The ERN Cryo-EM Federated Instrument Pilot Project: Phase 1

1st Maureen Dougherty
Ecosystem for Research Networking
New Brunswick, USA
maureen.dougherty@ernrp.org

2nd Jason Kaelber
Institute for Quantitative Biomedicine
Rutgers University
New Brunswick, USA
jason.kaelber@rutgers.edu

3rd James Barr von Oehsen
Office of Advanced Research Computing
Rutgers University
New Brunswick, USA
jbv9@rutgers.edu

4th Morgan Ludwig
TechSquare
Cambridge, USA
mludwig@techsquare.com

5th Kenneth Dalenberg
Office of Advanced Research Computing
Rutgers University
New Brunswick, USA
kd541@oarc.rutgers.edu

6th Jeremy Schafer
Office of Advanced Research Computing
Rutgers University
New Brunswick, USA
jschafer@rutgers.edu

7th Michael Zink
Electrical and Computer Engineering
University of Massachusetts, Amherst
Amherst, USA
zink@ecs.umass.edu

8th Boyd Wilson
Omnibond
Clemson, USA
boyd@omnibond.com

9th John Goodhue
Massachusetts Green High Performance Computing Center
Holyoke, USA
jtgoodhue@mghpcc.org

10th Wolf Hey
Institute for Computational and Data Sciences
Pennsylvania State University
State College, USA
wzh5214@psu.edu

11th D. Balamurugan
Office of Advanced Research Computing
Rutgers University
New Brunswick, USA
bala.desinghu@rutgers.edu

Abstract—The Ecosystem for Research Networking (ERN) CryoEM Remote Instrument Pilot Project was launched in response to feedback and survey data collected from hundreds of participants of the ERN series of NSF (OAC-2018927) funded community outreach events revealing that Structural Biology instrument driven science is being forced to transition from self-contained islands to federated wide-area internet accessible instruments. Its goal is to facilitate multi-institutional collaboration at the interface of computing and electron microscopy through the implementation of the ERN Federated OpenCI Lab's Instrument CI Cloudlet design. The conclusion will be a web-based portal leveraging federated access to the instrument, workflows utilizing edge computing in conjunction with cloud computing and real-time monitoring for experimental parameter adjustments and decisions. The intention is to foster team science and scientific innovation, with emphasis on under-represented and under-resourced institutions, through the democratization of these scientific instruments. This paper discusses the latest Phase 1 deployment efforts.

Index Terms—Research Computing, Federation, Cloud Services, Edge Computing, Core Facilities

I. INTRODUCTION

Transmission electron microscopes used in cryoEM cost millions of dollars and are increasingly located in national or regional facilities [1]. They can collect hundreds of multi-frame images per hour, with compressed dataset sizes in the terabytes. If scientists can monitor results of their experiments in real-time [2], they can adjust parameters as an experiment goes forward or even optimize instrument utilization because they may know when their target has been achieved—or recognize and terminate a fruitless run. Currently, insufficient data transfer rates, and significant latency and HPC queue wait times do not permit on-the-fly processing in either public clouds or university clusters. Additionally, due to the high cost of acquiring and operating such instruments, only scientists at select research centers and (R1) universities typically have access to these instruments and can be trained adequately. The instruments, the computing and data infrastructure and the expertise needed to support and make effective use of them have been largely beyond the reach of faculty and researchers at many under-resourced colleges and universities.

In response to the requests from the research community gathered from workshops and survey data collected from out-

reach efforts, and the need for remote access to research instruments during Covid-19 pandemic restrictions, the Ecosystem for Research Networking (ERN) [3] has partnered with the Rutgers CryoEM & Nanoimaging Facility (RCNF) to launch the ERN CryoEM Federated Instrument Pilot Project. The project's goal is to simplify collaborative research by removing many of the barriers encountered when accessing remote research resources, with emphasis on under-represented and under-resourced institutions. Instruments in the RCNF include a Thermo Fisher Scientific (TFS) Talos Arctica transmission electron microscope (TEM), indirectly attached to the university network via a bridge system, with limited remote access through a secured VNC server, behind the university's VPN. The ERN is creating an easy to use, secure, web-based resource portal (leveraging Open OnDemand) providing remote federated authorized access to the lab's TEM; real-time workflows for experiment parameter adjustments and decisions; local edge computing image pre-processing of raw data; and additional analysis on either Rutgers' private HPC cluster, Amarel, or public cloud resources. This paper discusses this ongoing project and the latest Phase 1 implementation efforts, updating an earlier published paper [4] with the inclusions of additional step 2 efforts since that publication.

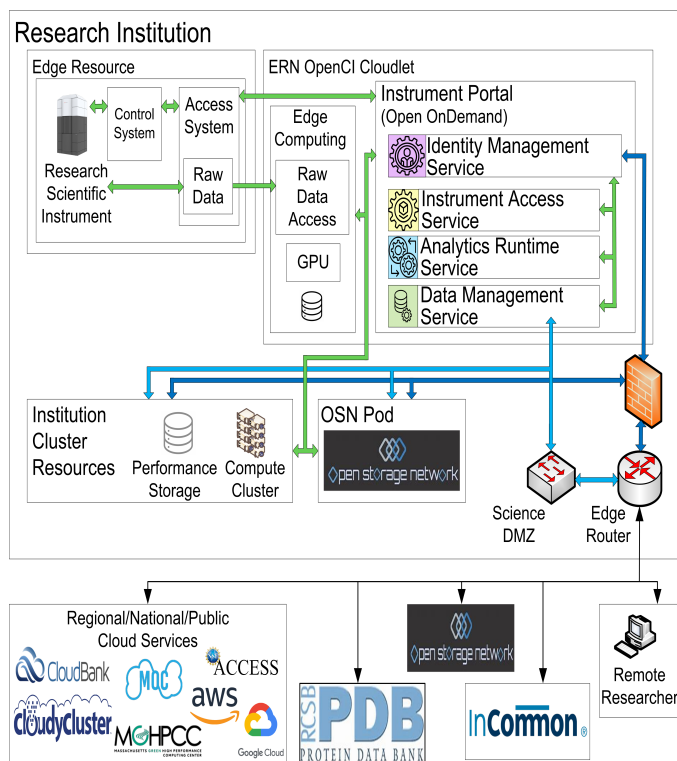## II. CRYO-EM FEDERATED INSTRUMENT PILOT DESIGN

Fig. 1. Overview of the ERN Cryo-EM Federated Instrument Pilot Design

The initial pilot design (see Fig. 1) is based on the ERN Federated OpenCI Lab (OCIL) [5] collaborative infrastructure model's Instrument CyberInfrastructure (CI) Cloudlet. The OCIL presents a common framework for federated access to a variety of research resources. The Cloudlet is designed to be the bridge device to the scientific instrument through a web-based Instrument Portal. From the portal, remote researchers will be able to login via credentials from their home institution, identify the portal's available resources, request access for particular times and/or levels of control, and utilize the instrument and resources in real time on low-latency connections, including cloud computing for analysis. Instrument administrators will register and manage their scientific instruments, review requests and either accept or decline them.

The Cloudlet represents a hybrid between instrument-local computing and cloud-remote computing whereby the cloudlet performs select, typically high-I/O operations locally through its edge computing resources, while seamlessly integrating with a remote High-Performance Computing (HPC) architecture. The Instrument Portal and its supporting services, present a secure, user friendly, web-based environment for real time analytics, reviewing, tuning, and filtering. The edge computing approach of the Cloudlet is a novel way to address a key infrastructural limitation faced by the field and is synergistic with those advances.

Developed for simple deployment, it utilizes Ansible playbooks and containers for support, maintenance, and portability. The CryoEM pilot is based on the minimum viable configuration (MVC) of the OCIL Instrument CI Cloudlet, deploying the following OCIL components: scientific instrument (edge resource), research computing (edge computing), performance storage, advanced networking, instrument access portal, identity management service, instrument access service, analytics runtime service and data management service. While this project focuses on a particular set of instruments, the design attempts to address broader research community resources. The design leverages open-sourced applications, software, tools, vendor interfaces and concepts, with the institution's existing resources and policies, to present a secure user environment.

Open OnDemand [6] with CILogon for authentication was selected as the foundation of the portal. It is a well established web portal for accessing HPC resources, built on open source software, securely hosted behind a web proxy incorporating federated authentication. It has a customizable landing page, whose default installation includes onDemand applications for file management, job management and monitoring, and command line access, aligning with several Cloudlet services. OOD is well supported with training, documentation, and workshops, and encourages the community to customize their native installation to meet the needs of their researchers. There is a strong user community providing their solutions to common needs and wants through open sourced repositories. By implementing OOD for our instrument portal landing page, we can leverage these existing resources, and contribute our efforts to this thriving community.

Podman [7], an open-source container management tool using Open Container Initiative (OCI), and Buildah [8], an open-source, command-line tool for building OCI compatible container images, both developed by Redhat, were selected for container development and management. Buildah is being

used to create OCI images from scratch, and allows several changes in one layer, as opposed to Docker's one layer per step. Podman provides full image management, modifying and maintaining Buildah provided images and containers; container life cycle management; and supports groups of containers that share resources and are managed together (pods). Both have no running manager daemon, can create rootless containers, and supports multiple container image formats. Together they provide added layers of security with rootless containers, resource isolation of containers and pods, and because there is no running manager daemon, should still be responsive as utilization increases.

Significant challenges are anticipated with identity management (authentication, access and authorization) throughout the environment, particularly with the container user environment, instrument interface, and data management throughout the architecture. Additional challenges include: how to simplify guest researcher accounts; interfacing with a variety of hardware and vendor interfaces potentially using different access methods; concerns about data and metadata ownership, access and management; financial responsibility for institution and cloud resources; local institution's security and network policies; secure container management; and making support, maintenance and portability as simple as possible. The planning and development process addressed some of these challenges, by contacting various entities for their input and feedback on design solutions, including the TEM vendor, ThermoFisher Scientific. It was determined a small, phased approach to the deployment of the design would be optimal for achieving our goals, starting with one institution and on-boarding others as we progress.

III. PILOT - PHASE 1 INSTRUMENT ACCESSIBILITY

Phase 1 of the pilot builds the basic framework for accessing and utilizing the scientific instrument through the Instrument Cloudlet for internal researchers, achieved in two steps. It establishes VNC/noVNC communications between the remote researcher and the scientific instrument. The Cloudlet is installed, with the appropriate hardware, software and configurations to provide remote access to the TEM through a noVNC browser connection. Remote researchers will launch real-time processing workflows leveraging the Cloudlet's edge computing to generate pre-processed images, and the Amarel cluster for 2D and 3D mapping and structure prediction. Step 2 will provide the Cloudlet's base Instrument Portal interface, a secure, user friendly environment for instrument accessibility (step 1) for remote researchers, in a manner that is easy to maintain and manage, leveraging open source wherever possible. Phase 1 required team discussions with vendors, systems and network infrastructure upgrades to address integration, accessibility, latency and throughput across the workflow pathway.

To determine the optimal method to access the TEM's Instrument Access System and Controls and routing requirements, the team had several discussions with TFS regarding their existing and developing toolsets. The TEM currently utilizes TFS' Remote Access Program for Interactive Diagnostics (RAPID) application via Virtual Network Computing (VNC) for communications through a proprietary Instrument Access system connected to the lab's private network. VNC is leveraged by a number of instrument vendors for communications, but has a known polling latency issue, which could have a negative impact on workflows. A review with TFS of their projects under development identified TeamViewer, part of RAPID2, as a potential option. After some investigation, this was determined not to be a viable solution due to its inability to route through the proprietary server, necessary for OOD implementation. Additional efforts with TFS would be more vendor specific, and team discussions noted a universal solution to address instrument access from multiple vendors would have a broader community impact, so we moved forward with VNC. Ultimately the VNC polling latency issue was ameliorated by polling from a noVNC instance on the Cloudlet. To address Phase 1 routing from the institutions edge to the TEM, the firewall whitelist was updated to reflect the necessary VLANS, limited IP address ranges, and appropriate service ports to provide secure access to the Cloudlet and it's access into the lab's private network.
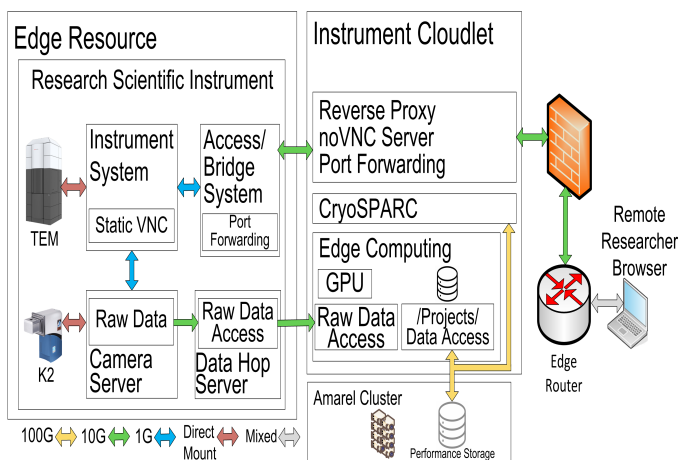

Fig. 2. Step 1 basic remote accessibility diagram

Step 1 (See Fig. 2) establishes remote access of the TEM through a noVNC browser connection, and runs a real-time workflow with the Cloudlet's edge computing. The Instrument Cloudlet was provisioned to host the web-based portal, provide edge-computing resources, and a data transfer node for further analysis either locally or at remote public and private clouds. Installed in a data center as close to the edge resource as possible, it contains dual AMD EPYC 7452 32-Core processors, 512GB memory, dual Nvidia A100 GPUs, and 2.2TB SSD storage, running Rocky Linux 8.4 with Podman, Apache HTTP Server, iperf3, CryoSPARC and an institution supported instance of LDAP. Switches, interfaces and cables were procured and installed to provide 10G (and 100G after step 1 was achieved), interfaces to the Cloudlet to support live data stream processing transfer rate requirements. The configuration has been ansiblized for ease of management and maintenance.

The TEM's camera output raw data file system can not be directly mounted onto the Cloudlet without additional network hardware, so an existing data hop server (DHS) was used to cross mount the file system to the Cloudlet. Using 10G connections and mounting read-only Common Internet File System (CIFS), the file system is first mounted on the data hop server, then onto the Cloudlet. Multiple hops is inefficient, introducing additional latency impacting performance. At a future stage of the project the DHS will be retired and the raw data folder will directly mount to the Cloudlet, eliminating the performance hit. A direct mount of the Rutgers' Amarel cluster projects directories on the Cloudlet would also achieve optimal performance for the workflow. This General Parallel File System (GPFS) requires licensing and network adjustments for a natively mounted directory, not available at this time. SSH File System (SSHFS) provides a low-performance but generalizable solution for directory access. Normally the performance of SSHFS is insufficient to support real-time analysis of CryoEM datasets, however, because the most data-intensive operations are performed on the Cloudlet, the required network I/O to and from the remote cluster is reduced by 1-2 orders of magnitude, so SSHFS is viable for this pilot study.

In November 2021, from an off-institution remote system, a Rutgers' VPN session was established, and from the remote system's local browser, a noVNC address of the Access/Bridge System submitted. Network routing directed the connection to the Cloudlet's noVNC server and port forwarding passed the connection to the Access/Bridge System, which port forwarded to the Instrument System's static VNC process. After a successful login, a workflow was launched implementing the Cloudlet's local CryoSPARC application. Through CryoSPARC, the cross mounted raw data directory was monitored for data. When detected, datasets were run through pre-process imaging analysis with the edge computing GPUs, writing output to the cross mounted projects Amarel file system. Monitoring of the analysis provided the ability for real-time adjustments to the experiment. CryoSPARC then submitted jobs to Amarel based on the data sets presence in the projects directory, for 2D alignment and 3D structure refinement. We collected ~320 images per hour of a novel complex of the transmembrane protein TolC over a two-day period, processing 2.5TB. Computation output was less than three minutes behind actual data acquisition with no increase over the course of the experiment. Graphically monitoring the output and analytics, the quality of the incoming data was assessed and experimental decisions made. Bandwidth measurements confirm that the requisite data transfer rate from the instrument exceeded 1 Gbps, but the data transfer rate between the cloudlet and cluster was low enough to be accommodated even by a 1G connection. Assessment of the workflow determined that data transfer was the main bottleneck, particularly regarding the SSHFS mount. Once the data hop server is eliminated with the direct mount, this test will be rerun. Since the pre-process output was significantly smaller than the raw data, data transfers for next step analysis will take

less time than transferring the full raw data set. Step 1, with its associated network, software and hardware installations and modifications, successfully achieved its goals. Lab researchers are actively, remotely, utilizing the Cloudlet's edge computing and workflow process for their research efforts.
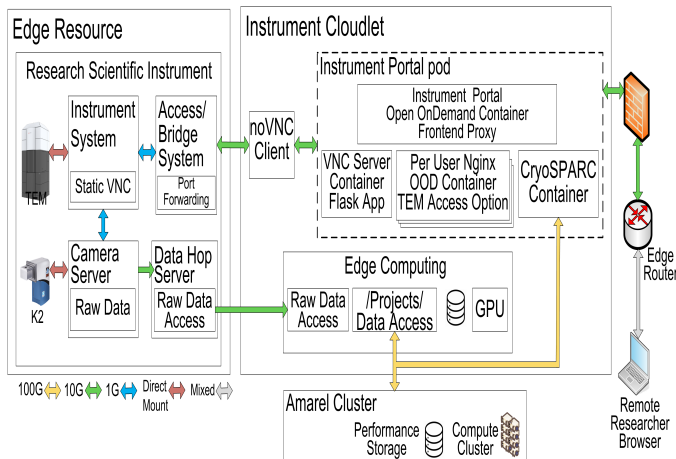
A. Step 2 - Design

Fig. 3. Step 2 design

Step 2 attempts to simplify the researcher's efforts under step 1 by providing VNC access of the TEM through the Instrument Portal by way of a navbar option and removing some of the manual steps, including the use of the instrument's Access/Bridge system's unique URL. The Step 2 design (See Fig. 3), has been updated from the initial architectural design, reflecting changes during implementation. The Instrument Portal is an ansiblized, OOD, rootless podman container (OOD front-end proxy), customized via parameterized scripts and persistent configuration files for OS independence, flexibility and portability spawning per user nginx images built dependent on authorized GIDs for new functionality like the TEM access. For SSO authentication, the OOD configuration utilizes CILogon, authenticating through the researcher's own institution's authentication process, returning user credentials that are then mapped against a local institution LDAP instance. With a successful match, the local host's username, UID:GID and /home directory are passed on for the OOD per user nginx (PUN) container image build and navbar customization by podman and a buildah stack. This presents a consistent and secure user environment, avoiding issues particularly with data access, ownership and permissions.

The rootless context container would launch for the authenticated researcher, within the same pod as the Instrument Portal container, identified as the Instrument Portal pod. The user environment presented within the container initially would include home directories as a placeholder, raw data and /projects/ directories mounted as in step 1, and applications would be available via the Cloudlet or the OOD navbar. The VNC Server container and the CryoSPARC application container are also part of the Instrument Portal pod. These will be

leveraged by the PUN to access the TEM and run the real-time experiment workflow. OOD customization is well documented, but still requires expertise to coordinate and implement these modifications, and develop the VNC server access to the TEM.
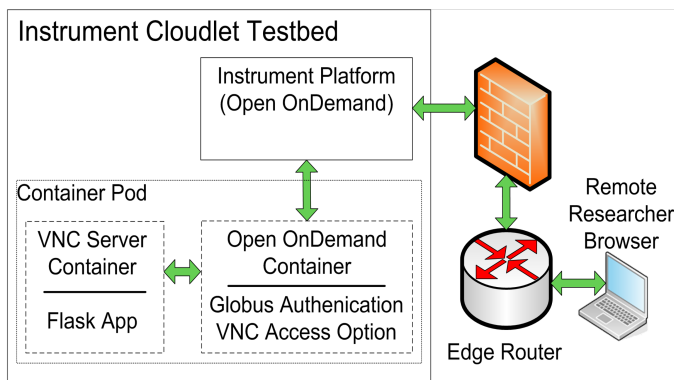
B. Step 2 - Testbed Environment

Fig. 4. Step 2 testbed implementation

The initial step 2 implementation was achieved by the team OOD expert at a remote site, without access to a TEM, using Globus for authentication, mapping against a local user file, and utilized local SSL certificates for Apache http communication between OOD and the host apache. An ansible, containerized OOD was deployed, and to simulate the TEM's static VNC server, a containerized VNC Flask application was launched, running an off-the-shelf VNC server.(See Fig. 4). Upon a successful login to the portal, an OOD image was built based on the UID:GID from the authentication user mapping, and a rootless context container was launched. The host UID is mapped with the container UID to allow consistent host/container file permissions for mounted folders, while the white-listed GID added the VNC Server option to the navbar. After successful VNC Server password submission, an HTML template was called that provides the appropriate URI which OOD then reverse proxied, redirecting to the VNC server container. The workflow ran seamlessly in the testbed environment. The initial step 2 Cloudlet deployment would utilize the testbed template for a stable base, then apply modifications to address the desired environment. Knowing that the Cloudlet implementation at other institutions may be different than our initial install, this was an opportunity to document migration steps.

C. Step 2 - Instrument Portal Implementation

Step 2 implementation on the Cloudlet required certain prerequisites: a public facing hostname and IP address, SSL certificates, a pathway to the public internet, the OIDC ClientID and Secret from the upstream OpenID Connect IDP like Globus or CILogon, and the appropriate OOD template, scripts and configuration files. In addition, a new VLAN was created to isolate the network traffic on the Cloudlet's public interface, and network routing and firewalls were updated. These efforts were achieved through collaborative efforts between the project team and Rutgers' Office of Information Technology and Office of Advanced Research Computing's networking and security teams. Once these items were addressed, the ansiblized OOD container installation was attempted, which encountered some complications.

Rootless podman containers running SELinux are by default rather restrictive, necessary for security, but cause complications with the portal implementation. Podman prevents processes from running almost all SELinux operations, runs a separate user namespace from the host, and has some default UID:GID configuration limits. SELinux by default restricts httpd network relays and network connections. Initial issues with network communications prevented the authentication process from functioning. This was resolved by minor Apache configuration tweaks and SELinux modifications to enable the network connections and httpd relays. From there updates to podman's /etc/subid and /etc/subgid were made to allow for ids greater than 65,536, resolving the insufficient number of IDs errors encountered. During troubleshooting, the inability to restart the container's Apache required the container to be restarted to implement changes, as systemd is not accessible for a non-root user. A fix for systemd access developed in the testbed will be incorporated into the production environment after verification.

Addressing the issues arising from separate user namespaces was more complicated. Since the container's user namespace is separate from the hosts, the UID:GID in the container may not necessarily map to the same UID:GID on the host, and could cause issues with the access pipeline to the instrument, data access and remote analysis. The current solution is to manually maintain OOD usermap container /etc/passwd and /etc/group files for pre-approved users, based on the host's files, to be mounted and used by the running container, thus ensuring continuity between namespaces. This is achieved through the configuration of the buildah stack launching the PUN. After confirmation that the user name provided in the returned OpenIDC email credentials from Globus or CILogon is found in the local LDAP database and matches the entry in the manually maintained usermap file, PUN is launched under the appropriate user name's UID:GID using the OOD usermap files.

Focus then shifted to the connection failure from the Instrument Portal using the VNC flask application to the Instrument System via the Bridge System. Attempts on the Cloudlet outside the portal to login via local VNC client was successful, but access through a browser were not. Upon examination, it was identified that the VNC server on the vendor's, proprietary, target system, the Instrument System, does not have the appropriate SSL httpd port available. It was determined that updates to the proprietary server was not an option, as a more general solution for use with any vendor instrument was deemed more practical, and has a broader impact. The flask application was updated to leverage OOD's reverse proxy and a noVNC client running on the Cloudlet. With these changes the remote researcher, through their web browser, is able to login to the Instrument Portal with their

institution's credentials, and login to the remote instrument.

Current efforts to provide stable, reproducible, software updates and documentation in the github repository, before running a real-time experiment workflow similar to the one in step 1, encountered a new issue with the mounting namespace in the container. Initial troubleshooting indicates this issue is with SELinux. Efforts are now focused on resolution.

A rootless container for the CryoSPARC application is also under development. The step 1 experiment workflow utilized CryoSPARC and associated scripts to monitor the incoming raw data, pre-process the image with the edge computing resource, monitor the results, and launch 2D and 3D protein structure analysis on the Amarel cluster. CryoSPARC has a master/worker and supporting database installation with a SLURM integration. Initial efforts encountered some challenges, including UID:GID user mapping and systemd issues in the workflow, as seen in the portal implementation. Investigation into solutions to these issues are underway and include the lessons learned from the the portal implementation. Concerns of where the CryoSPARC database should run, are also under discussion. The testbed will be used to evaluate solutions, eventually to be incorporated into the Cloudlet design.

IV. Conclusion

The initial phase 1 efforts of the ERN Cryo-EM Federated Instrument Pilot Project demonstrates remote access to edge scientific instruments by remote internal researchers for analytical workflows using edge computing is both feasible and beneficial to their projects. Step 1 confirmed the manual, bare bones, ability to access the scientific instrument and run an experiment workflow with real time adjustments. Step 2 will provide the ansiblized, containerized, secure, instrument portal presenting a user friendly, web-based interface for accessing the scientific resource, simplifying the the real-time experiment workflow, similar to the one run in Step 1. Implementation has been presented with various obstacles, some rather unexpected.

Launched during Covid-19, restrictions impacted experienced staff's availability for activities requiring an on-campus staff presence, staff coordination efforts between multiple departments/institutions, and procurement procedures and timelines due to stock availability and transport. These issues delayed efforts in implementation and troubleshooting. With the ease of pandemic restrictions, most of these subsided. Technical challenges encountered when implementing the project design required expertise and collaborative efforts by the team and the institution's core network and security departments to find solutions that were not simply site specific, but include resources of the broader community. Additional phases of the project will build off and expand the base foundation built here.

In phase 1, CryoSPARC was used to manage the experiment workflow. To address a broader range of applications and research domains, exploration is underway to incorporate the Pegasus Workflow Manager [9] into the Cloudlet design. This will provide a flexible and robust system for experiment workflows which includes data management and integration with private or public cloud resources. The final pilot design solution, software, container templates, ansible recipes, and documentation will be shared with the community through the github repository.

The current phase 1 efforts have been well received by the RCNF lab, who are actively making use of the currently deployed resource. Moving forward, we anticipate the ecosystem being developed through this pilot will break down resource access barriers, foster team science and the democratization of scientific instruments to benefit the research community, with emphasis on under-represented and under-resourced colleges and institutions, who typically do not have access to these resources, nor expertise to take full advantage of them.

The Github repository, https://github.com/mghpcc/ERN-Remote-Scientific-Instrument, hosts software, templates, and documentation developed through the project, including the current podman containerized OOD template. The research community is encouraged to utilize these resources, offer feedback and contribute to the development of the project.

References

[1] S. Subramaniam, "The cryo-EM revolution: fueling the next phase," *IUCrJ*, vol. 6, no. 1, pp. 1–2, Jan 2019. [Online]. Available: https://doi.org/10.1107/S2052252519000277

[2] R. F. Thompson, M. G. Iadanza, E. L. Hesketh, S. Rawson, and N. A. Ranson, "Collection, pre-processing and on-the-fly analysis of data for high-resolution, single-particle cryo-electron microscopy," *Nature Protocols*, vol. 14, no. 1, pp. 100–118, Jan 2019. [Online]. Available: https://doi.org/10.1038/s41596-018-0084-8

[3] E. for Research Networking, "The ecosystem for research networking (formerly eastern regional network)," 2021. [Online]. Available: https://ernrp.org

[4] M. Dougherty, M. Zink, J. von Oehsen, K. Dalenberg, B. Designhu, J. Kaelber, J. Schafer, J. Goodhue, W. Hey, M. Ludwig, B. Wilson, and C. McKnight, "The ERN Cryo-EM Federated Instrument Pilot Project," *ACM PEARC22: Practice and Experience in Advanced Research Computing (PEARC '22)*, vol. 52, no. 1, pp. 1–4, 2022. [Online]. Available: https://doi.org/10.1145/3491418.3535141

[5] M. Dougherty, M. Zink, and J. von Oehsen, "Identifying Research Collaboration Challenges for the Development of a Federated Infrastructure Response," *ACM PEARC21 conference (PEARC '21)*, vol. 40, no. 1, pp. 1–4, 2021. [Online]. Available: https://doi.org/10.1145/3437359.3465594

[6] D. Hudak, D. Johnson, A. Chalker, J. Nicklas, E. Franz, T. Dockendorf, and B. L. McMichael, "Open OnDemand: A web-based client portal for HPC centers," *Journal of Open Source Software*, vol. 3, p. 622, 05 2018. [Online]. Available: https://joss.theoj.org/papers/10.21105/joss.00622

[7] Podman, *Podman*, 2022. [Online]. Available: https://podman.io

[8] Containers, *Buildah*, 2022. [Online]. Available: https://buildah.io

[9] Pegasus, *Pegasus*, 2022. [Online]. Available: https://pegasus.isi.edu