# Technical Brief: SigDrive

Version: 1.0

Date: January 2026

## 1. Executive Summary

Modern spectrum operations face a "Data Gravity" crisis. As sensors move to wideband capture and high-fidelity sampling, a single mission can generate terabytes of IQ data in minutes. Traditional infrastructure, reliant on local hard drives, "sneakernet" logistics, and proprietary file formats, cannot scale to meet this demand, leaving critical intelligence trapped in unindexed silos.

**SigDrive** is a purpose-built **Enterprise RF Data Lake** designed to solve this bottleneck. By decoupling massive binary storage from lightweight metadata indexing, SigDrive transforms isolated "binary blobs" into a queryable, secure, and AI-ready intelligence asset. Architected with an "Air-Gap First" philosophy, it delivers cloud-native capabilities to the tactical edge, ensuring compliance with **MOSA** mandates and enabling the **JADC2** data fabric.

## 2. The Core Challenge: Gravity and Chaos

SigDrive addresses two fundamental failures in legacy RF engineering workflows:

1. **The Physics of Data Gravity:** RF data is massive. A standard high-bandwidth recording (e.g., 500 MHz at 16-bit IQ) generates gigabytes per second. Moving 5TB of files across bandwidth-constrained networks (such as SIPRNet) is operationally unfeasible, forcing data to remain "dark" and inaccessible at the point of collection.

2. **Format Chaos (The Tower of Babel):** The ecosystem is fragmented by proprietary formats (e.g., X-DAT, TMP, and specialized binary formats). This locks government programs into specific hardware vendors and strips data of its context, separating binary samples from critical metadata such as center frequency and location.

# 3. System Architecture

SigDrive utilizes a microservices architecture that separates concerns between storage, compute, and metadata management. This design supports horizontal scaling and ensures that heavy binary data is moved only when necessary.

## 3.1 The Storage Layer (The Vault)

- **Object Storage (MinIO):** SigDrive replaces traditional file systems with high-performance, S3-compatible object storage. This allows for the retention of petabytes of data in a single namespace with erasure coding for fault tolerance.

- **Resumable Uploads:** Utilizing the open **TUS protocol**, the system supports reliable, chunked uploads of files exceeding 5TB, ensuring data integrity even over unstable tactical networks.

## 3.2 The Metadata Engine (The Brain)

- **Canonical Schema (SigMF):** Upon ingestion, SigDrive's **Pluggable Ingestion Engine** automatically normalizes disparate legacy formats (WAV, Midas Blue, Raw IQ) into the open-source **Signal Metadata Format (SigMF)**. This creates a universal, JSON-based interoperability layer.
- **Geospatial Indexing (PostGIS):** Metadata is indexed in a **PostgreSQL** database extended with **PostGIS**. This enables complex geospatial queries impossible in file-based systems, such as *"Find all signals recorded within this 50km polygon between 08:00 and 12:00"*.

## 3.3 Zero-Download Visualization

- **WebGL Spectrograms:** The frontend utilizes a React-based Single Page Application (SPA) with WebGL acceleration to render high-fidelity spectrograms and waterfall plots directly in the browser.

- **Tile-Based Streaming:** Similar to modern map servers, SigDrive streams visualization "tiles" rather than the raw file. This allows an analyst to visually inspect a 5TB recording instantly without downloading the underlying binary data.

# 4. Operational Capabilities

## 4.1 Air-Gap Ready Deployment

SigDrive is engineered for disconnected environments (SIPRNet, JWICS, Submarines) where internet access is prohibited.

- **Bundled Dependencies:** The system is delivered via **Zarf** packages or hardened Docker containers (Iron Bank UBI8 base images) that include all dependencies.

- **Offline Licensing:** Licensing is handled via cryptographically signed keys that validate locally, eliminating "phone-home" requirements.

- **Local Mapping:** Includes a containerized tile server (Martin/TileServer-GL) to serve NGA-standard vector tiles (MBTiles) for offline geospatial context.

## 4.2 AI & Machine Learning Readiness

SigDrive functions as a **Feature Store** for Cognitive EW and AI pipelines (e.g., Project Linchpin).

- **Region of Interest (ROI) Annotation:** Analysts can draw bounding boxes around signals directly in the browser. These annotations are saved as SigMF metadata, creating high-quality "Ground Truth" labels.

- **TorchSig Integration:** The normalized data structure integrates seamlessly with ML loaders such as TorchSig, enabling direct streaming from the Data Lake into PyTorch tensors for model training.

## 4.3 Security & Compliance

- **RBAC & IAM:** Integrated **Keycloak** identity management supports PKI/CAC/PIV authentication and granular Role-Based Access Control (RBAC) to enforce "Need-to-Know" data segregation.

- **Immutable Audit Logs:** The system logs every search, view, download, and metadata edit, satisfying the "Trustworthy" pillar of the DoD Data Strategy and NIST 800-53 Audit (AU) controls.

# 5. Strategic Alignment

## MOSA Compliance (Title 10 U.S.C. 2446a)

SigDrive directly enables compliance with the **Modular Open Systems Approach**. By utilizing SigMF as a documented, machine-readable interface, the government retains "Data Sovereignty," breaking the cycle of vendor lock-in and avoiding proprietary sustainment costs.

## JADC2 Interoperability

As a vendor-agnostic data layer, SigDrive facilitates the "Data Fabric" vision of **Joint All-Domain Command and Control**. It ensures that RF data collected by a Navy sensor can be easily queried, accessed, and analyzed by Army or Air Force assets, regardless of the original hardware manufacturer.

# 6. Technical Specifications Summary

| Component | Technology Stack |
|---|---|
| **Ingestion** | Python FastAPI, Pluggable Extractors, TUS Protocol |
| **Storage** | MinIO (S3 API), Erasure Coded, FIPS 140-2 Encryption |
| **Database** | PostgreSQL + PostGIS (STIG Compliant) |
| **Schema** | SigMF (JSON + Binary) |
| **Frontend** | React, WebGL, OpenLayers/MapLibre |
| **Deployment** | Kubernetes (Helm), Docker Compose, Zarf |
| **Security** | Keycloak (OIDC/SAML), mTLS, NIST 800-53 Controls |

## Ready to Modernize Your Spectrum Operations?

SigDrive transforms your RF data from a logistical burden into a strategic asset. Contact our engineering team to schedule a demo or discuss an evaluation deployment for your lab.
**info@sigdrive.com**