INDIS Workshop SC26 Chicago

Measured WAN Traffic Reduction from Federated Edge AI in a Distributed Cosmic-Ray Sensor Network

Capacity, flexibility, consistency, and collaboration on a citizen-science sensor network.

Carlyn Lee Caltech / Oligo Space

Harvey Newman Caltech

Yifan Wu Cal Poly Pomona

Spencer Axani University of Delaware

Companion demonstration · SC26 NRE115 · Caltech booth #5438

Scientific data networking

> *"Data networks are the circulatory system for science, discovery, society, and industry."*

01

Capacity

More data, from more places, without the transfer becoming the experiment.

02

Flexibility

New instruments, sites, and participants, without rebuilding the system each time.

03

Consistency

What arrives can be trusted, including across domains you do not control.

04

Collaboration

The people who hold the data are able to take part, and willing.

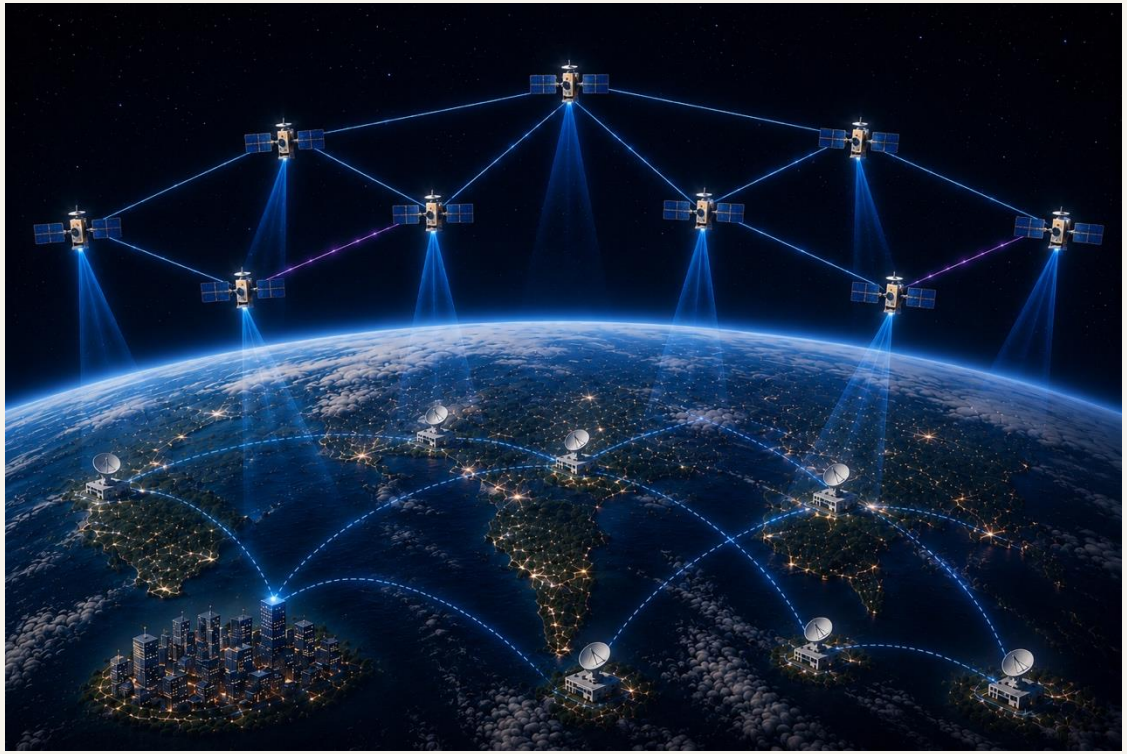
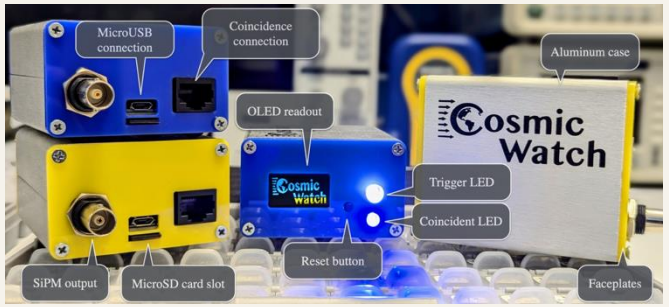


Image from xiaojian-wang.github.io Dr. Xiaojian Wang, Department of Computer Science & Engineering, CU Denver

Architecture

https://github.com/spenceraxani/CosmicWatch-Desktop-Muon-Detector-v3X



On the phone

- No install on the device.
- Covered lens
- Hot pixels masked, Clusters give three features: log brightness, cluster size, elongation.
- A 60×60 crop is kept for archive.

The model

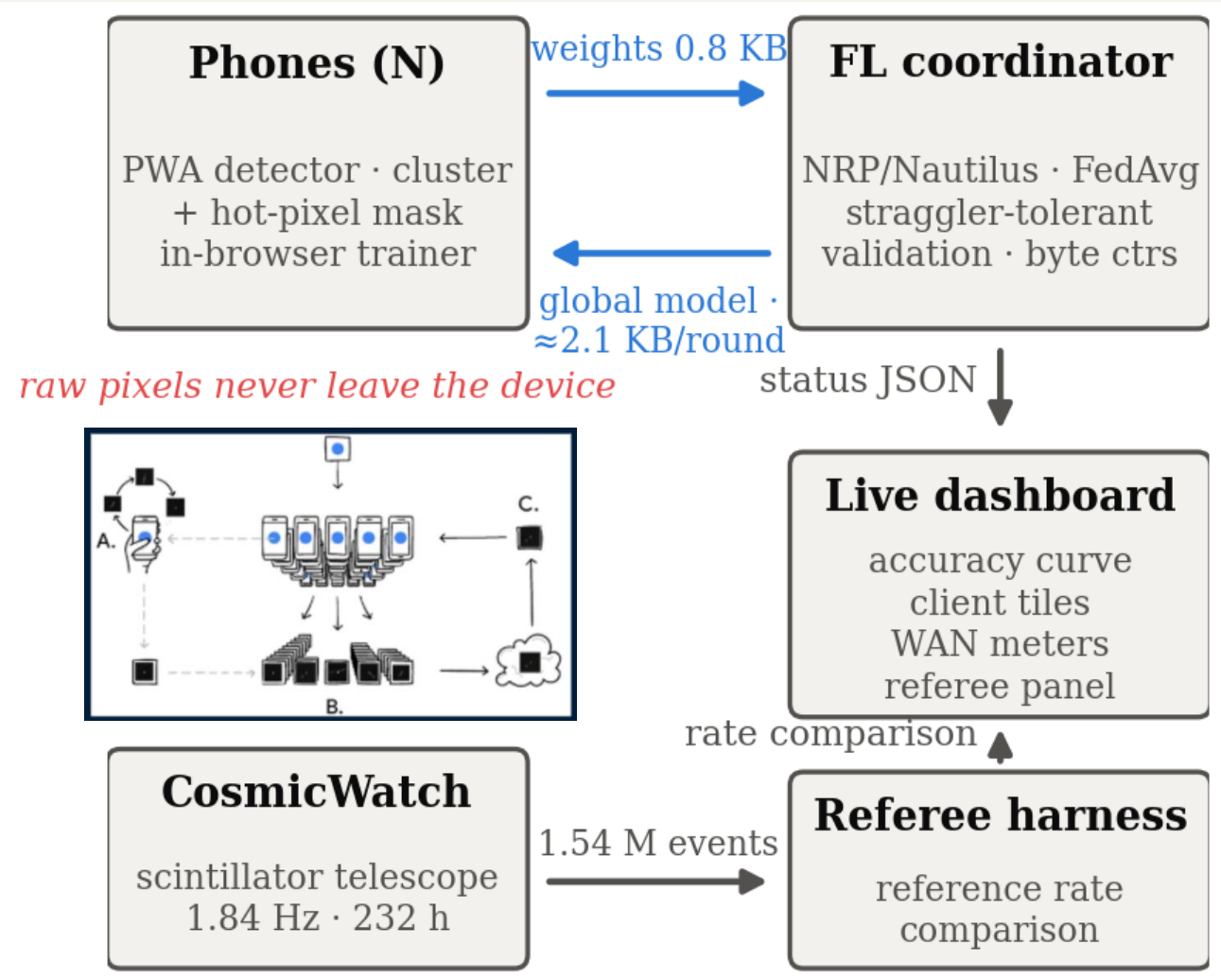
3 → 8 → 1, leaky ReLU, 41 parameters. SGD runs on the device browser and matches the server reference to ≤ 10^{-16}

Parity test

Compared with model.py (the NumPy standalone version) and the JavaScript in the progressive web app, model.js:

Stage	Largest absolute difference
Features	5.6×10^{-17}
Forward pass, initial weights	1.1×10^{-16}
Weights after each epoch	about 10^{-18}
Loss after epoch 1	5.6×10^{-16}
Forward pass after training	1.1×10^{-16}

NRP NATIONAL RESEARCH PLATFORM



THE EXPERIMENT

Phone data + Federated rounds on NRP.

For K clients holding S detections each, explore which architecture puts fewer bytes on the WAN to reach 0.95 validation accuracy

Corpus	3,131 real samples from 1,569 archived 60×60 crops. No simulated hits. 2,349 train, 782 validation. Labels are morphological.
Model	The 41-parameter network, from scratch in both modes. Same initialization seed.
Target	Stop the comparison at 0.95 on the held-out 782. Centralized training goes on to 0.999; that does not add WAN bytes.
Accounting	Those rounds were counted on the NRP coordinator: register, model out, weights back. Headers excluded. One seed, same shards.

Two procedures

Mode A · upload once

1. Draw S detections for each of K clients.
2. Charge each detection its real wire size, from the 1,569 archived documents. Mean 1,673 bytes.
3. Sum those sizes. That sum is the WAN cost. Training then runs centrally and adds no further wide-area bytes.

2,471 KB

One upload of the ten shards.

Sizes are the real documents. The measured sum is 2,471 KB.

Central training then reaches 0.999. No further WAN bytes.

Mode B · rounds until 0.95

1. The same shards stay on the clients.

Each round the NRP coordinator sends the global model (~1.0 KB). The client trains once and returns the weights (~1.1 KB).

NRP averages by sample count, scores the 782, and adds those bodies. Stop at 0.95, or at 40 rounds.

703 KB

34 rounds to 0.95, federated on the NRP coordinator.

About 2.1 KB per client per round, both ways. 34 × 10 × 2.1 KB is 714 KB. Measured total, with registration, is 703 KB.

The 1,500 detections are real phone crops. The exchange that cost 703 KB ran on NRP.

Sweep

Fleet size at fixed data: K = 5, 10, 25 with S = 150. Data per client at fixed fleet: S = 50, 150, 450 with K = 10.
The crops are real. Every federated round in this sweep ran on the NRP coordinator.

CAPACITY, RESULTS MEASURED AND COMPARED

Reduction follows data per client

1.1× **50 detections / phone**

Parity. Centralizing is competitive.

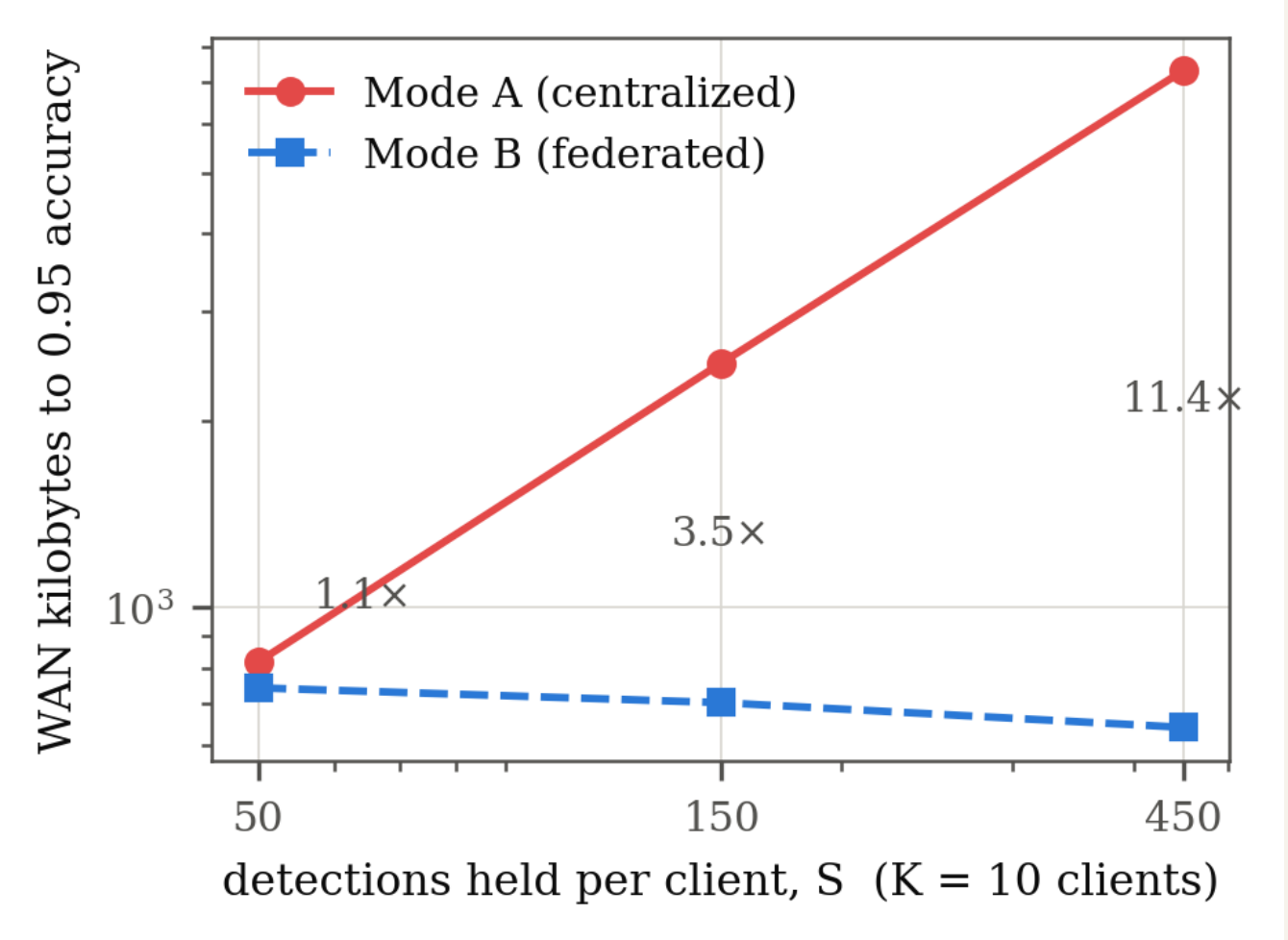
3.5× **150 detections / phone**

The operating point of the study.

11.4× **450 detections / phone**

Federated bytes stay flat as the data grows.

Fleet size barely moves the ratio: 3.2× at 5 phones, 3.9× at 25, same 150 detections. Both modes scale with K. The ratio is a property of S.



Break-even data volume

reduction ≈ (S × 1,673 bytes) / (rounds × 2.1 KB)

Upload avoided, divided by the rounds spent to reach 0.95.

Below ~50

Send the detections. On these crops, a short federation costs as much as the upload.

Above ~50

Send the weights. Further detections increase Mode A and leave Mode B flat, or slightly cheaper.

If the crop grows

Break-even moves to fewer detections per phone. Full frames would make 11× look small. 11× is the small-crop case.

Mode B at 450 detections cost 642 KB. Mode B at 50 cost 743 KB. The federated wide-area bill shrank as each phone held more data, because the run needed fewer rounds.

Once a phone holds more than about 50 detections, sending weights is the cheaper wide-area path.

Looking Ahead

- Onboard spacecraft intelligence — Process and prioritize sensor and payload data locally before transmitting to Earth.
- Federated learning across spacecraft — Share model updates across distributed satellites without continuously exchanging raw data.
- Bandwidth-aware autonomy — Adapt data transmission and processing to limited communications windows and available network capacity.
- Hardware-validated edge AI — Extend the approach toward flight-like systems through simulation and hardware-in-the-loop testing.

Questions?

SC26· Chicago
NRE115 · Caltech booth #5438