



# Introduction to IOMETE

The self-hosted data lakehouse platform

MARCH 2025

IOMETE pioneers a new approach to data management for complex enterprise data infrastructure environments with large data sets.

# Free your data from SaaS—self-host with ease.

SAAS

## The old guard

While SaaS solutions may offer convenience for smaller organizations, larger enterprises often face one-size-fits-all-rigidity, vendor lock-in, data leaks and runaway costs.

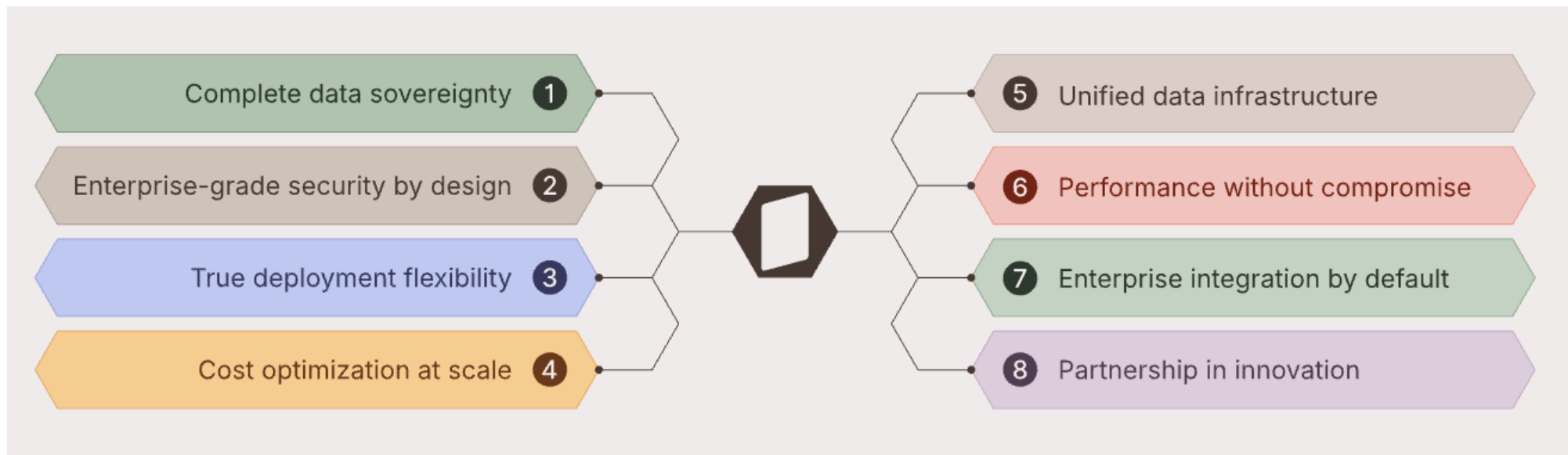
SELF-HOSTED

## IOMETE

Deploy IOMETE to your own infrastructure. No data ever leaves your security boundary, providing enterprises full control over their data at significantly lower costs. Ideal for highly regulated industries like Financial Services, Healthcare, Government and Technology.

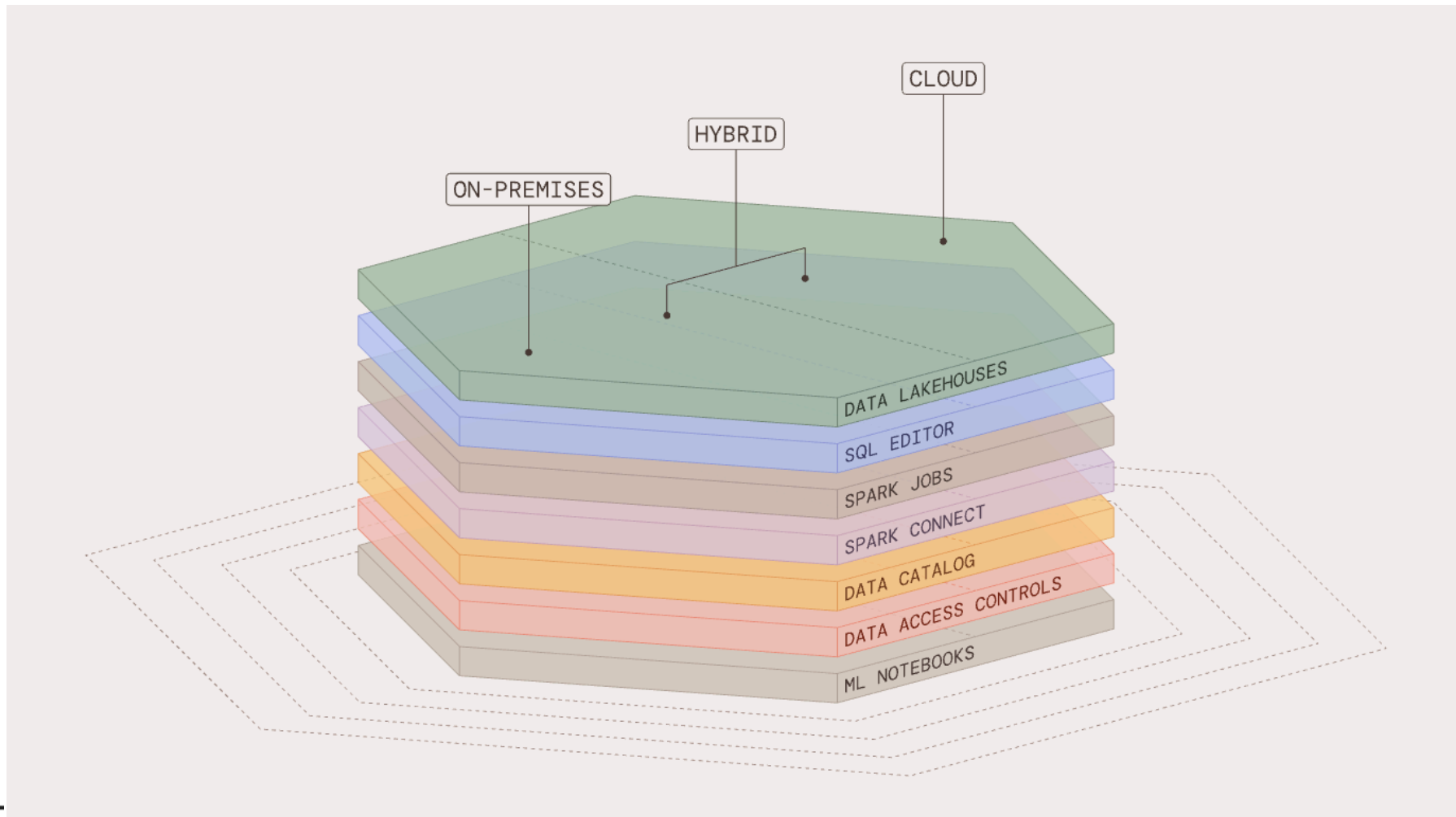
# Design principles

IOMETE was designed based on recent real-world feedback from enterprise clients.



# The platform

IOMETE includes everything a modern enterprise data team needs to easily manage and extract value from large data sets,



■ USE CASE

Dell is building the largest  
on premise data  
lakehouse in the world on  
IOMETE

*“IOMETE is a game changer in how  
Dell manages and extracts business  
value from petabytes of data.”*

**DELL**Technologies

Manikandan Rathinavelu

DIRECTOR OF DATA PLATFORMS, DELL TECHNOLOGIES

■ PLATFORM FEATURES

## See what IOMETE can do

Compute

Jupyter notebook

SQL editor

Query history

Spark application

Data catalog

Database explorer

Data security

Domain

Business Intelligence

PLATFORM FEATURES

# Compute

Easily spin up multiple lakehouses tailored to specific needs - separate environments for development, testing, and production workloads. Use different lakehouses for different business units or data domains.

Spark connect lets you seamlessly run Spark workloads from your preferred development environment. Code in Python, R, or SQL using familiar tools like Jupyter notebooks while leveraging IOMETE’s distributed computing power. Connect, analyze and process data at scale without leaving your trust perimeter.

Admin Admin

Development

Home

Engine

Compute

Jupyter Kernels

Workspace

SQL Editor

Applications

Spark Applications

Streaming Jobs

Job Templates

Governance

Data Catalog

Data Explorer

Data Products

Other

Settings

Compute

Filter by namespace

Filter by status

Search ID or Name

Refresh

Sort

Name	Driver	Executor	Auto suspend
<a href="#">medium-compute-cluster</a> ad26e0aa-e941-4f25-9cec-dcfcb0a70810	⏸ Stopped driver-medium 2vCPU • 15.6GiB	No running executors exec-medium 7.6vCPU • 60.5GiB • x4	✅ 30 minutes
<a href="#">data-engineering-team</a> 546c1e95-ba49-4a6e-b0c0-49810c797f15	🟢 Active xs-node 1vCPU • 2GiB	🟢 Running 4/4 exec-small 2vCPU • 15.6GiB • x4	✅ 30 minutes
<a href="#">arrow-flight-compute</a> d752a93e-c5bf-45d2-8d40-c52672786259	🟢 Active xs-node 1vCPU • 2GiB	🟢 Running 2/2 exec-medium 7.6vCPU • 60.5GiB • x2	⚠ Disabled

All items: 3 < 1 > 50 / page

Resources

Virtual Lakehouses

Get started to create your first compute

[Learn more](#)

Lakehouse cluster statuses

More about Lakehouse statuses

[Explore](#)

Spark Connect

Begin the process of crafting your cluster by referring to the step-by-step instructions provided in the documentation.

[Learn more](#)

Spark Connect quickstart guide

Jumpstart your journey with Spark Connect by using our quickstart guide. This repository includes hands-on PySpark Python and notebook examples, along with detailed instructions on how to use Spark Connect in your projects.

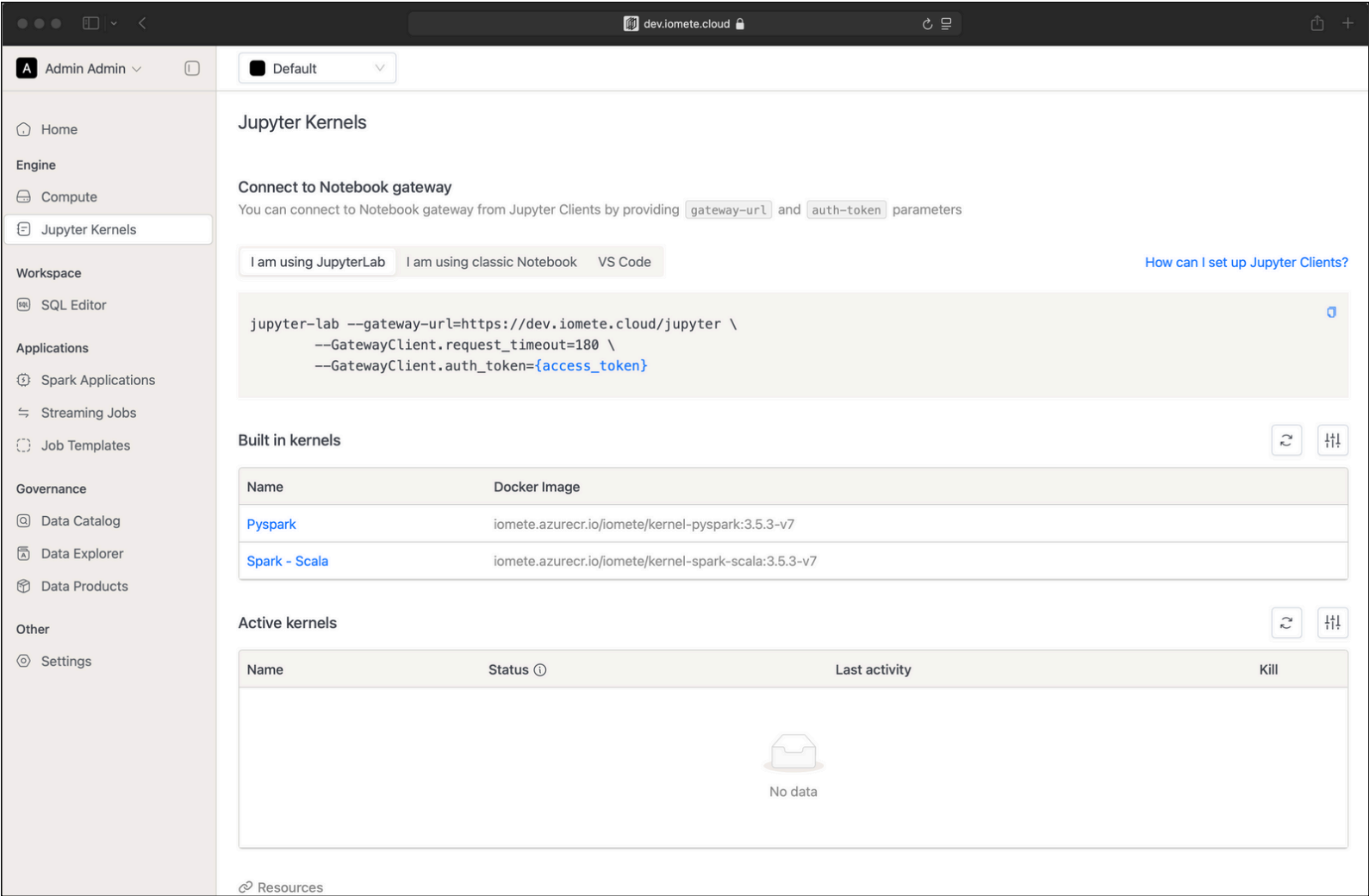
[Visit the Spark Connect quickstart repository](#)



▶ PLATFORM FEATURES

# Jupyter notebook

Connect to IOMETE’s lakehouses through interactive Jupyter notebooks. Run Python, R, or SQL code against your data using familiar notebook interfaces, with built-in kernels that handle distributed processing seamlessly.



PLATFORM FEATURES

# SQL editor

Write, analyze, and optimize SQL queries with IOMETE's built-in SQL editor. Features intelligent autocompletion, query history, and visual explain plans to help you write efficient queries against your lakehouse data.

Admin Admin

Development

Home

Engine

Compute

Jupyter Kernels

Workspace

SQL Editor

Applications

Spark Applications

Streaming Jobs

Job Templates

Governance

Data Catalog

Data Explorer

Data Products

Other

Settings

Worksheets

My Workspace

Data Products

IOMETE Samples

branch.sql

external-sources.sql

time-travel.sql

tpcds.sql

hudi\_experiment

Welcome.sql

Shared

Data Engineering

TPCDS Queries

Scheduled Queries

Repositories

gitlab-example

iom-test

tpcds.sql

arrow-flight-compute 2/2

spark\_catalog

tpcds\_10gb

▶

⋮

```
1 WITH customer_total_return AS
2 ( SELECT
3     sr_customer_sk AS ctr_customer_sk,
4     sr_store_sk AS ctr_store_sk,
5     sum(sr_return_amt) AS ctr_total_return
6 FROM store_returns, date_dim
7 WHERE sr_returned_date_sk = d_date_sk AND d_year = 2000
8 GROUP BY sr_customer_sk, sr_store_sk)
9 SELECT *
10 FROM customer_total_return ctr1, store, customer
11 WHERE ctr1.ctr_total_return >
12     (SELECT avg(ctr_total_return) * 1.2
13 FROM customer_total_return ctr2
14 WHERE ctr1.ctr_store_sk = ctr2.ctr_store_sk)
15 AND s_store_sk = ctr1.ctr_store_sk
16 AND s_state = 'TN'
17 AND ctr1.ctr_customer_sk = c_customer_sk
18 ORDER BY c_customer_id
19 LIMIT 100;|
20
21
```

6s 399ms • 100 rows

Data Chart SQL

	ctr_customer_sk ↓↑	ctr_store_sk ↓↑	ctr_total_return ↓↑	s_store_sk ↓↑	s_st
1	327680	86	2076.00	86	AAA
2	458752	58	2895.36	58	AAA
3	458752	26	1818.00	26	AAA
4	135168	46	2093.75	46	AAA
5	401408	91	3091.20	91	AAA
6	12288	98	6782.22	98	AAA
7	86016	46	1851.94	46	AAA
8	348160	86	1641.98	86	AAA
9	479232	68	1865.44	68	AAA
10	28672	68	3222.63	68	AAA
11	94208	26	1596.36	26	AAA

PLATFORM FEATURES

# Query history

Track and review all previously executed queries with detailed runtime statistics, execution plans, and performance metrics. Easily reuse successful queries and troubleshoot performance bottlenecks.

Admin Admin

Default

Home

Engine

Compute

Jupyter Kernels

Workspace

SQL Editor

Applications

Spark Applications

Streaming Jobs

Job Templates

Governance

Data Catalog

Data Explorer

Data Products

Other

Settings

Query History

Search

Last 1 hour

show tables;

spark\_catalog

show DATABASES;

compaction\_demo\_catalog

show DATABASES;

compaction\_demo\_catalog

show tables;

compaction\_demo\_catalog

SELECT \* FROM external\_oracle\_db.IOME...

compaction\_demo\_catalog

select \* from spark\_catalog.tpcds\_1000...

spark\_catalog.tpcds\_1000gb

show databases in rocco\_catalog;

spark\_catalog.tpcds\_1000gb

show databases in rocco\_catalog;

spark\_catalog.tpcds\_1000gb

select \* from parquet.`s3a://lakehous...

compaction\_demo\_catalog

SELECT \* FROM external\_oracle\_db.IOME...

compaction\_demo\_catalog

SELECT \* FROM external\_oracle\_db.IOME...

compaction\_demo\_catalog

Query details

Details

Data Result

ID

169d248d-f3d0-489c-ac6b-8b15a342fb25

Status

Completed

Compute

listener-test

Catalog & Schema

compaction\_demo\_catalog

Timing

Duration

3s 333ms

Start time

23 Feb, 2025 19:07 +04:00

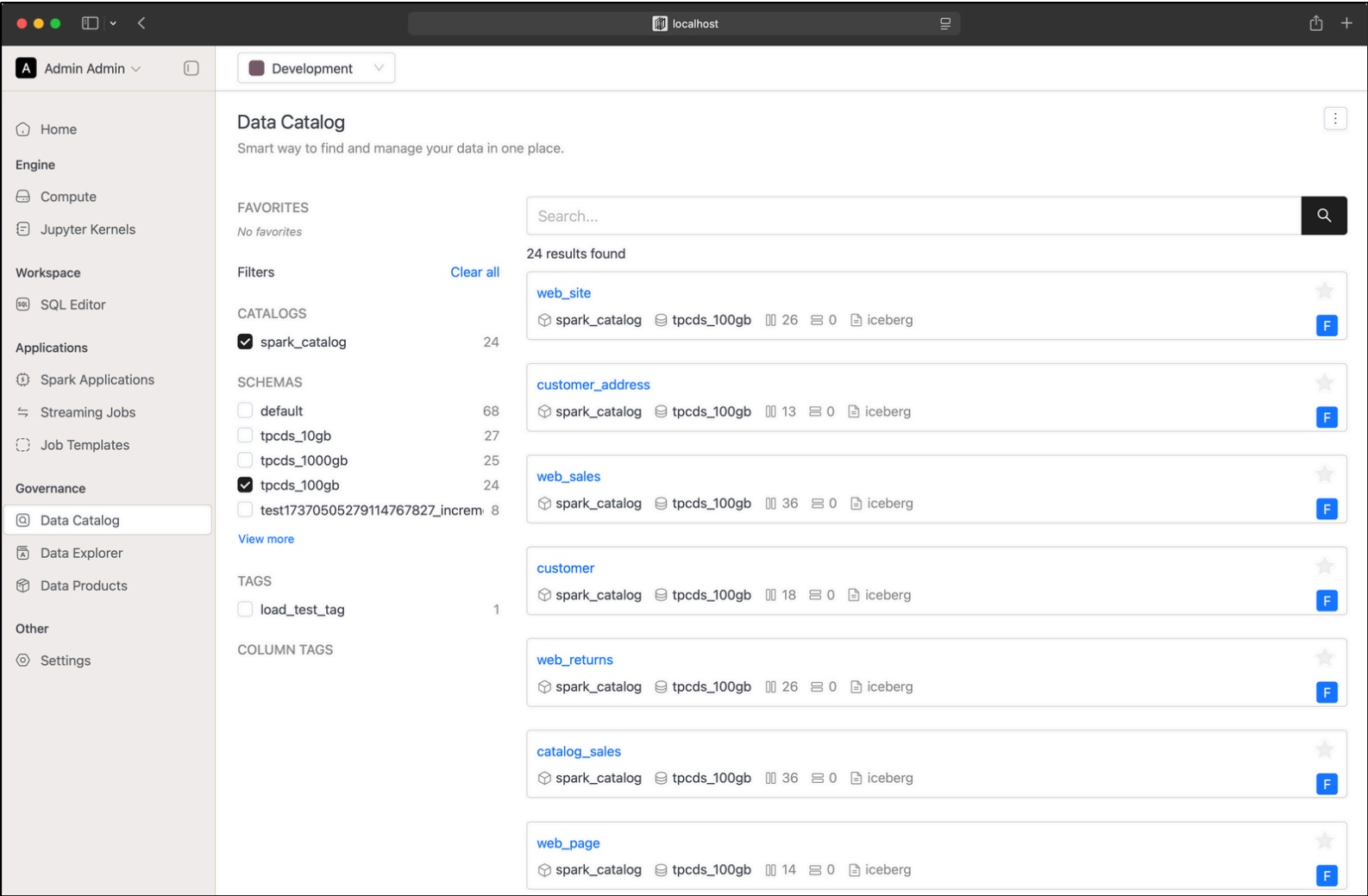
Finish time

23 Feb, 2025 19:08 +04:00

▶ PLATFORM FEATURES

# Data catalog

Easily discover, understand, and trust data assets with a centralized repository that provides metadata, lineage, and quality information.



PLATFORM FEATURES

Data security

Securely manage and control access to your data at the table, row, column as well as the user, group, team, and department level, ensuring that only authorized individuals and teams can access the specific data they need.

Protect sensitive data by automatically replacing it with realistic but fictional values, allowing you to safely use data for development, testing, and analytics without compromising privacy.

Use tags to enforce access policies based on data categorization. Organize and categorize your data using custom tags, making it easier to search, filter, and manage your data assets. This helps improve data discovery, data security, and understanding across your organization.

Admin Admin

Admin Portal

Select Domain

Domains

Monitoring

IAM

Users

Groups

Admin Roles

LDAP

Single Sign-on

Data Governance

Spark Catalogs

Data Security

Data Policies

Audit

Compute

Node Types

Volumes

Namespaces

Docker Registries

Administration

System config

Event Logs

Other

Settings

Data Security Policies

+ Create

Search

Refresh

Filter

Name	Policy Type	Expiration Time	Status
<a href="#">Access Policy to Compaction Catalog</a> 16cb0678-cf50-4d6a-b470-39e03a23601b	Access Policy		Active
<a href="#">Deny Access to PII and PCI Table</a> 8e5311bd-08aa-4d54-9416-d117ddbc2c5e	Access Policy		Active
<a href="#">Filter VIP Customers</a> 11ab5171-0599-4332-b773-6bf57c2b10b2	Row Filter		Active
<a href="#">Mask PII Data Policy</a> 771cc53f-b7eb-488c-95ac-155e6c3662de	Data Masking		Active
<a href="#">Policy All Users - View</a> 53dd3ce6-be83-4cfe-8e26-bc9cb60c2213	Access Policy	15 May, 2025 03:01 +04:00	Active
<a href="#">Policy for External Audit Team</a> 4d691176-f588-47ad-bdb6-2377384a7b60	Access Policy	23 Feb, 2025 02:50 +04:00	Expired

All items: 6 < 1 > 20 / page

Resources

Data Access Policy

Guidelines for accessing sensitive data securely.

[Learn more](#)

Data Masking Policy

Introduction to data masking techniques to protect sensitive information.

[Learn more](#)

Row Level Filter

Implementing row-level filtering to control access to specific rows of data.

[Learn more](#)

▶ PLATFORM FEATURES

# Data domains

IOMETE lets you organize and govern your data by subject area (e.g. customer, product, finance) to promote data ownership, quality and to simplify data discovery, access control, and compliance with regulations like GDPR.

Admin Admin

Admin Portal

Select Domain

Domains

Monitoring

IAM

Users

Groups

Admin Roles

LDAP

Single Sign-on

Data Governance

Spark Catalogs

Data Security

Audit

Compute

Node Types

Volumes

Namespaces

Docker Registries

Administration

System config

Event Logs

Other

Settings

Domains

+ Create

Search

↺

⌵

Name	ID	Created by	Created at
AI team	ai_team	admin	a month ago
Product	product	admin	a month ago
Development	development	admin	a month ago
Marketing	marketing	admin	a month ago
ML team	ml_team	admin	a month ago
Finance	finance	admin	21 days ago
Management	management	admin	21 days ago
Default	default	admin	a month ago
OPS Domain	ops_domain	admin	6 days ago

All items: 9

< 1 >

20 / page

# Connect BI

Easily connect to your favorite BI tool by using the SQL endpoint.

Admin Admin

Development

Home

Engine

Compute

Jupyter Kernels

Workspace

SQL Editor

Applications

Spark Applications

Streaming Jobs

Job Templates

Governance

Data Catalog

Data Explorer

Data Products

Other

Settings

data-engineering-team

MetricsSpark UIConfigureRestartTerminate

DetailsConnectionsLogsKubernetes events0 / 15ActivityConfiguration

PythonJDBCDBTTableauPower BISupersetMetabaseRedashSpark ConnectArrow Flight

Tableau - Spark SQL Driver

Parameter	Value
Connection	SparkThriftServer (Spark 1.1 and later)
Server	dev.iomete.cloud
Port	443
Authentication	Username and Password
Username	admin
Password	<a href="#">Access token</a> <small>Generated Tokens are used instead of <u>password</u> for connecting to compute.</small>
Transport	HTTP
HTTP Path	/data-plane/spark-resources/lakehouse/data-engineering-team
SSL	<input checked="" type="checkbox"/> Require SSL

Tableau - Spark SQL by CData

Parameter	Value
General	
Server	dev.iomete.cloud

At IOMETE, we understand that at the heart of great engineering is the drive to solve problems. That's why IOMETE is not just a platform, we consider ourselves an extension of your engineering team.



# Straightforward pricing

Our generous Free Tier provides access to all core features (up to 100 vCPUs). Our Enterprise Plan - \$500 per vCPU per year - adds advanced data access and security features with enterprise-grade support, and the Business Critical Plan delivers hybrid and multi-region deployment with dedicated expert assistance.

PLAN TIER

Free

The world's most generous free plan

LICENSE COST

Zero

MAX # OF VCPUS

Max 100 vCPUs

DEPLOYMENT OPTIONS

Self-hosted, on-premises deployment

FEATURE SET

Core features

SUPPORT PLAN

Community support

FREE

Get started with IOMETE's Free tier today.

Start now ↗

PLAN TIER

Enterprise

Scale and performance with peace of mind

LICENSE COST

\$500 license cost per vCPU per year <sup>1</sup>

MAX # OF VCPUS

Unlimited vCPUs

DEPLOYMENT OPTIONS

Self-hosted, on-premises deployment

FEATURE SET

Core features + Data masking, Role level security, Disaster zone, Auditing

SUPPORT PLAN

Enterprise support

ENTERPRISE

Find out more about the Enterprise plan.

Get in touch

PLAN TIER

Business Critical

Tailored for mission-critical infrastructure

LICENSE COST

Custom license cost <sup>2</sup>

MAX # OF VCPUS

Unlimited vCPUs

DEPLOYMENT OPTIONS

Self-hosted, on-premises or hybrid deployment

FEATURE SET

Enterprise features + Multi-region

SUPPORT PLAN

Enterprise support + Dedicated engineers

BUSINESS CRITICAL

Learn more about the Business Critical plan.

Get in touch

<sup>1</sup> \$50,000 per year (100 vCPU) minimum commitment

<sup>2</sup> \$150,000 per year minimum commitment

Thank you.

