databricks

Mapping ITGC Best Practices to Databricks' Customer Capabilities

Version 1.0

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE

Strategy Operations and Governance	Elect a Business Sponsor and Steering Committee.
	
	Designate a business sponsor and steering Committee to oversee and govern your security and compliance program.
		Databricks recommends that customer organizations consider creating leadership oversight to govern their IT compliance program.
		
		■ Determine whether your Databricks Platform is within your IT General Controls (ITGC) compliance program's scope.
		■ Based on the size and type of your organization, an IT security and compliance steering committee may be created and represented by:
		  ■ Chief Compliance Officer (CCO)
		  ■ Chief Financial Officer (CFO)
		  ■ Chief Information Officer (CIO)
		  ■ Chief Information Security Officer (CISO)
		  ■ Chief Legal Officer (CLO)
		  ■ Chief Technology Officer (CTO)
		■ Customers who implement such cross-functional leadership steering committees may benefit from regular oversight updates to provide further accountability, security, and compliance over their organizational data.

Strategy Operations and Governance	Create a Center of Excellence (CoE) to sustain your IT compliance. Build a program with all relevant stakeholders and senior leaders to ensure transparency and success of your IT compliance obligations.	Databricks recommends that customer organizations determine where Databricks fits into their IT compliance program concerning their ITGC compliance obligations. Learn more about the Databricks Shared Responsibility Model and closely partner with your Databricks account team as required to support the technical, security, and compliance requirements. The Databricks shared responsibility model outlines data and services security and compliance obligations of Databricks, the cloud service provider, and the customer. Access documentation for AWS, Azure, or GCP here. Further, Databricks offers an extensive learning platform that includes self-paced and instructor-led courses. Please refer to the entire training catalog here.
Access Controls	Onboarding/Offboarding of Access: develop procedures for formal access approval requests for onboarding and offboarding. Ensure that procedures include role transfers to ensure appropriateness.	Databricks customer organizations are responsible for ensuring that authorized users are appointed as organizational administrators, and they should follow a defined logical access management process to manage and protect the admin account user credentials to administer the lakehouse architecture, where applicable. Customers are responsible for appropriately managing user access to the lakehouse architecture and customer-managed AWS, Azure, or GCP account(s). Databricks recommends that customer organizations take advantage of the following: ■ Capability to configure single sign-on and unified login. AWS | Azure | GCP ■ SCIM synchronization of users and groups. AWS | Azure | GCP Please refer to the Databricks SOC 1 Type II Control(s): Complementary User Entity Controls: # 1 and 2.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Access Controls	Implement role-based access control (RBAC), which requires formal requests and approvals for access. Ensure users have the minimum access necessary for their roles, minimizing unnecessary access.	Databricks access control lists (ACLs) enable you to configure permissions for accessing and interacting with workspace objects, including folders, notebooks, experiments, models, clusters, pools, jobs, Delta Live Tables pipelines, alerts, dashboards, queries, and SQL warehouses. https://docs.databricks.com/en/security/auth/access-control/index.html Databricks Unity Catalog provides a unified data governance solution for customers to manage their identity and user access solution for Databricks. Customers may create and use the metastore to register metadata about securable objects (such as tables, volumes, external locations, and shares) and the permissions to govern access. In the Unity Catalog, users must be in a workspace attached to a metastore in their region. https://docs.databricks.com/en/data-governance/unity-catalog/create-metastore.html To learn how to configure identity in Databricks best, see Identity Best Practices here. Roles may follow the best practice below, where there are three types of Databricks identities: Users: ■ User identities are recognized by Databricks and represented by email addresses. Service principals: ■ Identities for use with jobs, automated tools, and systems such as scripts, apps, and CI/CD platforms. ■ Databricks recommends creating service principals to run production jobs or modify production data. ■ All processes that act on production data should run using service principles. Interactive users do not require write, delete, or modify privileges in production, which reduces the risk of a user accidentally overwriting production data. Groups: ■ Groups simplify identity management, making assigning access to workspaces, data, and other securable objects easier. ■ It is best practice to assign access to workspaces and access-control policies in Unity Catalog to groups instead of users individually. All Databricks identities may be designated as members of groups, and members inherit permissions assigned to their group. Account Administration*: ■ Databricks has administrative roles that can manage Databricks identities: ■ Account admins: can add users, service principals, and groups to the account and assign them admin roles. They can give users access to workspaces as long as those workspaces use identity federation. ■ Workspace admins: can add users and service principals to the Databricks account. They can also add groups to the Databricks account if their workspaces have been enabled for identity federation. Workspace admins can grant users, service principals, and groups access. ■ Group managers: can manage group membership. They can also assign other users the role of group manager. ■ Service principal managers: can manage the roles of a service principal.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE

		Workspace Administration
		■ Manage identities and access control:
		  ■ Workspace administrators are responsible for adding users to the workspace, assigning appropriate roles (User/Admin) to principals, and managing entitlements (ACLs) at the workspace level.
		  ■ Oversee implementation of role-based access controls (RBAC) for various workspace objects, such as clusters, pools, jobs, and tables.
		■ Create and manage compute resources:
		  ■ Workspace admins create and manage SQL warehouses and clusters for workspace users.
		  ■ Regulate compute usage through the implementation of cluster policies, management of workflow tags, and the enforcement of naming conventions.
		  ■ Further, they perform cluster health reviews to ensure best practices are implemented.
		■ Configure workspace features and settings:
		  ■ Workspace administrators are responsible for overseeing and managing workspace behavior and settings. This includes turning features on/off at the workspace level, configuring SSO and cluster policies, and directing security and data protection measures. Further, they oversee the creation of queries and dashboards to monitor critical systems and processes.
		*Databricks recommends a limited number of account admins per account and workspace admins in each workspace.

Access Controls	Least Privilege: ensure users have the minimal access necessary to perform their job duties.	Databricks access control lists (ACLs) enable you to configure permissions for accessing and interacting with workspace objects, including folders, notebooks, experiments, models, clusters, pools, jobs, Delta Live Tables pipelines, alerts, dashboards, queries, and SQL warehouses. https://docs.databricks.com/en/security/auth/access-control/index.html Databricks Unity Catalog offers customers granular control to define user privileges in their workspaces. Please see the documentation here: https://docs.databricks.com/en/data-governance/unity-catalog/manage-privileges/privileges.html

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE

Access Controls	Data Masking and Redaction: where feasible, use data masking to restrict access to sensitive data for users who do not need complete visibility.	ITGC compliance demands robust protection of sensitive financial data. Unity Catalog offers fine-grained security controls that allow organizations to implement sophisticated data protection measures.
		
		These include:
		
		- Column-level Masking
		  - Control and restrict access to a specific column (of potentially sensitive information) to a particular group of users, which is helpful when working with sensitive data (PII, PCI PAN data).
		  - Use the mask function to apply data masking to specific columns. https://docs.databricks.com/en/tables/row-and-column-filters.html
		- Row-Level Security
		  - Limit data access based on user roles or attributes
		  - Control access to specific rows in your data, which is useful when different users or user groups should have access to different subsets of the data.
		  - Use row filters to apply a filter to a table so that subsequent queries only yield the appropriate results.
		- Mapping tables
		  - Use mapping tables to create an access-control list
		  - Define a mapping table (or access-control list) to achieve row-level security. Each mapping table is a comprehensive mapping table that encodes which data rows in the original table are accessible to certain users or groups. Mapping tables are helpful because they offer simple integration with your fact tables through direct joins.
		  - This methodology proves beneficial in addressing many use cases with custom requirements. Examples include:
		  - Customers may impose restrictions based on the logged-in user while accommodating different rules for specific user groups.
		  - Customers may create intricate hierarchies, such as organizational structures, requiring diverse rules.
		  - Customers may replicate complex security models from external source systems.
		  - By adopting mapping tables in this way, you can effectively tackle these challenging scenarios and ensure robust row-level and column-level security implementations. https://docs.databricks.com/en/data-governance/unity-catalog/row-and-column-filters.html
		
		These features enable organizations to implement a defense-in-depth strategy for data protection, aligning with ITGC requirements for safeguarding financial information.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Access Controls	Maintain appropriate separation of duties (SoD) for all users with access to production and development systems to establish proper checks and balances in place. Ensure workflows and access permissions to segregate duties among users, preventing conflicts of interest and reducing risk.	Databricks customers are responsible for ensuring that authorized users are appointed as organizational administrators and follow a defined logical access management process to manage and protect the admin account user credentials to administer the lakehouse architecture, where applicable. Customers are responsible for appropriately managing user access to the lakehouse architecture and customer-managed AWS, Azure, or GCP account(s). Databricks recommends maintaining separate admin accounts from standard user accounts. Catalogs, which are part of Databricks Unity Catalog offering, may help provide segregation across your organization's information architecture. They may correspond to a software development environment scope, team, or business unit. Customers may use workspaces as a data isolation tool, with different workspaces for production and development environments or a specific workspace designated for sensitive data. In that case, you can also bind a catalog to specific workspaces to handle all specified data processing in the appropriate workspace. Further, Databricks maintains distinct administrative roles and various capabilities. See below: Account Administrator may: ■ Provision Principals (Groups/Users/Service) and SSO at the account level ■ Identity Federation refers to assigning Account Level Identities access to workspaces directly from the account ■ Configuration of Metastores ■ Setting up Audit Log ■ Monitoring Usage at the Account level (DBU, Billing) ■ Creating workspaces according to the desired organization method ■ Managing other workspace-level objects (storage, credentials, network, etc.) ■ Automating dev workloads using IaaC to remove the human element in prod workloads ■ Turning features on/off at the Account level, such as serverless workloads and Delta sharing. Workspace Administrators may assign/define: ■ Appropriate roles (User/Admin) at the workspace level to Principals ■ Appropriate entitlements (ACLs) at the workspace level to Principals ■ Setting SSO at the workspace level ■ Cluster Policies to entitle Principals to enable them to ■ Compute resources (Clusters/Warehouses/Pools) ■ Orchestration (Jobs/Pipelines/Workflows) ■ Turning features on/off at the Workspace level ■ Entitlements to Principals ■ Data Access (when using internal/external hive metastore) ■ Manage Principals' access to compute resources ■ Managing external URLs for features such as Repos (including allow-listing) ■ Controlling security and data protection ■ Turn off / restrict DBFS to prevent accidental data exposure across teams ■ Prevent downloading result data (from notebooks/DBSQL) to prevent data exfiltration

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
		■ Enable Access Control (Workspace Objects, Clusters, Pools, Jobs, Tables etc.) ■ Defining log delivery at the cluster level (i.e., setting up storage for cluster logs, ideally through Cluster Policies ■ For more, review the introduction to Databricks administrator privileges and responsibilities. Further roles feature-specific admin roles, which have narrower sets of privileges, include: Databricks Databricks Marketplace admins: ■ Manage their account's Databricks Marketplace provider profile, including creating and managing Databricks Marketplace listings. Metastore admins: ■ Manage privileges and ownership for all securable objects within a Unity Catalog metastore, such as who can create catalogs or query a table. https://www.databricks.com/blog/2022/08/26/databricks-workspace-administration-best-practices-for-account-workspace-and-metastore-admins.html Granular Permissions: ■ Databricks allows customer organizations to define granular permissions to catalogs, schemas, tables, and columns, which provides necessary control over who may access data assets. AWS | Azure | GCP

Access Controls	Authentication Mechanisms: Implement and enforce robust authentication methods, such as multi-factor authentication (MFA), to access your Databricks Workspaces.	Databricks provides security features, such as single sign-on, to configure strong authentication. Admins can configure these settings to help prevent account takeovers, where a user's credentials are compromised using phishing or brute force, giving an attacker access to all the data accessible from the environment.
		
		Single sign-on enables you to authenticate your users using your organization's identity provider. Databricks recommends configuring SSO for greater security and improved usability. Once SSO has been configured, you can enable fine-grained access control, such as multi-factor authentication, via your identity provider. Unified login allows you to manage one SSO configuration in your account, which is used for the account and Databricks workspaces. If an account has been created before June 21, 2023, you can manage SSO individually on your account and workspaces. See SSO in your Databricks account console and Set up SSO for your workspace.
		
		Multi-Factor Authentication (MFA)
		
		■ MFA event types are considered account events and logged at the workspace level here- https://docs.databricks.com/en/admin/account-settings/audit-logs.html
		■ Users can review these MFA event types:
		  ■ The user registers a new security key.
		  ■ The user deletes a security key.
		  ■ The user logs into Databricks using MFA.

Access Controls	Access Reviews: Perform regular reviews of user access rights for appropriateness and alignment with job responsibilities.	As part of data governance, Databricks customers are encouraged to leverage the Unity Catalog to perform ongoing reviews of user access roles and privileges in alignment with their business requirements.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE

Access Controls	Logging and Monitoring: Enable logging for access and use of Databricks resources to detect unauthorized access and potential security breaches.	Unity Catalog provides access to System Tables, which are a repository of analytical data for an account that offers insights into how the platform functions. System tables allow organizations to monitor the usage, performance, and behavior of the platform's components. Some examples of system tables include audit logs, table lineage, and predictive optimization.
		
		The benefits for System Tables include but are not limited to:
		
		■ Comprehensive Observability and Monitoring
		  ■ A centralized analytical repository for operational and historical data across the account enables customers to gain deep insights into the Data Intelligence Platform's status and health (resource usage, costs, performance).
		■ Enhanced Data Governance and Security
		  ■ Offers centralized access control and auditing features to organizations.
		  ■ Organizations can monitor data access, track usage patterns, and ensure compliance with their security policies.
		  ■ Customers may audit who accessed what data and when the activity may have occurred.
		
		Further, the Databricks Unity Catalog offers customers audit logs of activities performed by Databricks users to allow your enterprise to monitor detailed Databricks usage patterns.
		
		Audit logging is essential because it provides a detailed account of system activities (user actions, changes to settings, and so on) that could affect the system's integrity. While standard system logs help developers troubleshoot problems, audit logs provide a historical activity record for compliance and other business policy enforcement purposes. Maintaining robust audit logs can help identify and ensure preparedness for threats, breaches, fraud, and other system issues.
		
		Databricks provides access to audit logs of activities performed by Databricks users, allowing your customer organization to monitor detailed Databricks usage patterns.
		
		Databricks recommends customer organizations to:
		
		■ Set up comprehensive audit logging to track all data access and modifications. Customers must configure audit logs properly, ensure they capture all necessary details, and store them securely for compliance.
		■ Configure audit logs for Unity Catalog events to capture all data access and modifications.
		  ■ Databricks can deliver audit logs for your workspace into your cloud-specific bucket, including S3/ADLS/GCS. AWS | Azure | GCP
		  ■ Additionally, Databricks makes audit logs available via our System Tables: AWS | Azure | GCP
		■ Set up and configure log delivery. AWS | GCP
		■ Align and implement log retention requirements in your cloud storage bucket.
		  ■ If a customer has yet to select an external cloud storage option, Databricks will retain various audit logs for up to 365 days. AWS | Azure | GCP
		
		Databricks Unity Catalog table lineage is fed into system tables. It offers customers the following benefits to support better governance and visibility into customer workspaces. This includes but is not limited to:
		
		■ Enhanced Data Observability and Governance
		  ■ This is a comprehensive view of data lineage across the entire lakehouse environment. It further provides insights into data origins, transformations, and dependencies and improves understanding of how data flows through the system.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Change Management	Change Control Procedures: Establish formal change management policies and procedures. Ensure changes to Databricks configurations and code undergo formal change management, including review, approval, and adequate documentation.	Ensure your organization maintains a formal change and release management policy and procedure to address changes within your Databricks account. While audit logs are available to customers, Unity Catalog offers a data governance layer to support visibility into events in your workspace further. Databricks offers System Tables (including data lineage and audit log events) to govern changes in your Databricks environment. See more details here: https://docs.databricks.com/en/admin/system-tables/index.html
Change Management	Testing of All Changes: ensure all changes are tested in a non-production environment before deployment.	Databricks recommends customers maintain secure software development lifecycle controls to protect source code. Examples of this may include but are not limited to customers applying the following to their change/release management procedures: ■ Maintain a secure source code management system ■ Require authentication/authorization for access to source code management system upon role (RBAC) using SSO and MFA solution. ■ Separate environments in place between non-production and production environments ■ Code undergoes dynamic and static analysis scanning tools as part of a secure CI/CD pipeline. ■ Developed code undergoes security and peer review/approval before check-in. ■ Formal change/release management procedures govern code deployment from non-production to production environments. Databricks Unity Catalog enables users to provide segregation across their organization's information architecture. Customers may use specific catalogs for various use cases (development, staging, production, non-customer data, etc.) to correspond to a software development environment scope, team, or business unit. If you use workspaces as a data isolation tool—for example, different workspaces for production and development environments or a specific workspace for working with sensitive data. In that case, you can also bind a catalog to specific workspaces to ensure that all specified data processing is handled in the appropriate workspace.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Change Management	Maintain a Secure Version Management System to track notebooks, configurations, and data pipeline changes.	Databricks Git folders is a visual Git client and API in Databricks. It supports everyday Git operations such as cloning a repository, committing and pushing, pulling, branch management, and visual comparing diffs when committing. https://docs.databricks.com/en/repos/index.html Databricks customers may use delta Tables to track changes to data, including versioning and lineage. Unity Catalog customers may review relevant events using audit logs (system tables). https://docs.databricks.com/en/admin/system-tables/index.html Data Classification Customers are responsible for defining and implementing their organizational data classification strategy to understand all data types used across their internal and external environments (including Databricks). The Databricks' Unity Catalog enables users to describe and tag data assets to identify and categorize financial data assets clearly with a search interface to help consumers find data. Further, Databricks Lakehouse Monitoring also provides proactive alerts for quality issues and errors in data and ML model pipelines, including automatic classification and identification of personally identifiable information (PII) using AI-based data classification technology. Further please refer to the Databricks SOC 1 Type II Control(s) Complementary User Entity Controls #3, #4 and #8 covering data classification and security measures below: ■ 3) Customers are responsible for data classification and implementing encryption features available where deemed necessary by customer-defined requirements. Additionally, customers are responsible for choosing libraries to prevent the accidental transmission of sensitive data (including PHI) over the wire in an unencrypted fashion. ■ 4) Customers are responsible for securing and protecting account credentials and security tokens for REST API access. ■ 8) For Customer-managed storage, customers are responsible for backing up, securing access to, and encrypting customer data in the AWS S3 bucket, Azure Blob storage, or Google GCS bucket.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Data Management	Data Encryption: Implement data in transit and at rest encryption to protect organizational information.	Databricks provides encryption features to protect your data. For more information, see Databricks' data security and encryption section here. Data in Transit ■ All communications between the control plane and data plane use minimum TLS 1.2+ for encryption in transit. ■ Encryption of data in use with confidential computing Data at Rest ■ AES-256-bit encryption is in place to protect the control plane. ■ The data plane supports local encryption; customers can use encrypted storage buckets. ■ Customer-managed key encryption is available to support encryption for control plane data and external cloud storage. See the related documentation here: AWS, Azure, and GCP. Further, customers may refer to SOC 1 Type II Control(s) Complementary User Entity Controls #3, #4 and #8. See below: ■ 3.) Customers are responsible for data classification and implementing encryption features available where deemed necessary by customer-defined requirements. Additionally, customers are responsible for choosing libraries to prevent the accidental transmission of sensitive data (including PHI) over the wire in an unencrypted fashion. ■ 4.) Customers are responsible for securing and protecting account credentials and security tokens for REST API access. ■ 8.) For Customer-managed storage, customers are responsible for backing up, securing access to, and encrypting customer data in the AWS S3 bucket, Azure Blob storage, or Google GCS bucket.
Data Management	Backup and Recovery: develop and maintain data backup and recovery policy and procedures and apply them to your Databricks implementation to ensure business continuity in case of data loss or corruption.	Databricks customers may develop backup and recovery strategies to align with their organizational responsibilities. ■ Backup of your organization's account and workspace here: https://github.com/databrickslabs/databricks-sync ■ Set Recovery Point Objectives (RPO) and Recovery Time Objectives (RTO) using best practices here: https://www.databricks.com/blog/2022/04/25/disaster-recovery-overview-strategies-and-assessment.html ■ Perform a Disaster Recovery Impact Assessment here: https://www.databricks.com/wp-content/uploads/2022/04/disaster-recovery-impact-assesment.pdf ■ Deploy Databricks disaster recovery services for Databricks to meet the organization's DR requirements here: https://docs.databricks.com/en/admin/disaster-recovery.html Customers may refer to SOC 1 Type II Control(s) Complementary User Entity Control #8: ■ 8.) For Customer-managed storage, customers are responsible for backing up, securing access to, and encrypting customer data in the AWS S3 bucket, Azure Blob storage, or Google GCS bucket.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Incident Management	Incident Response Plan: Create and update an incident response plan to handle security incidents	Customers are encouraged to maintain their own organizational incident management policies and procedures. For more information, customers may refer to the Databricks SOC 1 Type II Control(s) Complementary User Entity Controls. ■ 7.) Customers are responsible for communicating issues and incidents related to security, availability, and confidentiality to the support personnel (Databricks and Microsoft Azure, as appropriate) through identified channels. Customers are to regularly inspect their assigned "Data Security Contact" to ensure that it remains current and up to date.
Incident Management	Detection and Reporting: Implement mechanisms to detect and report security incidents promptly.	Customers are encouraged to maintain their own organizational incident management policies and procedures. For more information, customers may refer to the Databricks SOC 1 Type II Control(s) Complementary User Entity Controls. ■ 6.) Customers are responsible for notifying Databricks of any unauthorized use of any password or account or any other known or suspected breach of security related to using the Lakehouse Platform. ■ 7.) Customers are responsible for communicating issues and incidents related to security, availability, and confidentiality to the support personnel (Databricks and Microsoft Azure, as appropriate) through identified channels. Please reference the following resources: https://www.databricks.com/company/contact
Incident Management	Incident Analysis: Conduct root cause analysis of incidents to prevent recurrence and improve response measures.	Customers are encouraged to maintain their own organizational incident management policies and procedures, including conducting root cause analysis of incidents to prevent recurrence and improve their response measures. Audit logs and system tables can be used for various purposes, from user activity, model serving events, and cost monitoring to audit logging. Databricks recommends that customers configure system tables and set up automated monitoring and alerting to meet their needs. The blog post Improve Lakehouse Security Monitoring Using System Tables in Databricks Unity Catalog is a good starting point to help customers get started. Customers using Enhanced Security Monitoring or the Compliance Security Profile can monitor and alert on suspicious activity detected by the behavior-based malware and file integrity monitoring agents.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Security Configuration	Platform Security: Review and update Databricks security configurations, including workspace settings, cluster policies, and network security, regularly.	Customers are encouraged to follow the Databricks Unity Catalog best practices documentation covering the configuration of access control and cluster configuration located here: https://docs.databricks.com/en/data-governance/unity-catalog/index.html# Customers are encouraged to proactively scan their Databricks workspaces using the Databricks Security Analysis Tool (SAT) tool. SAT enhances Databricks security by flagging best practice deviations and simplifying security health checks for workspaces. It conducts routine workspace scans to ensure persistent alignment with security best practices and boosts data sensitivity confidence. Leveraging Databricks' deep security expertise, SAT offers recommendations and documentation links for implementing robust security measures. *The Security Analysis Tool (SAT) is an experimental productivity tool and is not intended to be a certification for your deployments. The SAT project is regularly updated to improve the correctness of checks, add new checks, and fix bugs.
Security Configuration	Compliance Monitoring: Continuously monitor compliance with security policies and regulatory requirements.	Audit logs and system tables can be used for various purposes, from user activity, model serving events, and cost monitoring to audit logging. Databricks recommends that customers configure system tables and set up automated monitoring and alerting to meet their needs. The blog post Improve Lakehouse Security Monitoring Using System Tables in Databricks Unity Catalog is a good starting point to help customers get started. Customers using Enhanced Security Monitoring or the Compliance Security Profile can monitor and alert on suspicious activity detected by the behavior-based malware and file integrity monitoring agents. AWS | Azure | GCP
Security Configuration	Vulnerability Management: Regularly perform scans for vulnerabilities and apply patches or mitigation techniques promptly.	As a part of our Threat and Vulnerability Management program, we perform weekly scanning of the AMIs as they progress from dev to production. These scans are reviewed and plugged into our vulnerability management process, including creating tickets for resolution. How does the scan itself work? The Databricks SDLC naturally moves code from dev to stage and production; we use various tools to scan a host in the staging environment to identify vulnerabilities before capturing a machine image that would be released to customer environments. Vulnerability scans use authentication with all checks enabled. At least weekly, we can host and then identify and remediate any vulnerabilities per our Vulnerability Management SLA. Internally, we have automated vulnerability management to effectively track, prioritize, coordinate, and remediate vulnerabilities in our environment. We perform daily authenticated vulnerability scans of Databricks and third-party/open source packages used by Databricks, along with static and dynamic code analysis (SAST and DAST), using trusted security scanning tools before we promote new code or images to production. The product security team also triages critical vulnerabilities to assess their severity in the Databricks architecture. Databricks has a vulnerability response program that monitors emerging vulnerabilities before they are reported to us by our scanning vendors. We accomplish this using internal tools, social media, mailing lists, and threat intelligence sources (e.g., US-CERT and other government, industry, and open source feeds). Databricks monitors open vulnerability platforms, such as CVE Trends and Open CVDB. We have an established process for responding to these so we can quickly identify the impact on our company, product, or customers. This program allows us to reproduce quickly reported vulnerabilities and resolve zero-day vulnerabilities. As Databricks is a shared multi-tenant environment, we do not permit vulnerability scanners against our management infrastructure. Customers are responsible for restarting their cluster to receive the latest patch version.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE

System Operations	Monitoring and Logging: Implement comprehensive monitoring of Databricks systems to detect performance issues and anomalies.	Databricks Unity Catalog lets you easily access and query your account's operational data, including audit logs, billable usage, and lineage, using system tables (Public Preview). https://docs.databricks.com/en/data-governance/index.html
		
		Customers can access and review their account's audit logs in the Unity Catalog via system tables. For more details, please see the Databricks Audit Logs System Table Reference.
		
		Databricks recommends:
		■ It is essential to set up comprehensive audit logging to track all data access and modifications. Customers must configure audit logs properly, ensure they capture all necessary details, and store them securely for compliance.
		■ Configure audit logs for Unity Catalog events to capture all data access and modifications.
		  ■ Databricks can deliver audit logs for your workspace into an S3/ADLS/GCS bucket of your choosing: AWS | Azure | GCP
		  ■ Additionally, Databricks makes audit logs available via our System Tables: AWS | Azure | GCP
		■ Set up and configure log delivery. AWS | GCP
		■ Align and implement log retention requirements in your cloud storage bucket.
		  ■ If a customer has yet to select an external cloud storage option, Databricks will retain various audit logs for up to 365 days. AWS | Azure | GCP

System Operations	Operational Procedures: Develop standard operating procedures (SOPs) for routine operations and maintenance tasks.	Customers are encouraged to build and maintain their own organizational standard operating procedure documents that align with their information security policy and standards. These may include but are not limited to: ■ Change Management Standard ■ Cryptographic Algorithm Standard ■ Cryptographic Key Management Standard ■ Data Classification and Data Handling Standard ■ Data Retention / Records Management Standard ■ Logical Access Management Standard ■ Logging and Monitoring Standard ■ Physical Access Management Standard ■ Release Management Standard ■ Software Development Lifecycle Standard
System Operations	Capacity Management: Plan and manage system capacity to handle current and future workloads effectively.	Customers are encouraged to review and monitor the following: ■ Their cloud provider-specific resource limits. ■ Databricks Platform limits: review platform limits. ■ Unity Catalog limits: Unity Catalog AWS resource quotas | Azure resource limits | GCP Resource Limits

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE
Vendor Management	Third-Party Risk Management: Assess and manage risks with third-party services integrated with your Databricks account regularly.	Customers are encouraged to maintain and regularly review their own organizational third party risk management program. Please see the Databricks Security and Trust Center for related documentation here: https://www.databricks.com/trust/compliance
Vendor Management	Service Level Agreements (SLAs): Define and enforce SLAs with third-party vendors to ensure they meet security and performance requirements.	Databricks customers are encouraged to subscribe to availability notification details here: https://docs.databricks.com/en/resources/status.html and https://status.databricks.com. Databricks provides a variety of customer support plans (Business, Enhanced, Production, Mission Critical) to provide you with dedicated support and timely service for the Databricks Platform and Apache Spark™ that best align with your business needs. Please see tiering options located here- https://www.databricks.com/support
Training and Awareness	User Training: Regularly train users on Databricks security practices and compliance requirements.	Customers are encouraged to design regular security awareness and training curricula to educate users on the latest industry-related threats. For customer organizations, please see: ■ Security Best Practices Guide here ■ Cloud Service Provider Specific Security and Compliance Guides here: AWS | Azure | GCP
Training and Awareness	Awareness Programs: Conduct an ongoing security awareness program to educate employees and contractors about new threats and best practices.	Customers are encouraged to design their organizational security awareness and training curricula, continuously educating users on the latest industry-related threats. Databricks regularly publishes white papers, tools, videos, and blogs to help you harden your Databricks deployments using security best practices and maintain the security of your systems and data.
Documentation	Policy and Procedure Documentation: Maintain up-to-date documentation for all policies, procedures, and controls related to Databricks usage.	Customers are encouraged to build and maintain up-to-date documentation for information security policies, standards, and procedures that align with their security controls and cover cloud-based solution implementations such as Databricks. Databricks recommends customers to: ■ Regularly review the Databricks Security and Trust Center security and compliance documentation here ■ Regularly review Unity Catalog best practices here- AWS | Azure | GCP ■ To determine that you are following Databricks security best practices and monitor the overall security health for all your account workspaces, run the Security Analysis Tool (SAT)*. The Security Analysis Tool (SAT) analyzes customers' Databricks account and workspace security configurations and provides recommendations that help them follow Databricks security best practices. When a customer runs SAT, it will compare their workspace configurations against security best practices and deliver a report. For more information, see here: https://docs.databricks.com/en/security/index.html#security-analysis-tool *The Security Analysis Tool (SAT) is an experimental productivity tool and is not intended to be a certification for your deployments. The SAT project is regularly updated to improve the correctness of checks, add new checks, and fix bugs.

CONTROL FAMILY	CONTROL	DATABRICKS FEATURE MAPPING/GUIDANCE

Documentation	Audit Trails: Ensure all actions and changes within Databricks are well documented and traceable to support audit requirements.	Databricks Unity Catalog lets you easily access and query your account's operational data, including audit logs, billable usage, and lineage, using system tables (Public Preview). https://docs.databricks.com/en/data-governance/index.html
		
		Customers can access and review their account's audit logs in the Unity Catalog via system tables. For further details, please review the Datarick Audit Logs System Table Reference.
		
		Databricks recommends:
		■ It is essential to set up comprehensive audit logging to track all data access and modifications. Customers must configure audit logs properly, ensure they capture all necessary details, and store them securely for compliance.
		■ Configure audit logs for Unity Catalog events to capture all data access and modifications.
		  ■ Databricks can deliver audit logs for your workspace into an S3/ADLS/GCS bucket. AWS | Azure | GCP
		  ■ Additionally, Databricks makes audit logs available via our System Tables: AWS | Azure | GCP
		■ Set up and configure log delivery. AWS | GCP
		■ Align and implement log retention requirements in your cloud storage bucket.
		  ■ If a customer has yet to select an external cloud storage option, Databricks will retain various audit logs for up to 365 days. AWS | Azure | GCP

Compliance and Audit	Compliance Assessment: Regularly assess compliance with relevant standards (e.g., ISO 27001/17/18, SOC 1/2, GDPR) and regulatory requirements.	Customers are encouraged to build and maintain up-to-date documentation for information security policies, standards, and procedures that align with their security controls and cover cloud-based solution implementations such as Databricks. Databricks recommends customers to: ■ Regularly review the Databricks Security and Trust Center Compliance Page for relevant third party audit reports here- https://www.databricks.com/trust/compliance ■ Regularly review Unity Catalog best practices here- AWS | Azure | GCP ■ To determine that you are following Databricks security best practices and monitor the overall security health for all your account workspaces, run the Security Analysis Tool (SAT)*. The Security Analysis Tool (SAT) analyzes customers' Databricks account and workspace security configurations and provides recommendations that help them follow Databricks security best practices. When a customer runs SAT, it will compare their workspace configurations against security best practices and deliver a report. For more information, see here- https://docs.databricks.com/en/security/index.html#security-analysis-tool *The Security Analysis Tool (SAT) is an experimental productivity tool and is not intended to be a certification for your deployments. The SAT project is regularly updated to improve the correctness of checks, add new checks, and fix bugs.
Compliance and Audit	Audit Readiness: Prepare for internal and external audits by ensuring that all controls are in place and properly documented.	