LF EDGE

Real-Time Utility Outage Detection at the Edge With EdgeLake

How UC Santa Cruz students are using EdgeLake to detect utility outages in real time and capture the electrical evidence behind them

EDGELAKE

UC SANTA CRUZ | Baskin Engineering

At a Glance

Project
EADS: Electrical Anomaly Detection System

Platform
LF Edge's EdgeLake Project

Hardware
Raspberry Pi 3 with a custom sensor HAT sampling at 7,680 Hz

Students
Bode Polhemus
Bronsen Lasala

Faculty Advisors
Dr. Patrick Mantey &
Dr. Richard Julig, UC Santa Cruz

Project Advisors
Dr. Roy Shadmon, AnyLog

Deployment
Five EADS boxes delivered to field participants

Executive Summary

When a home loses power, detecting the outage is only the first step. Understanding what happened, and where the fault originated, requires evidence from the electrical network itself.

EADS, the Electrical Anomaly Detection System, is a low-cost hardware and software solution built by UC Santa Cruz students on LF Edge's EdgeLake platform. Each EADS box samples the AC voltage waveform 128 times per 60 Hz cycle (7,680 events per second), detects anomalies locally in real time, and stores what it captures in an EdgeLake-managed local database. EdgeLake connects the boxes into one distributed system that users, applications, and AI agents query through a single API, with no cloud service anywhere in the data path.

Five units have been delivered to the field. The goal is to give utilities a clearer account of why a home went dark: sensing fast enough to capture the disturbance, distributed enough to help localize it, and resilient enough to keep observing through the outage.

The Challenge

The scale of the problem is substantial. In 2024, U.S. electricity customers experienced an average of 11 hours of power interruptions, nearly twice the annual average of the preceding decade. Major weather events accounted for 80% of that interruption time.

Utilities have invested heavily in modernizing their infrastructure. Between 2006 and 2009, the California Public Utilities Commission authorized approximately $2.2 billion for PG&E's advanced metering deployment and subsequent upgrade. These investments expanded access to energy-use information and improved metering capabilities.

But diagnosing electrical faults requires more than consumption readings or an outage alert. It requires measurements that capture the disturbance, observations from multiple locations, and monitoring that continues when grid power disappears. EADS is designed to address those three requirements:

- **Capture disturbances at the timescale they occur.** PG&E's interval consumption data can be recorded in hourly or 15-minute increments, depending on the meter configuration. Those readings describe energy use over time but do not preserve the electrical waveform within each interval. Electrical disturbances can unfold within milliseconds, and a brief voltage sag or transient may leave diagnostic evidence that an interval consumption total cannot reveal.
- **Keep monitoring through the outage.** Smart meters can send a "last gasp" notification when power is lost, but a notification alone does not provide a continuous record of conditions before, during, and after the interruption. EADS is designed to operate independently of grid power so that monitoring can continue through the event it is intended to observe.
- **Connect observations across the network.** A disturbance recorded at one location provides a local view. Comparing measurements across distributed sensors can help establish which locations were affected, how their electrical signatures differ, and where a fault may have originated. EADS aims to turn those distributed observations into timely diagnostic information, helping utilities investigate events with evidence captured as they happen.

The Approach

Sampling the waveform, not just the total

Each EADS box combines a Raspberry Pi 3 with a custom sensor HAT that samples the AC voltage waveform at 7,680 Hz, 128 measurements per 60 Hz cycle. This resolution captures the waveform's shape and the electrical signatures that can help identify and characterize a fault.

Detecting anomalies on the box, cycle by cycle

Lightweight detection algorithms analyze the stream directly on the box, cycle by cycle, flagging four classes of power quality anomaly in real time: full outage, undervoltage (covering both sags and sustained brownouts), overvoltage, and waveform distortion. The first three are identified by Root Mean Square (RMS) deviation from the 120 V standard; the fourth by comparing the measured cycle against an ideal gain-matched sine, which catches malformed and harmonically distorted waveforms that an RMS test alone would pass.

Each box maintains a rolling two-second waveform buffer in memory. When an anomaly is detected, it preserves the buffered measurements leading up to the event alongside the event itself, providing valuable context for diagnosis.

One distributed system, one API

Every EADS box runs an EdgeLake operator node that ingests the captured waveform data and stores it in a local SQL database. EdgeLake connects the boxes into one distributed system. Analysis and Monitoring Stations running on mini PCs host an EdgeLake metadata manager node that synchronizes metadata describing each EADS box and the data it hosts. Using this shared metadata, EdgeLake's query protocol identifies the relevant boxes, routes requests to them for local execution, and combines their results into a unified response.

Users, applications, and AI agents access the distributed data through a single API, as though they were querying a centralized cloud database, while the measurements remain stored locally on the EADS boxes. Through this API, the monitoring application lets operators and AI agents view live waveforms and statistics on a dashboard and track sensor locations and status on a map, without any dependency on centralized or cloud infrastructure.

As new EADS boxes are deployed, they automatically join the same distributed EdgeLake data network, making their measurements accessible through the same API. Monitoring coverage can expand across electrical grids, sites, homes, neighborhoods, and distribution networks without changing how applications access the data. This allows the team to focus on anomaly detection and fault analysis while EdgeLake handles data ingestion, discovery, and query coordination across the network.

Comparing neighbors on a shared clock

A shared time reference allows operators to compare measurements from neighboring boxes during the same event. These comparisons help establish whether a disturbance was confined to one home or affected a wider area, providing the evidence needed to narrow down a fault's location.

Why It Has to Run at the Edge

The cloud dependency is self-defeating. If a sensor must reach a data center to be analyzed, then the moment the grid goes down—taking local internet, and often local power, with it—you lose exactly the visibility you built the system to provide.

> A monitoring architecture whose failure mode coincides with the event it monitors is not a monitoring architecture.

Everything about this system's design follows from taking that constraint seriously.

Any sensing system capturing 7,680 samples per second across 1,000 devices generates more than 27.6 billion samples in one hour. Assuming four bytes per sample, that amounts to approximately 111 GB of raw data per hour—before adding timestamps, metadata, or transmission overhead. Continuously sending that volume over residential internet or 4G wireless links would impose substantial bandwidth and storage costs. EADS addresses this challenge through local processing: each box analyzes and stores its measurements locally, while selected anomaly captures, summaries, and query results move across the network as needed.

Keeping measurements local reduces data transfer and storage costs while giving homeowners control over sensitive electrical telemetry. High-frequency measurements can reveal daily routines and appliance use, so access policies limit what utilities and other authorized parties can retrieve. Utilities can access selected anomaly records and summaries to investigate outages, while detailed household measurements remain private.

For example, a utility could compare voltage-sag events across a neighborhood to narrow down a fault’s location and share relevant findings with repair crews to help restore power sooner. With access restricted to those records and summaries, the utility cannot query the underlying household telemetry to monitor residents’ daily routines or individual appliance use.

Results to Date

The project has moved from an earlier proof of concept that relied on synthetic data and a central server to a complete distributed pipeline running on real hardware.

Distributed pipeline: Live voltage data flows from a physical sensor, through on-box detection, into a local EdgeLake operator, across an encrypted overlay to the metadata manager and query nodes, and out to dashboards, with no cloud service anywhere in the data path.

Sustained 7,680 Hz acquisition: Continuous cycle-accurate sampling on a Raspberry Pi, alongside detection, local database storage, overlay networking, and time synchronization, all within a very tight CPU and memory budget. Making this work required dedicating a core to acquisition and hand-writing the ingest path in C.

Local anomaly detection: Four classes of power quality anomaly detected locally, cycle by cycle, with captured waveform windows for each.

Waveform-resolution visualization: The system can display the actual 60 Hz sine wave, sample by sample, pulled live from a sensor through the distributed query layer. Confirming this end to end required tracing correctness through every layer of the stack—sensor timestamping, schema, distributed query, and dashboard—which surfaced and fixed several real defects along the way.

Plug-and-play field deployment: A participant receives a box, plugs it in, and joins it to their home Wi-Fi through a captive portal on their phone. The box then enrolls itself with the metadata manager over pinned TLS, receives its network identity, and joins the fleet with no per-device configuration by the team. A box that is reflashed or wiped recovers the same identity, keyed to its hardware.

Retention policies: Routine measurements roll over within a fixed retention window. When local storage fills, routine data is automatically summarized and pushed to long-term storage servers running EdgeLake within the deployment environment, freeing space for new measurements. Captured anomaly segments are preserved without automatic deletion.

Boxes in the field: Five units have been delivered. With boxes in the field, the system is no longer one sensor and a monitoring station, but multiple independent nodes, capturing and storing their own data, and answering fan-out queries from the monitoring station as a single logical network

This is the point where the distributed architecture stops being theoretical.

Real-World Resilience

When a box goes quiet

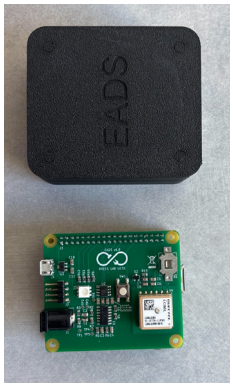
A site that has lost power and a box whose software has stalled are operationally different problems demanding different responses. A two-layer outage classifier distinguishes between them. When a node goes quiet, the monitoring station checks data freshness and dials the box directly over the overlay before deciding which it is. The same logic also refuses to guess: if the query node itself is unreachable, nodes are reported as unknown rather than falsely declared outages, so a monitoring failure never masquerades as a grid failure.

When the internet goes out

Measurements remain on the devices where they are captured, and monitoring continues through normal operation, power interruptions, and internet outages, recording electrical activity, loss of voltage, and anomalies locally. Queries execute across devices on the local network without a cloud dependency, and remote access resumes when connectivity returns.



A heartbeat from the Raspberry Pis, shown in Grafana



An EADS box deployed in the field

What's Next

The immediate next step is placing boxes in several different homes served by the same distribution line. This is the configuration the whole architecture was built for.

Monitoring a shared section of the grid

Two sensors inside one house share a service drop, so they see essentially the same electrical conditions. Sensors in separate houses on the same feeder do not, and the difference between what they see is where the diagnostic value lives.

A disturbance visible at one address is a problem inside that home. The same disturbance arriving at four addresses simultaneously is a problem upstream, and the set of sensors that saw it, along with the set that did not, brackets the affected segment of the line. Getting from "an outage occurred somewhere" to "the fault is on this span" requires exactly this: multiple independent observers on a known shared circuit, each with its own record of what happened.

It also puts the distributed query layer under the load it was built for. Correlating an event across houses means asking every node what it saw during a specific window and merging the answers, which is precisely the operation EdgeLake's single-system-image abstraction exists to make trivial.

Further planned steps

- **Hardware autonomy.** An onboard battery and cellular modem are the two components that would let a box fully outlive the event it is measuring, moving the system from "detects outages at sites with backup power" to "detects outages."
- **Automatic node location.** Each box carries a GPS receiver, and the path from a satellite fix to a marker on the outage map is built end to end. Because GPS rarely locks indoors, where these boxes live, an external antenna or a fallback location method is the practical next step. Until then, coordinates are entered once per site by hand.
- **Outage prediction.** Equipment often degrades before it fails, and that degradation frequently has an electrical signature: recurring sags, creeping distortion, intermittent excursions that resolve on their own. With enough recorded anomaly history across enough nodes, the goal is to flag developing problems before they become outages—the difference between a scheduled repair and an unplanned one.
- **Low cost cellular connectivity.** The next step is to also equip EADS boxes with 4G connectivity and evaluate low-cost cellular IoT services. The pilot will measure bandwidth usage, connectivity costs, and remote query performance while each box continues analyzing and storing measurements locally, extending monitoring to locations without reliable wired internet.

Watch the **EADS video demo** to see the system capture and display live waveforms.

Get Started with EdgeLake

EdgeLake empowers businesses to unlock the true potential of their edge data. By turning the edge into a virtual data lake, companies extract real-time insights from their distributed edge data without centralizing the data.

GitHub: github.com/EdgeLake

Web: lfedge.org/projects/edgelake