## SC26 Network Research Exhibition: Demonstration Preliminary Abstract

# Network-Enabled Transcontinental CryoEM Over AmLight and FABRIC

**NRE Proposal Abstract**

AmLight-Next Frontier Team: sdn@amlight.net; NICIS: KPillay@csir.co.za; UCT: jeremy.woodward@uct.ac.za; UCSC: jweekley@ucsc.edu; CIENA: gshami@ciena.com, akara@ciena.com; FABRIC: pruth@email.unc.edu;

**Abstract**

Over the past decade, cryo-electron microscopy (Cryo-EM) has emerged as an important technique in structural biology, dramatically expanding our capacity to visualize biological macromolecules at near-atomic resolution. This advancement has been particularly impactful in biomedical research, enabling scientists to determine the detailed structures of critical targets such as the HIV vaccine research and SARS-CoV-2 spike protein. Insights obtained from Cryo-EM have directly accelerated vaccine and therapeutic development, demonstrating its potential to transform our approach to tackling infectious diseases, many of which disproportionately burden the Global South. Despite these advances, researchers in resource-constrained settings, particularly in Sub-Saharan Africa and other Global South locations, face significant barriers in accessing Cryo-EM infrastructure due to the high cost, sophisticated instrumentation, and specialized technical expertise required.

To address these access and infrastructure constraints, we extend FABRIC to South Africa, introducing a transcontinental testbed that brings advanced networking capabilities closer to under-resourced regions while exposing the practical limits of latency, data movement, and distributed coordination at intercontinental scale. We then apply FABRIC Loom AI to explore network-embedded single-particle (SPA) Cryo-EM processing through selective in-network placement of Cryo-EM data preprocessing stages, reducing reliance on centralized facilities and enabling more accessible, distributed analysis workflows.

### Goals

For SC26, the Cryo-EM collaboration will report on advances in distributed workflow development leveraging applications deployed across international research networks, including AmLight's 100G infrastructure, SANReN, and newly deployed FABRIC nodes in South Africa, establishing a transcontinental testbed that enables evaluation of distributed applications under intercontinental latency, data movement, and coordination constraints. This NRE builds on collaborative work conducted for SC25. The primary objective this time is to extend prior proof-of-concept efforts toward distributed Cryo-EM data sampling, processing, and analysis over end-to-end Data Transfer Nodes (DTNs), while sustaining high-performance network utilization in service of geographically distributed and resource-constrained research environments.

Specifically, this work seeks to move beyond prior benchmarks, including sustained 100G throughput demonstrated by the SANReN team during SC25 "flood-the-pipe" experiments, by applying these capabilities to a scientifically meaningful, data-intensive Cryo-EM use case that directly addresses barriers to access and infrastructure. In parallel, the experiment will evaluate FABRIC Loom AI as a framework for orchestrating distributed workflows and enabling selective placement of Cryo-EM preprocessing stages across geographically disparate sites, including new deployments in South Africa.

The anticipated results will demonstrate the integration of programmable network infrastructure and coordinated orchestration across the AmLight-ExP 100G link over the SACS submarine cable system, validating a transcontinental testbed for distributed, data-intensive applications under real-world intercontinental constraints. This work advances toward network-embedded SPA workflows by enabling selective placement of pre-processing stages across distributed resources. The experiment further highlights the role of South Africa's NREN within the broader ecosystem of international research and education networks participating in the Global Network Advancement Group (GNA-G), supporting next-generation, data-intensive science at intercontinental scale.

### Impacts

This project demonstrates how advanced, programmable networking can support distributed SPA workflows across geographically separated and resource-constrained environments, enabling global collaboration in biomedical research. By leveraging

FABRIC, AmLight, and interconnected high-capacity networks including Internet2, ESnet, Florida LambdaRail, and SANReN, along with submarine and terrestrial 100G links, the project supports distributed SPA workflows, allowing researchers in Africa to transfer and analyze datasets at scales that were previously inaccessible.

This next phase extends this proof of concept by introducing FABRIC Loom to explore the placement of selected SPA preprocessing steps (beam-induced motion correction and Contrast Transfer Function - CTF - estimation) within the network, alongside continued workflow execution on connected cyberinfrastructure such as the National Research Platform. While specific processing components remain under active collaboration with domain experts, this direction aims to reduce data movement, improve end-to-end efficiency, and enable adaptive orchestration of distributed resources. These efforts advance toward a sustainable, globally distributed framework that expands access to advanced imaging and supports inclusive, data-intensive scientific discovery.

**Resources**

This NRE demonstration will be conducted using resources from the collaborating domains, as listed in the Involved Parties section. Resources will involve SACS/WACS/AmLight 100G link. For integration of FABRIC resources at Ciena Booth 5035, one single-mode dark fiber pair with duplex fiber drops (LC connectors) from the SCinet NOC to Booth 5035. Ciena is providing a pedestal with a screen and network to display our presentation at Ciena's SC booth.

**Involved Parties**

**Florida International University (FIU)**: Julio Ibarra-julio@fiu.edu, Jeronimo Bezerra - jbezerra@fiu.edu, Vasilka Chergarova - vchergar@fiu.edu, Javier Franco-jafranco@fiu.edu, Juliana Hernandez-jmherna@fiu.edu, Eric Brletic- ebrletic@fiu.edu
**Information Sciences Institute (ISI), University of Southern California (USC)**: Heidi Morgan - hlmorgan@isi.edu
**South African National Research Network (SANReN), NICIS, CSIR**: Ajay Makan - ajay@sanren.ac.za, Kasandra Pillay - KPillay@csir.co.za, Johann Hugo - jhugo@sanren.ac.za
**TENET South Africa**: Paul Sullivan - paul@tenet.ac.za
**Centre for High Performance Computing (CHPC), NICIS, CSIR**: Bryan Johnston - BJohnston@csir.co.za, Werner Janse van Rensburg - WJvRensburg@csir.co.za
**University of California, Santa Cruz (UCSC)**: Jeffrey Weekley - jweekley@ucsc.edu, Vitor Hugo Balasco Serrao - vbalasco@ucsc.edu
**RNP**: Marcos Schwarz - marcos.schwarz@rnp.br, Renata Frez De Lima - renata.frez@rnp.br
**University of Cape Town (UCT)**: Jeremy Woodward - jeremy.woodward@uct.ac.za
**National Institute for Communicable Diseases of South Africa (NICD):** Danielle Martin daniellem@nicd.ac.za, Thandeka Moyo-Gwete - thandekam@nicd.ac.za
**Caltech**: Harvey Newman newman@hep.caltech.edu
**University of Chicago:** Greg Nawrocki - greg@globus.org, Lev Gorenstein - lev@globus.org
**CIENA:** Rod Wilson - rgwilson@ciena.com, Gauravdeep Shami - gshami@ciena.com, Marc Lyonnais - mlyonnai@ciena.com, Akbar Kara - akara@ciena.com,m
**Ecosystem for Research Networking:** Maureen Dougherty - maureen.c.dougherty@gmail.com
**Renaissance Computing Institute (RENCI):** Path Ruth - pruth@email.unc.edu, Mert Cevik - mcevik@renci.org