dremio

Dremio Cloud

Shared Responsibility Model

Dremio Cloud 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 while relieving their operational burden since Dremio operates and manages the environment, including the Dremio control plane and the Dremio software that runs computing resources in the data plane.

This document:

- Summarizes the shared responsibility model to help better understand how to configure the Dremio environment for optimal operation.
- Details areas of ownership between Dremio and the customer in creating a production, enterprise-grade environment.

Please see the Dremio Cloud 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.

November 2024

Area	Dremio Responsibility	Customer Responsibility
Platform	**Platform Services** ● Secure the Dremio control plane. ● Harden deployed images and operating systems. Maintain the Dremio control plane with updated software and images. **Managed Resources** ● Securely deploy and terminate Dremio-managed systems. ● Track security configurations against industry standard baselines. ● Deploy the latest applicable source code and system images upon launch of customer data plane hosts.	**Account Management** ● Create the prerequisite cloud and project resources. ● Manage account configurations, administration, subscriptions, and cloud resources. **Adoption** ● Ensure use cases comply with documented service limits. Ensure datasets comply with supported file limits. ● Integrate third-party services and clients using Dremio's REST API or other supported endpoints.
Network Connectivity	**Network Communications** ● Deploy cloud security groups within the customer data plane to isolate workloads. ● Use secure defaults for network access controls and security groups within the control plane. ● Separate the control plane from the customer data plane using multiple layers of network security controls.	**Data Plane Communications** ● Create or provide network resources as part of the setup. ● 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.
Identity	**Dremio Operations** ● Authenticate Dremio Cloud personnel using industry best practices. **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 SSO IP allow lists to limit IDP access to Dremio and other authorized applications. ● Use the least-privilege principle for cross-account Identity and Access Management (IAM) roles, such as access to the project store.

Area	Dremio Responsibility	Customer Responsibility
Access Control	**Dremio Operations** ● Define Dremio operations employee privileges consistent with least-privilege principles. ● Limit access to systems processing customer data to employees with roles that warrant access. ● Secure interactions with the customer cloud environment. ● Secure storage and policy enforcement of secrets scope. ● Perform quarterly access reviews to maintain minimal access posture. **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 memberships, including the admin role. ● Implement access management best practices, including regular user access audits, at least every 6 months.
Data	**Data Security** ● Transmit customer content using TLS 1.2 or higher between the customer client and the control plane and the control plane to the data plane. ● Encrypt customer data at rest within the control plane using AES-256 bit equivalent or higher. **Data Access** ● Provide role-based access privileges and data access policies to manage access. **Git for Data** ● Provide an integrated lakehouse management catalog based on Project Nessie to perform Git-for-data activities such as branching and versioning.	**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. ● Configure automatic optimization of Dremio lakehouse resources or perform periodic optimization of Iceberg tables in other catalogs. **Data Security** ● Deploy and manage customer-owned encryption keys.

Area	Dremio Responsibility	Customer Responsibility
Jobs & Query Execution	**Engine Autoscaling** - Automatically scale engines based on the customer's configuration parameters. Engine replicas are started and stopped as required by monitoring query load and engine replica health. **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.	**Engine 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 and routing configurations. - Periodically ensure engines and workload routing rules are appropriate for query workloads. **Query Management** - Periodically review query performance and take action to improve performance. Use Dremio raw or visual profiles to understand runtime behavior and pinpoint bottlenecks. - Employ the Dremio job overview and raw and visual profiles to understand performance factors.
Metadata	**Metadata Configuration** - Enable data source configuration of metadata collection and refresh.	**Metadata Management** - Periodically optimize the metadata refresh process. Define and tune engines, routing configurations, and refresh rates appropriate for the metadata refresh workload. - Configure metadata expiration to minimize inline metadata refresh due to potential negative query performance impact. - Refresh metadata on demand using Dremio SQL commands when required.
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. - Periodically optimize the reflection refresh process. Right-size the reflection refresh engine and tune routing configurations and refresh rates.

Area	Dremio Responsibility	Customer Responsibility
Monitoring	**Security Monitoring** ● Deploy security detection capabilities, including those provided natively by cloud service providers. ● Maintain an intrusion detection system across computing resources. ● Employ an incident response framework to manage and minimize the effects of unplanned security events. ● Notify customers of security breaches per data protection laws and customer agreements. **Audit & Query History** ● Generate audit and query history showing the use of the platform; make history available through system tables.	**Cloud Monitoring** ● Deploy verbose cloud service security monitoring (e.g., AWS CloudWatch and CloudTrail). Monitor cloud provisioning, configuration changes, and usage. **Query & Audit History** ● Periodically review workload sizes, characteristics, and historical trends; configure any required changes in engine resources or routing configuration.
Availability	**Platform Availability** ● Maintain availability and security of the Dremio control plane and engines that run on the customer environment. ● Provide continuous service status on status.dremio.com, including customer notifications of service interruptions. For further information, see the Dremio Cloud Terms of Service. **Control Plane DR** ● Provide resilience to zonal failures. ● Review Business Continuity and Disaster Recovery plans annually; conduct Business Continuity and Disaster Recovery drills annually. ● Conduct periodic backups of the Dremio control plane.	**Data Plane DR** ● Define DR processes appropriate for the customer’s data sources, the Dremio project store, and other customer-owned resources. Back up these resources regularly.

Area	Dremio Responsibility	Customer Responsibility
Platform Security	**Vulnerability Management** ● Maintain a vulnerability management program; see the Dremio Vulnerability Management Policy. ● Publish an updated list of security fixes and responses to security vulnerabilities impacting Dremio through the supply chain under the Dremio Cloud Changelog. **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. **Service Management** ● Build and manage infrastructure using infrastructure as code. The cloud production infrastructure is regularly monitored for compliance violations and security anti-patterns using Cloud Infrastructure Security Posture Management (CISPM).	**Operational Best Practices** ● Minimize the number of Dremio administrators; grant selected workgroup users the Dremio administrative privileges required by workgroup teams. Additional admin information is available in the support knowledge base. ● Automate the promotion of Dremio objects from lower environments using the Dremio REST API.
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.