+33 (0)1 48 10 09 46
formation@ambient-it.net
140 Avenue Jean Lolive 93500 Pantin - France

Updated on 29/09/2025

Sign up

Apache OpenDAL Training

3 days (21 hours)

Overview

Apache OpenDAL is a unified data access layer for cloud-native environments. Multi-backend, it enables seamless interaction with object storage, files and distributed services, while offering performance and scalability.

Our OpenDAL training course will teach you how to abstract your data access, secure exchanges and industrialize your ETL/ELT pipelines with complete observability and cost control.

You'll learn how to exploit optional APIs for fine-grained control, integrate OpenDAL into your CI/CD chains and apply optimization patterns adapted to Data & Big Data workloads.

At the end of the course, you'll be able to build cloud-agnostic pipelines, make your processing more reliable and publish production-ready SRE runbooks.

Like all our training courses, this one covers the latest stable release and its new features, with a resolutely practical and operational approach.

Objectives

- Master OpenDAL and its optional APIs (read/write/list/stat/delete)
- Design high-performance, cloud-agnostic ETL/ELT pipelines
- Enhance security, resilience and observability
- Industrialize via CI/CD and IaC (Terraform/OpenTofu)
- Optimize costs and SLO with the right patterns

Target audience

- Cloud Engineers
- Data Engineers & Big Data teams
- Data platforms, SRE/DevOps

Prerequisites

- Cloud databases (AWS/GCP/Azure) and object storage
- Programming skills (Rust/Go/Python/Java)
- CI/CD and Terraform knowledge recommended

Our Apache OpenDAL training program

[Day 1 - Morning]

Fundamentals & getting started

- Understanding Apache OpenDAL: unified multi-storage data access layer
- Key principles: Open Community, Solid Foundation, Fast Access, Extensible Architecture
- Overview of backends: S3, GCS, Azure Blob, HDFS, POSIX...
- Operator / services / layers model and capabilities management
- Data & Cloud-native use cases: ingestion, ETL/ELT, feature store, serverless
- Practical workshop: OpenDAL & oli project, list/read S3 objects via temporary credentials.

[Day 1 - Afternoon]

APIs & options: clean data manipulation

- API read/write/list/stat/delete and Options-based APIs (ReadOptions, WriteOptions...)
- Metadata management, ranges, multipart and streaming
- Resilience strategies: retries, backoff, idempotence
- Multi-languages: Rust, Go, Python, Java (bindings)
- Best naming practices and prefixes for data lakes
- Practical workshop: read/write streaming and use of options (range, type, checksum).

Security & compliance

- Auth: SigV4, Service Accounts, SAS, OIDC
- TLS and rest-based encryption, vendor security policies
- Secret & variable management, Vault/KMS integration
- Account/project partitioning and governance
- Access logs & audits
- Practical workshop: S3 + KMS, assumed roles, audit trail verification.

[Day 2 - Morning]

Performance, reliability & industrialization

- Understanding latency & throughput; range requests
- Multipart upload, pipelining, batching, concurrency
- Cache-aside, write-through, read-through patterns
- Tiering & lifecycle; storage classes
- Observability: metrics, logs, traces; benchmarking with oli
- Practical workshop: multi-region bench + Prometheus/Grafana dashboards.

[Day 2 - Afternoon]

Robustness: real-life errors & testing

- Network errors, throttling and S3 200 + error cases
- Timeouts, circuit breakers, flow limiters
- Load and chaos tests
- Contracts & golden files
- Retries by operation
- Practical workshop: simulating throttling, adjusting retries/backoff, validating SLOs.

CI/CD & IaC

- Terraform/OpenTofu provisioning, CI secrets
- 12-Factor templates: config, blue/green, canary
- Packaging Rust/Go/Python/Java
- Security & policy as code scans
- Multi-environment promotion
- Practical workshop: CI provisioning pipeline + OpenDAL end-to-end testing.

[Day 3 - Morning]

Advanced integration & operations

- POSIX-like vs object store; parquet/csv/json implications
- Arrow/Polars/Pandas ecosystem + OpenDAL
- Pre-signed links, staging, atomic move
- Multi-cloud/hybrid and fallback strategies
- Data contracts & catalogs
- Practical workshop: ELT cloud-agnostic mini-pipeline with OpenDAL.

[Day 3 - Afternoon] Operations &

SRE

- Runbook incidents and operations
- Latency/error/throughput dashboard
- Alerting: p95, errors classified by backend
- Error budgets, SLO/SLI batch vs. interactive
- Cost optimization (listings, small files, multipart)
- Practical workshop: runbook + SLO & alert tests.

Production launch & team roadmap

- Security checklist, observability, performance, limits
- Versioning & compatibility; release tracking
- Prefix governance, environment separation
- Documentation: dev & operations guides
- Adoption plan and risk management
- Practical workshop: final review + simulated post-mortem.

Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced IT technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the forthcoming course, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or internal security difficulties

(intra-company or virtual classroom) that could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical training: 60% hands-on, 40% theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire is used to check that skills have been correctly acquired.

Certification

A certificate will be awarded to each trainee who completes the training course.