dremio

Dremio Admin as a Service

Shared Responsibility Model

Dremio operates on a shared responsibility model to ensure an optimal customer experience and mutual partnership.

This Shared Responsibility Model provides customers with a robust data lakehouse platform on several types of customer-managed infrastructures, providing a range of deployment options and configurations.

This document:

- Summarizes the shared responsibility model to help better understand how to configure and maintain the Dremio environment for optimal operation.
- Details areas of ownership between Dremio, the customer and a Dremio Admin as a Service (DAaaS) provider in managing a production, enterprise-grade environment.

Please see the Dremio Software documentation and resources below for additional information. Dremio also offers a variety of virtual and instructor-led training courses and other assets. Contact your Dremio representative for further details.

January 2025

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Platform	**Platform Software** ● Deliver secure, high-quality, compliant software. ● Harden software and images. **Specifications** ● For each deployment type, document minimum system compute and network requirements, such as EKS, AKS, standalone clusters, and YARN. ● Document requirements for metadata and distributed storage.	**Deployment** ● Provision system and storage infrastructure per Dremio's minimum requirements (or higher) for the selected deployment type. ● Install and configure software components, including primary and secondary coordinators, engines, and clients per Dremio's requirements, using Dremio-provided binaries and images. ● Configure persistent logging in Kubernetes environments. ● Manage account configurations, administration, security, subscriptions, and cloud resources. **Integration** ● Integrate third-party services and clients using Dremio's REST API or other supported endpoints.	**Operation and Expansion** ● Ensure engine sizing and auto-scaling are appropriate for the variety of workloads to be processed (either via static or elasticity configurations); including adding new secondary coordinators or new engines when necessary. ● Ensure adequate distributed, metadata, and data object storage performance, including sufficient space, bandwidth, and minimal latency. ● Ensure the deployment complies with Dremio limits regarding the catalog, query execution, metadata, query planner, and workload size. ● Conduct routine release upgrades, at least annually, to ensure the use of a supported Dremio version. Test and approve updates before installation.
Network Connectivity	**Specifications** ● Document minimum system and network requirements.	**Provision & Configure** ● Ensure networks meet Dremio's minimum requirements (or higher). ● Follow cloud provider best practices in creating virtual network environments, e.g., AWS and Azure. ● Manage the security, bandwidth, and performance of networks connecting to Dremio and within the Dremio deployment.	**Network Review** ● Assess network data to determine any issues with the network that are impacting Dremio

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Identity	**Identity Management** ● Support industry-standard authentication and Single Sign-On (SSO) services, including Oauth 2.0 / OpenID Connect.	**Identity Management** ● Configure integrated authentication (Active Directory or OpenID Connect) to centrally manage user accounts with strong password policies and SSO/multi-factor authentication (MFA). ● Enable System for Cross-domain Identity Management (SCIM) with the Identity Provider (IDP). ● Configure LDAP connectivity, caching, and searching according to your organization's authentication architecture and policies. ● Configure SSO IP allow lists to limit IDP access to Dremio and other authorized applications.	
Access Control	**Access Management** ● Provide role-based access privileges and data access policies to manage access to data and platform features with fine granularity.	**Access Management** ● Manage users and roles and determine role membership, including the admin role. ● Implement access management best practices, including regular user access audits, at least every 6 months.	

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Data	**Data Security Configurations** ● Support several storage and wire encryption configurations using TLS 1.2 or higher and HTTPS. **Data Access** ● Provide role-based access privileges and data access policies to manage access. **Git for Data** ● Dremio connections to Nessie catalogs enable you to perform Git-for-data activities such as branching and versioning.	**Data Security** ● Secure management of Dremio and your data infrastructure. Deploy and configure wire encryption to the level your use cases require. ● Deploy and manage customer-owned encryption keys. ● Manage expiry/refresh schedules of customer-owned encryption keys so that the customer knows to prepare new keys that can be swapped in prior to the expiry of old ones. **Data Governance** ● Utilize Dremio's role-based access control and data access policies to limit sensitive data access according to the least privilege principle. ● Revise roles, policies, privileges, and dataset ownership with user onboarding and offboarding. ● Construct a semantic layer following Dremio's best practices including - layer views into sub-layers, use roles for directory and dataset access, and leverage tags and wiki for dataset information. ● Follow best practices in the creation and use of Parquet-based datasets. ● Understand the operation of Apache Iceberg tables and catalogs.	**Iceberg Maintenance** ● Understand the operation of Apache Iceberg tables and catalogs. Periodically optimize Iceberg tables and vacuum unneeded snapshots.

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Jobs & Query Execution	**Engines** - Provide configurations to scale the control plane and engines for workload requirements. **Query Management** - Show a consolidated view of jobs and job details, configurable to include a variety of job states and statuses. - Provide job routing configurations to determine which engine to use for a given query. - Include a query profile for each query showing a runtime breakdown, reflections used, and other query performance information.	**Query Management** - Employ the Dremio job overview and raw and visual profiles to understand performance factors. - Periodically review query performance and take action to improve performance. Use Dremio raw or visual profiles to understand runtime behavior and pinpoint bottlenecks.	**Infrastructure Management** - Apply workload management rules and engine configurations to provide all jobs and job types with appropriately sized and configured engines. - Many workloads include extremely high-cost queries and significant workload variance; plan for these queries in engines, queues, and routing configurations. - Periodically ensure platform coordinators, engines and executors, and workload routing rules and queues are appropriate for query workloads.
Metadata	**Metadata Configurations** - Enable data source configuration of metadata collection and refresh.	**Metadata Management** - Refresh metadata on demand using Dremio SQL commands when required.	**Metadata Management** - Periodically optimize the metadata refresh process. Define and tune engines, queues, routing configurations, and refresh rates appropriate for the metadata refresh workload. - Configure metadata expiration to minimize inline metadata refresh due to potentially negative query performance impact.

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Reflections	**Reflection Definition** - A reflection is an optimized materialization of source data or a query, similar to a materialized view. - Dremio's query optimizer can accelerate a query against tables or views by using one or more reflections to partially or entirely satisfy that query rather than processing the raw data in the underlying data source.	**Reflection Management** - Utilize best practices in managing the lifecycle of reflections. - Understand when queries use reflections, the factors influencing the use of reflections, and how to apply reflection recommendations. - Verify reflection behavior in the query profile as required.	**Reflection Management** - Periodically optimize the reflection refresh process. Right-size reflection’s resources and tune routing configurations and refresh rates.
Monitoring	**Monitoring Technologies & Logs** - Provide a range of technologies, including REST, JMX, and Dremio system tables for integrating Dremio with enterprise monitoring. - Generate query and audit history logs showing the use of the platform.	**Implementation & Use** - Configure Audit Analyzer and Query Analyzer to analyze logs after storage in a persistent cloud location. - Employ monitoring tools that continuously monitor Dremio-recommended metrics and send alerts on service thresholds. - Deploy verbose cloud service security monitoring (e.g., AWS CloudWatch and CloudTrail) where appropriate.	**Review** - Periodically review workload sizes, characteristics, and historical trends; right size coordinators, engine resources, queues, or routing configurations. **Storage** - Monitor the distributed store and coordinator metadata store for sufficient headroom. Expand allocated space and clean coordinator metadata storage as required.

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Availability	**Business Continuity** ● Review Business Continuity plans and conduct drills annually. ● Provide standard practices and steps to conduct DR procedures on Dremio environments.	**High Availability** ● Install and configure the Dremio-recommended HA solution for the underlying infrastructure. **Disaster Recovery** ● Define a Recovery Point Objective (RPO) and Recovery Time Objective (RTO) on your lakehouse, including Dremio, all critical data sources, and other customer-owned resources. ● Perform incident response, disaster recovery, and contingency planning procedures.	**Backup & Restore Testing** ● Regularly back up the Dremio deployment, including the KV store, configuration files, logs, data sources, Dremio's distributed store, and other customer resources. Periodically test restore in the event of failure.

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Platform Security	**Vulnerability Management** ● Maintain a vulnerability management program; see the Dremio Vulnerability Management Policy. ● Publish all security-related notifications in the Security Bulletins section of Dremio's documentation. ● Publish an updated list of security fixes and responses to security vulnerabilities impacting Dremio through the supply chain under the Dremio Software Release Notes. **Application Security** ● Follow the Secure Software Development Lifecycle and employ tooling to detect vulnerabilities, including Static Analysis and Security Tooling (SAST), open source software scanning, and AMI scanning. ● Conduct third-party penetration tests at least annually. ● Periodically review cryptographic standards to select and update technologies and ciphers per assessed risk and market acceptance of new standards.	**Deployment & Operation** ● Understand the Dremio admin CLI and the activities of the Dremio administrator. Additional information is available in the support knowledge base. ● Minimize the number of Dremio administrators; grant selected workgroup users the Dremio administrative privileges required by workgroup teams. ● Automate the promotion of Dremio objects from lower environments using the Dremio REST API or Dremio Cloner.	

Area	Dremio Responsibility	Customer Responsibility	DAaaS Responsibility
Compliance	**Standards & Compliance** ● Maintain independent third-party audits, standards, and certifications of compliance: ○ ISO 27001 ○ SOC 2 Type II ○ HIPAA ● Adhere to privacy regulations such as GDPR and CCPA.	**Adherence to Standards** ● When processing sensitive data such as PII or PHI, adhere to relevant privacy regulations such as the GDPR, CCPA, or HIPAA. ● Comply with applicable laws and regulations.	**Adherence to Standards** ● When processing sensitive data such as PII or PHI, adhere to relevant privacy regulations such as the GDPR, CCPA, or HIPAA. ● Comply with applicable laws and regulations.