SC26 Network Research Exhibition: Demonstration Preliminary Abstract

Federated Edge AI for Distributed Cosmic Ray Sensor Networks

Carlyn Lee, California Institute of Technology HEP/Oligo Space, cblee@caltech.edu

Abstract

Cosmic rays continuously pass through Earth, carrying information about high-energy astrophysical processes such as supernovae, active galactic nuclei, and other extreme environments. Detecting scientifically meaningful events at scale is difficult. Candidate signals are rare, observations are noisy, and detector nodes may be geographically distributed across universities, schools, citizen-science participants, and edge platforms. Moving all raw data to a central site is inefficient and may not scale as detector networks grow.

We propose an SC26 Network Research Exhibition demonstrating a distributed cosmic ray sensor network that combines low-cost CosmicWatch muon detectors, smartphone-based data from the Cosmic-Ray Extremely Distributed Observatory (CREDO), real-time data streaming, and federated/edge machine learning. The demonstration will use SCInet and connected research infrastructure to demonstrate how federated/edge learning can make distributed scientific instrumentation more scalable by reducing raw-data movement while still improving a shared model across sites.

Proposed Demonstration

The demo will connect detector and data sources at multiple sites to a SC26 show-floor visualization and analysis environment. Candidate endpoints include:

1. CosmicWatch detector node at Cal Poly Pomona, Caltech, University of Delaware, or an SC26 booth. 2. CREDO smartphone image data and historical event samples.
3. National Research Platform (NRP) / Nautilus Kubernetes resources for data ingestion, model training, and dashboard services.
4. SC26 booth displays with real-time monitoring, model-update visualization, and event classification results.

This workflow will stream selected detector events and metadata into a shared analysis backend, visualize event rates and environmental correlations, and run distributed ML clients that classify or prioritize candidate events. We will use a centralized-learning workflow as a baseline to quantify the advantages of the federated / edge approach.

1. Centralized baseline: raw or slightly filtered event data is transported to central analysis service. 2. Federated/edge mode: detector-side or site-local clients train or update models locally, sharing compact model updates or embeddings rather than full raw data.

The SC26 demo will present live dashboards showing event flow, network traffic, model-update traffic, latency, bandwidth reduction, and candidate high-confidence detections.

Resources

The current prototype stack includes:

- Instrumentation: CosmicWatch Desktop Muon Detector hardware
- Citizen-science Data: CREDO smartphone cosmic ray events.
- Data services: Elasticsearch/Kibana for event storage and visualization.
- AI/ML: Pytorch models, DINO-style image embeddings for smartphone app images, BERT/GPT-style sequence models for event streams, and flower-based federated learning
- Infrastructure: Kubernetes-based deployment on NRP/Nautilus, with SCinet connectivity for show demonstration and WAN experiments.

Involved Parties

- Harvey Newman, California Institute of Technology, newman@hep.caltech.edu
- Yifan Wu, California State Polytechnic University, Pomona, yifanwu@cpp.edu
- Jacky Li, California State Polytechnic University, Pomona, jpli@cpp.edu
- Spencer Axani, University of Delaware, saxani@udel.edu