prior Simulations: Most of our effort focused on analysis of our existing simulations centered on two key themes:

1) Use of machine learning approaches for wind gust detection and quantification. We have compared our new deep learning-based approach with a straightforward (physics-based) post-processing parameterization of wind gusts and statistical approaches that employ linear methods. Our analyses show ANNs (Artificial Neural Networks) exhibit higher skill than logistic and linear regression models for occurrence and magnitude, respectively. Importantly, they exhibit higher skill in forecasting strong (17 to 25.7 ms-1) and damaging (≥ 25.7 ms-1) wind gusts. Deep ANN models capture up to half of all strong and damaging wind gusts at the three airports we have considered; Newark, Boston and Chicago O'Hare. Our work in this arena benefits from the Aristotle architecture in two keys ways; (i) Availability of large RAM. The datasets are multi-dimensional and large. (ii) Availability of multi-processors. Initial network design and testing for the machine learning is being conducted in a hierarchical framework wherein we are seeking to recursively examine the skill of ANNs as a function of (a) number and nature of predictor variables including inclusion of autoregressive terms (from 1 to 16 with all variable combinations considered, i.e. several hundred model forms of variable predictor numbers from; x1&x2 to x1..xn), (b) depth of the network (i.e. number of hidden layers, from 1 to 9, where typically 1-3 is typically regarded as optimal), and (c) the precise data record used for the network training (via bootstrapping, 1000 draws). We are using an informed (or constrained) sparse matrix approach where (c) is only applied to models from (a) and (b) that exhibit statistical skill above a threshold. Initial work largely focused on optimal network settings and so the time of training and testing of the network at a single site on a single processor of 3.1 minutes (on average) has not been too problematic. However, given the amount of uncertainty space we are seeking to explore (n > 15000), running on a single core would mean over a month of processing time per individual airport location.

2) Analysis of our first set of simulations of wind farm wakes from the east-coastal offshore lease areas using ultra-high resolutions with WRF. The simulations were performed on NERSC-Cori but were ported onto Aristotle for analysis. An abstract describing preliminary analysis results was submitted and was just accepted for oral presentation at the Wind Energy Science Conference in Hannover, Germany in May 2021: Barthelmie R.J., Shepherd T.J. and Pryor S.C. (2021): Offshore wakes in the U.E. east coast lease areas. *Wind Energy Science Conference (WESC)* Hannover, Germany May 2021. Our work in this arena benefits from the Aristotle architecture in three keys ways; (i) Availability of large RAM. The WRF output is multi-dimensional and large. (ii) Availability of multi-processors for analysis speed using parallel processing. (iii) Availability of large disk volumes so all WRF output can be hosted for analysis.

Journal manuscripts in review this quarter:

- Coburn J.J. and Pryor S.C.: Do machine learning approaches offer skill improvement for short-term forecasting of wind gust occurrence and magnitude? *Weather and Forecasting*.
- Aird J.A., Barthelmie R.J., Shepherd T.J. and Pryor S.C.: WRF-simulated Low-Level Jets over Iowa: Characterization and sensitivity studies. *Wind Energy Science Discussions* https://doi.org/10.5194/wes-2020-113, in review, 2020.

#1541215

Use Case 4: Transient Detection in Radio Astronomy Search Data

The Singularity container developed by Peter Vaillancourt has been used by graduate student Akshay Suresh on XSEDE's Bridges and allowed a very simple transition to Bridges-2, where the work is ongoing (described by Suresh as "seamless," which is often *not* the case for moving between specialized computational resources).

Use Case 5: Water Resource Management Using OpenMORDM

Bernardo Trinidade (now-graduated PhD student in the Reed Group) has provided CAC systems expert Bennett Wineholt a measurement script to run OpenMORDM with step timing to quantify overhead, which Wineholt is scheduled to investigate.

Use Case 6: Mapping Transcriptome Data to Metabolic Models of Gut Microbiota

The team focused on addressing reviewer comments for their submission to the American Society for Microbiology's *mSystems Journal*. One of the reviewers suggested that we compare our models to the Biolog dataset. This required simulating each of the models using FBA to determine growth under a new collection of media, and then comparing that to Biolog's experimental growth results. This work was completed and the paper was resubmitted. The Predicted Metabolic Function of the Gut Microbiota of *Drosophila melangogaster* by Nana Y.D. Ankrah, Brandon E. Barker, et al. was made available this quarter as a preprint at https://www.biorxiv.org/content/10.1101/2021.01.20.427455v1. We also tested the project's containerization facilities by rebuilding the VM from the project's NixOS configuration.nix. This resulted in a few updates to documentation, but otherwise, no hurdles were observed.

Use Case 7: Multi-Sourced Data Analytics to Improve Food Production & Security

Citrus Frost Prevention (Lindcove Research and Extension Center, Exeter, CA):
Implementation of a CFD model for airflow through the CUPS screenhouse structure is proceeding. Currently, a preliminary model captures air velocity and heat transfer through the porous mesh vertical wall with an open top. Science team graduate students are currently enhancing the model to account for a mesh top and an incident wind wall that is canted 45 degrees (to reflect the true geometry of the structure). Once the geometry is successfully incorporated, model development will turn to capturing vortical effects. A validation experiment is planned for the basic velocity and heat transfer model for the time period immediately following the lifting of COVID restrictions on University travel (currently estimated for early summer).

As part of the instrumentation needed to develop the predictive CFD model, the Science Team created a data visualizer for the growers who they are collaborating with on the project. Hosted in Aristotle, the prototype is undergoing user testing ahead of a production release scheduled for next month.

Other projects:
The UCSB Edible Campus Food Security project continues with at-home research by the students. Currently, the science team is developing an image analysis system that attempts to correlated windsock images with windspeed and direction measurements to provide the Edible Campus with low cost, high density airflow measurements. Data gathering at Sedgwick remains on hold pending the approval of a requested exemption from COVID restrictions that would permit the Science Team to visit the site.

#1541215

6.0 Community Outreach and Education

6.1 Community Outreach

- Michael Zhang, Chandra Krintz, and Rich Wolski shared their edge computing experiences with the research community by publishing "Edge-Adaptable Serverless Acceleration for Machine Learning IoT Applications" in the *Journal of Software: Practice and Experience:* https://sites.cs.ucsb.edu/~ckrintz/papers/ZhangSTOIC_SPE20.pdf
- Co-PI Rich Wolski and Aristotle use case scientist Chandra Krintz's work on cloud-based IoT solutions was mentioned in a *PCMag* article.
- A video on the SmartFarm Project that Aristotle is collaborating with and hosting data for is available on *YouTube*: https://www.youtube.com/watch?v=aTrvhM5erP8
- Planning continued for the deployment of the "New York zone" of Jetstream 2 (1,024 computer cores and 869TB storage) at Cornell. Installation is schedule for April 20th. It will be used to explore the federation of clouds and to make OpenStack enhancements that will be shared with the rest of the project team and disseminated to the broader research community. Cornell will draw on our Aristotle experiences to create campus software so that campuses can set up their own clouds. We recently tested and deployed Slurm-based MPI clusters on Aristotle and Jetstream. The availability of the New York zone will help to facilitate faster development and dissemination of cloud tools for research.

6.2 Education

- Aristotle REU students were featured in the Coalition for Academic Scientific Computation's *Science at Full Speed* brochure in the "Visionary Ideas for Virtual Education" section – https://www.cac.cornell.edu/about/pubs/2021CASCBrochure.pdf
- Aristotle User Guide documentation was updated; it is available to the broader community at GitHub: https://federatedcloud.org/using/gettingstarted.php.