MINIO

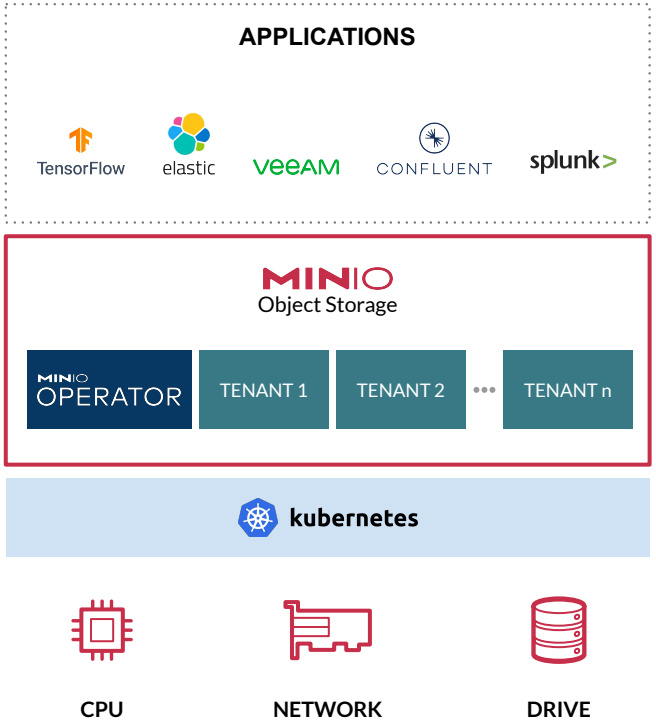
DISA Technical Introduction

October 2023

What is MinIO

MinIO is a **high performance, Kubernetes-native object store.** It is designed for large-scale data infrastructure.

It was built from scratch to be cloud native and is the most decorated storage company in the market.



Guiding Principles Align with Defense/Intel

Durability

Full suite of technology to ensure data is safe. Inline erasure coding, bit rot protection, active active replication, lifecycle management.

Simplicity

Seamless scalability. Launches with a single command.

Runs the same from the edge to the core.

Ideal for day two operations.

Portability

Entire binary is <100MB.

Runs on any HW, any cloud, with any application.

Full support for AI/ML.

Security

MinIO's state-of-the-art encryption schemes support granular object-level encryption using modern, industry-standard encryption algorithms, such as AES-256-GCM, ChaCha20-Poly1305, and AES-CBC.

Full Enterprise Feature Set

S3 COMPATIBLE

Amazon S3 API is the de facto standard for object storage. MinIO implements Amazon S3 v2/v4 API + S3 Select

KUBERNETES-NATIVE

MinIO was built natively for RESTful APIs - no connectors needed. More than 62% of MinIO instances are containerized with 43% of those managed via Kubernetes.

IMMUTABILITY

Complete suite of object locking capabilities is Cohasset certified - required by financial services and healthcare customers.

ACCESS MANAGEMENT

MinIO Server integrates with Identity Providers such as WSO2, Keycloak, Okta, Ping Identity to allow applications or users to authenticate and use Object Storage.

ENCRYPTION & TAMPER-PROOF

MinIO provides confidentiality, integrity and authenticity assurances for encrypted data with negligible performance overhead.

FULL INTERACTION SUITE

MinIO offers a developer-driven command line interface in addition to an IT-friendly graphical user interface.

LAMBDA COMPUTE

MinIO server triggers Lambda functions through event notification service. OCR, audit compliance are good examples of lambda computing.

ERASURE CODE & BITROT PROTECTION

You may lose up to half the number of drives and still recover from it. Data protection code is accelerated using SIMD instructions on x64 and ARM CPUs.

ACTIVE ACTIVE REPLICATION

A resilient and scalable architecture that can withstand a complete DC failure.

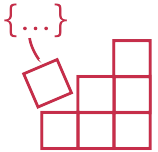
MONITORING

API driven capabilities with a suite of web hooks to your favorite apps.

MinIO Deployment Use Cases

AI & Machine Learning

HDFS replacements



High performance data lake/warehouse infrastructure

IOT and Airgapped Storage

Multi-cloud environments (portability)

Endpoint for streaming workloads

By deploying for performance, long term archival storage and disaster recovery are effectively free.

Delivering Business Value Across Customers

CROWDSTRIKE

- Data lake use case repatriated from AWS

- **100PB growing to 500PB over the next year**
- **2-3% GM improvement for the entire business**
- **Crowdstrike uses ⅓ of an FTE to manage 100 PB+ of data on MinIO**
- **250TB ingest per day**

AMERICAN EXPRESS

- HDFS migration to a modern data lake (6+PB) for SIEM solution

- **Reduced data infrastructure costs (as measured by servers) by more than 84%**
- **Improved performance by 33%**
- **Headcount down by 87%**

Walt Disney

- AWS data repatriation for streaming use case

- **500 PB going to 5EB over the next 2 years**
- **Reduced TCO by 50x**

PayPal

- 10 PB legacy appliance based solution needed S3
- Failed SLAs due to replication challenges on previous platform

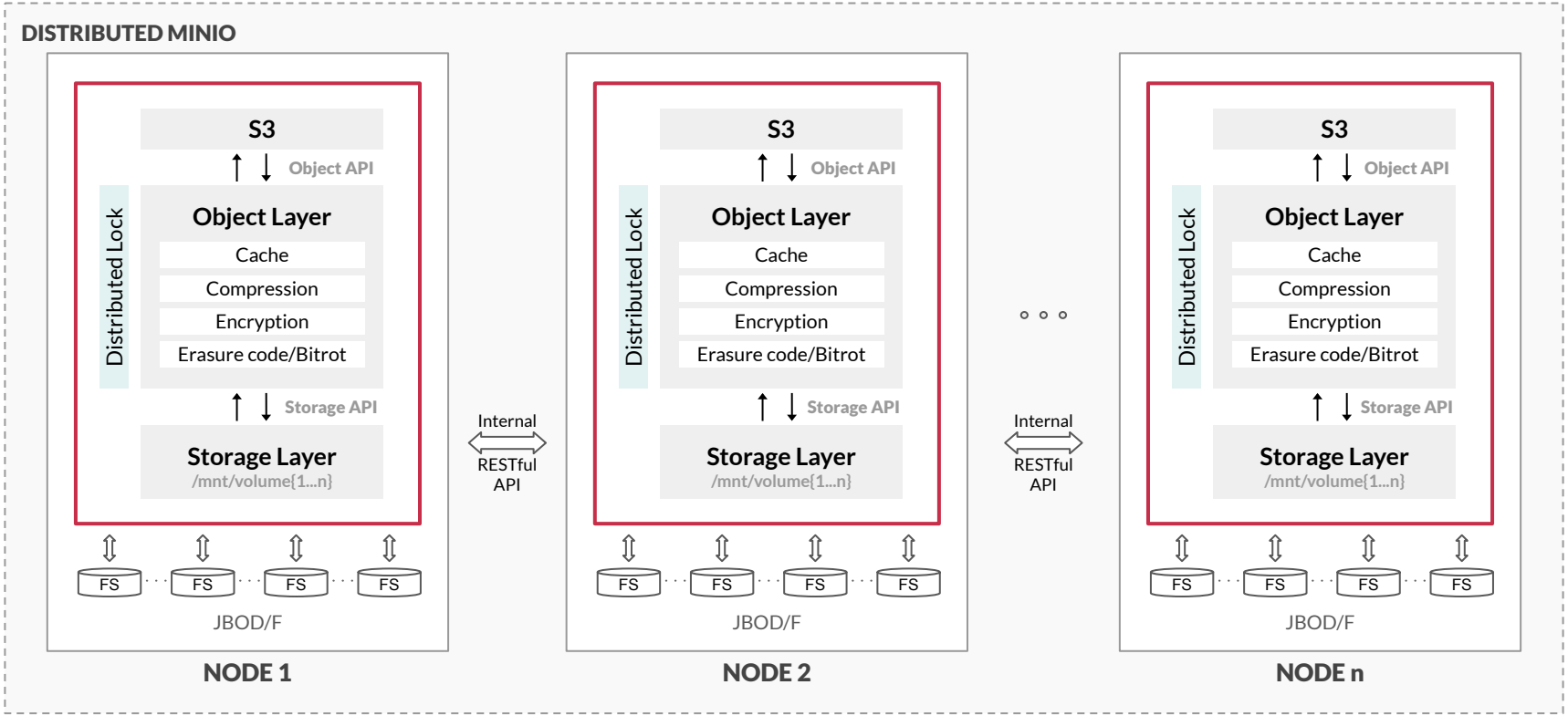
- **Merchant reporting, median Services and web crawler fully migrated**
- **8 clusters across 2 sites (800km apart) with over 5 billion objects**
- **Small object workload (avg 250kb)**

Strong Government Penetration

BAE		ARBETSFÖRMEDLINGEN SWEDISH PUBLIC EMPLOYMENT SERVICE	NASA	esa
9001-northanalysiscon-dozzui-b2db5hwzycu.ws.app.gitpod.gss.gov.uk	minio-sts.va.th.ten	minio-console.arbetsformedlingen.se	console.er7.minio.luna.nasa.gov	engdev141.dev.esoc.esa.int:9001
	GOUVERNEMENT Liberté Égalité Fraternité	Prävention Detektion Reaktion		U.S. AIR FORCE
minio-console.cfg.intranet.cnaf.info	minio-console.dev.kube.insee.fr	backups3prd-01.bpa.bund.de:9090	minio-py.min.io	minio.udn.dev:9001
Défense nationale National Defence	nationaal archief	Rijkswaterstaat Ministry of Infrastructure and Water Management	Skatteforvaltningen	
minio-console-cornerstone.apps.si01.lv1.ocp.c1b	production-console-minio-operator.apps.prod.frosting.ea.com	tst-minio-console.k8s.datalab.cnap-am2.intranet.rws.nl	test-minio-route-osm2-devops-02.ocpt.ccta.dk	minio-dashboard-gitlab-nswc-runner2.apps.hx-k8s.nswc.navy.mil
OAK RIDGE National Laboratory	EUROCONTROL	Sandia National Laboratories	JOHNS HOPKINS APPLIED PHYSICS LABORATORY	Federale Overheidsdienst FINANCIEN
deathstar.ornl.gov:19001	minio.cluster2.doublepulsar.win	127.0.0.1:59383	vhadr-meineca1.livelab.jhuapl.edu:9001	minio.tooling02.finbel.intra

MinIO Internals

MinIO Internals



The Standard for Performance

NVMe

8 Node NVMe 100 Gbe Network	Read/GET	Write/PUT
Distributed	46.54 GB/s	34.4 GB/s
Distributed/Encrypted	46.4 GB/s	34.6 GB/s

32 Node NVMe 100 Gbe Network	Read/GET	Write/PUT
Distributed	325 GB/s	171.3 GB/s

HDD

16 Node HDD 25 Gbe Network	Read/GET	Write/PUT
Distributed	10.81 GB/s	8.57 GB/s
Distributed/Encrypted	9.38 GB/s	6.91 GB/s

24 Node HDD 25 Gbe Network	Read/GET	Write/PUT
Distributed	16.3 GB/s	9.4 GB/s

Source: https://blog.min.io/scaling-minio-more-hardware-for-higher-scale/

What Makes MinIO Fast

SINGLE LAYER

We are a single layer, object only.

Multiple layers cause latency, complexity.

NO METADATA DATABASE

By writing object and metadata together you make all operations single and atomic.

Multiple steps for other vendors.

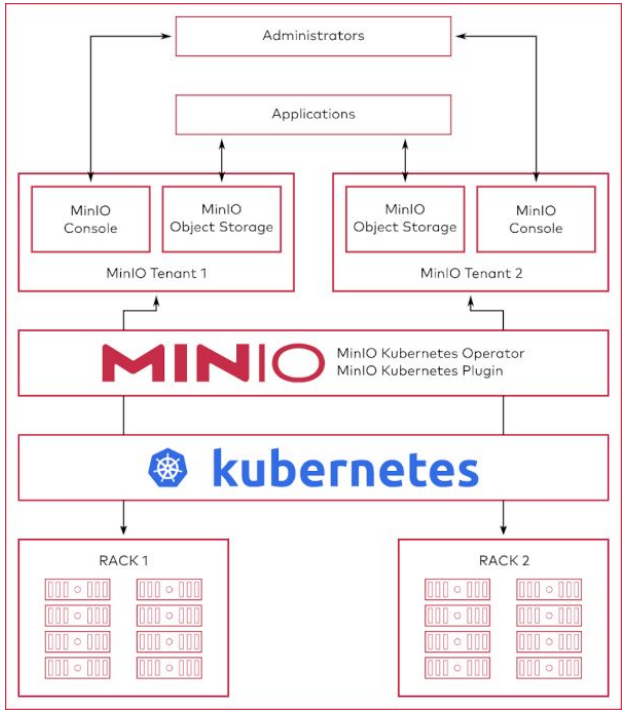
SIMD ACCELERATION

By writing the core parts of MinIO in assembly language (SIMD extensions, e.g. AVX512, NEON, VSX) we are hyperfast on commodity HW.

COMBINATION OF GO + GOASM

Delivering faster than C performance by combining GO + Assembly Language and targeting them to the task.

MinIO Deployment Option - On Prem

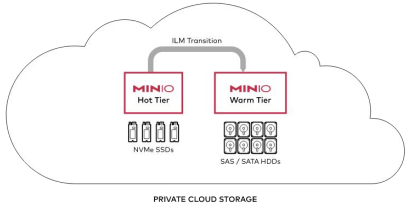


Deploy and manage multiple isolated tenants within the same Kubernetes cluster.

Tenants are separately configured for capacity, erasure coding, hardware resources, encryption, and identity providers.

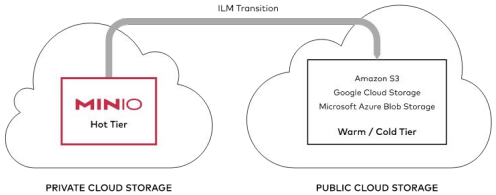
Tenants are fully isolated, protected from disruption from the others, and scales independently.

Tiering: Achieving Economics & Efficiency



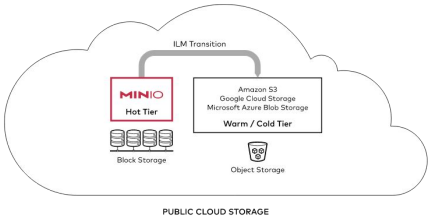
Across Storage Types

HDD → SSD



Private to Public

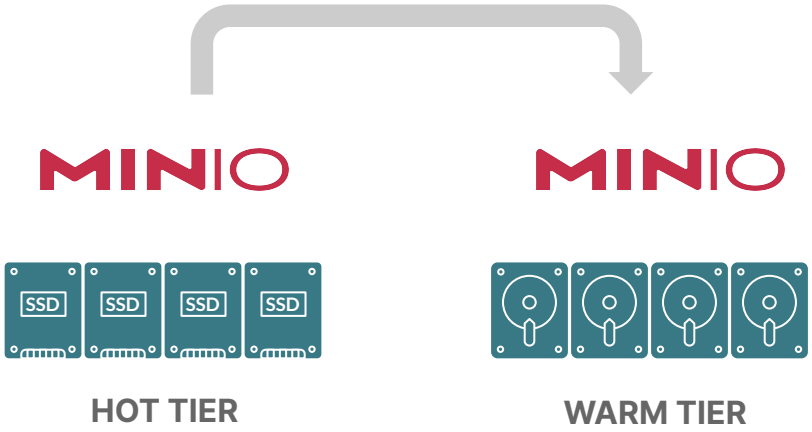
Private hot tier to public warm or cold (depending on requirements)



Within Public

Manage performance/price across ANY cloud.

MinIO Object Lifecycle Management



- Combine SSD for hot and HDD for warm tier
- Bucket level policy base tiering - names, tags, timeline
- Transition or Expire infrequently used objects
- Transparently fetch objects from warm tier

Object Retention, Legal Holds & Lifecycle

DATA RETENTION

Ensures that an object is protected (cannot be deleted or overwritten) for a set period of time

Operate in Compliance and Governance modes

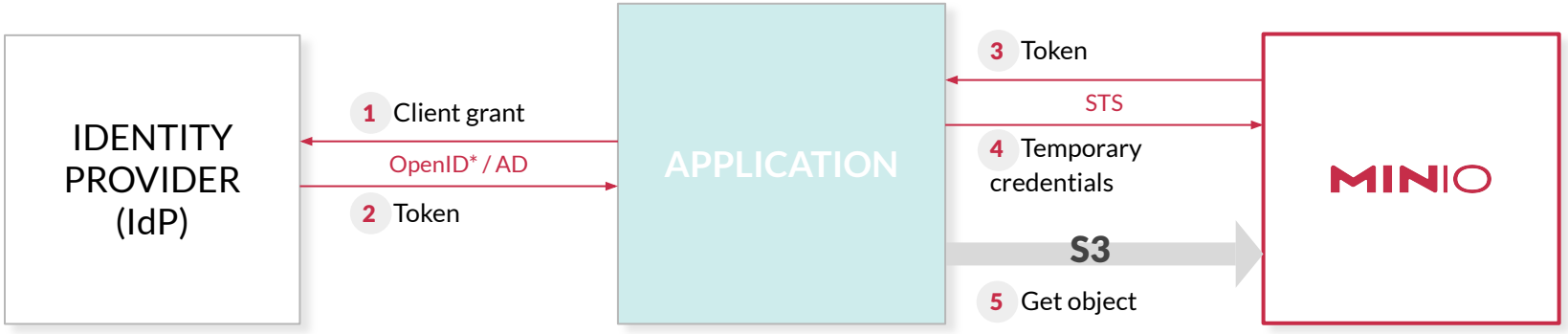
LEGAL HOLD

Offers the same protection as the retention period but it has no expiration date

OBJECT LIFECYCLE

Automate data lifecycle management activities such as lifecycle policies update, transition and deletion of data

Identity & Access Management



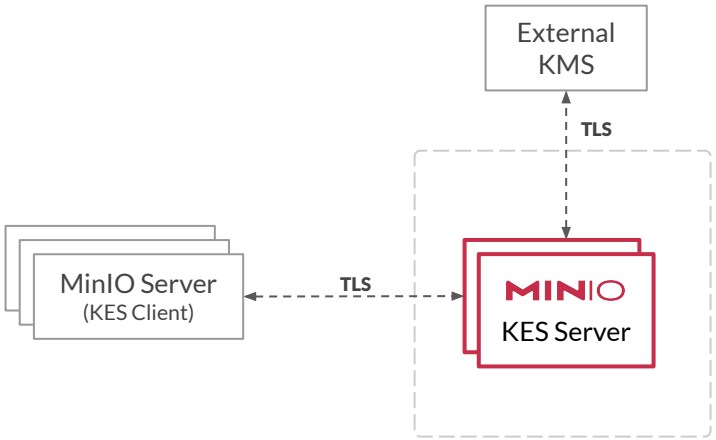
*

LDAP and OpenID (OIDC)-compatible providers

Azure Active Directory · Active Directory · Google Cloud Identity · OpenLDAP · okta · KEYCLOAK

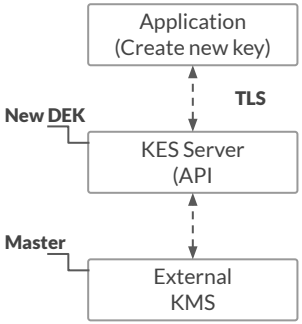
16

Encryption & Security

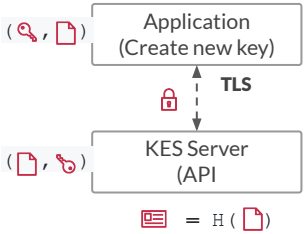


MinIO KES is a tool for managing and distributing secret keys at scale. In particular, it decouples a traditional key-management-system (KMS) - like AWS-KMS or Hashicorp Vault or Thales from large-scale and high-performance applications.

Components and Flow



Authentication and Authorization



Supported External Key Management Systems

HashiCorp

gemalto
security to be free

AWS Secrets Manager

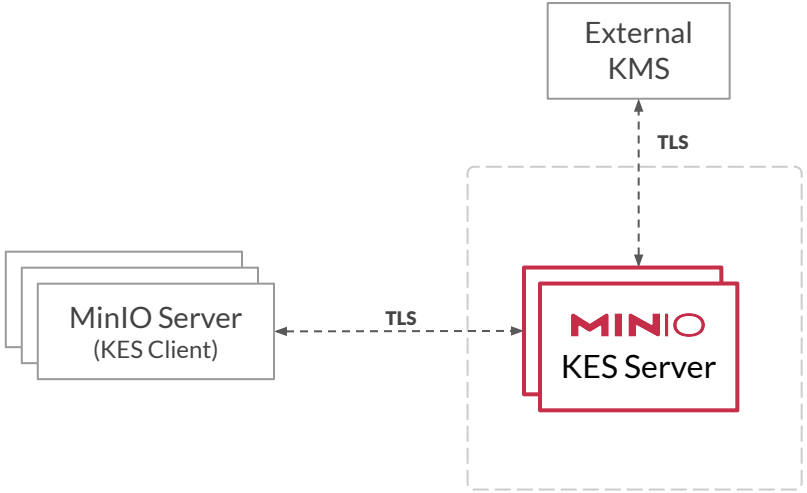
Secret Manager

OASIS STANDARD KMIP

ENTRUST

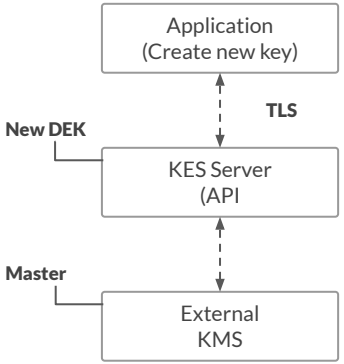
17

MinIO KES

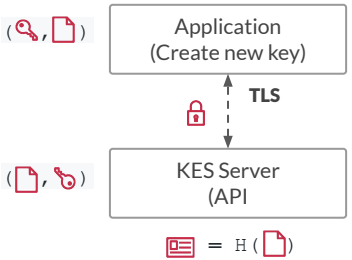


MinIO KES is a tool for managing and distributing secret keys at scale. In particular, it decouples a traditional key-management-system (KMS) - like AWS-KMS or Hashicorp Vault or Azure Key Vault from large-scale and high-performance applications.

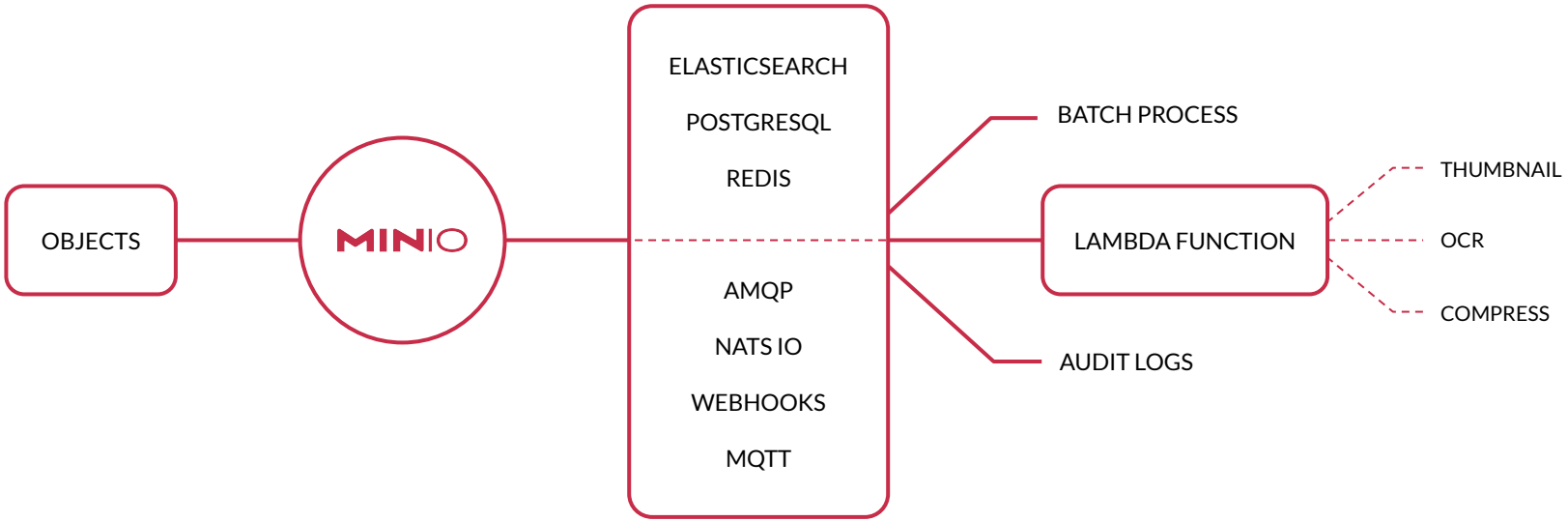
Components and Flow



Authentication and Authorization



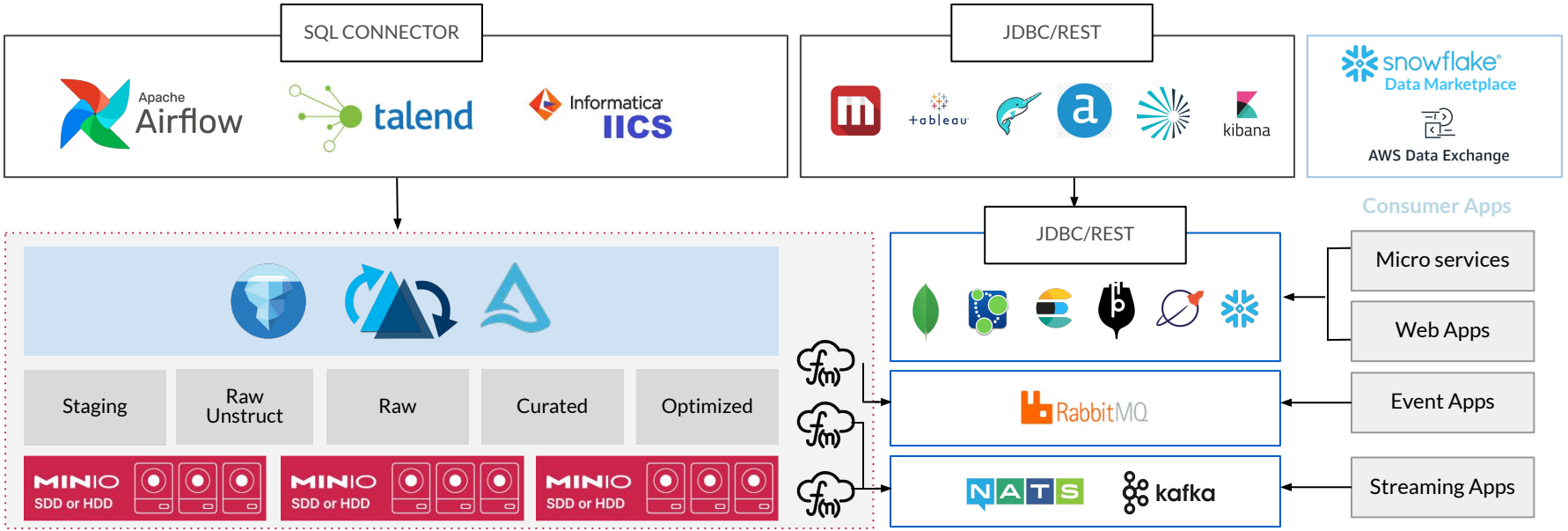
MinIO Lambda Functions



Data Consumption Patterns with MinIO

- Populate Optimized data to appropriate datastore for performance with Airflow pipeline/talend/IICS
- Leverage Tableau/Alteryx/Dremio for BI/Reporting and Self Service Analytics
- Leverage Pinot/Snowflake/Elastic for Portal Analytics
- Minio pre & post notifications to assist with streaming, cdc, event use cases

MinIO lambda functions can be used update various marketplaces for 3rd party consumption



Pricing and Summary

Commercial License Capabilities

License	MINIO STANDARD	MINIO ENTERPRISE
	Designed for customers who require a commercial license and can mostly self-support but want the peace of mind that comes with the MinIO's unique direct-to-engineer interaction model and suite of operational capabilities delivered through the SUBNET portal. The Standard version is fully featured but with SLA limitations. To learn more about the features and capabilities available to commercial license holders click here.	Designed for mission critical environments where both a license and strict SLAs are required. The Enterprise version is fully featured but comes with additional capabilities over the Standard license. To learn more about the features and capabilities available to commercial license holders click here.
FEATURES		
Term	Annual (monthly billing option is available)	
Software Release	Enterprise Hardened	
Release Support	1 Year Long Term Support	5 Years Long Term Support
SLA	Next Business Day	Less than 4 hours
Support	L4 Direct Engineering support via the SUBNET portal	L4 Direct Engineering support via the SUBNET portal, Phone or Web Conference
Panic Button	1 Per Year	12 Per Year
Critical Security and Bug Detection	Continuous Scan and Alert	
Health Diagnostics	24/7/365	
Performance Diagnostics	✓	✓
Annual Architecture Review		✓
Annual Performance Review		✓
Indemnification		✓
Security and Policy Review		✓

- True direct-to-engineer support
- Average SLA measured in minutes
- Completely compatible with air gapped systems

MINIO

Summary

Highly Adaptable, Horizontal System

MinIO features, flexibility, size and performance make it ideal for a wide range of use cases.

MinIO Drives Superior Economics

MinIO's cloud-operating model and legendary simplicity drive better business outcomes - from performance to cost.

Half the HW, 2x the performance, ¼ of the FTEs...

Broad Adoption in Governments Across Globe

Governments, Agencies, Defense Contractors and SIs all deploy MinIO to satisfy their objectives.

Significant opportunities engage at a deeper level to achieve more.

MINIO

Thank you

- @minio
- https://github.com/minio/minio
- https://slack.min.io
- https://min.io