Service Name

Amazon Security Lake

Service Overview

Amazon Security Lake is a fully managed security data lake service that automatically centralizes security data from AWS environments, SaaS providers, on-premises, and third-party sources into a purpose-built data lake stored in your AWS account. The service analyzes security data to provide a comprehensive understanding of security posture across the entire organization while improving protection of workloads, applications, and data. Security Lake automates collection of security-related log and event data, manages data lifecycle with customizable retention and replication settings, and converts ingested data into Apache Parquet format using the Open Cybersecurity Schema Framework (OCSF) standard.

Key Use Cases

- Centralized security data analysis: Aggregate and analyze security logs from multiple AWS services, SaaS applications, and on-premises systems for comprehensive threat detection
- Incident response and forensics: Provide security teams with normalized data in OCSF format for faster investigation and response to security events across multi-cloud environments
- Security analytics and machine learning: Enable advanced analytics using Amazon Athena, Amazon QuickSight, or third-party tools to identify patterns and anomalies in security data

Features

- **Data Aggregation**: Creates purpose-built security data lake in customer account with automated collection from cloud, on-premises, and custom sources
- **OCSF Normalization**: Transforms data from various sources to Open Cybersecurity Schema Framework standard for consistent analysis
- **Multi-Account Management**: Supports multi-account and multi-Region data management with rollup regions for centralized analysis
- **Subscriber Access Control**: Provides multiple levels of access for subscribers with data access and query access options
- **Lifecycle Management**: Manages data lifecycle with customizable retention and replication settings using S3 storage classes

Value Proposition

Amazon Security Lake provides unique value by offering a fully managed security data lake that eliminates the complexity of building custom security analytics infrastructure. Unlike traditional SIEM solutions, Security Lake stores data in customer-owned S3 buckets while providing automatic normalization to OCSF standard, enabling seamless integration with existing security tools:

- The service reduces operational overhead through automated data collection, transformation, and lifecycle management while maintaining data ownership
- Key differentiators include native AWS service integration, support for multi-account organizations, and ability to work with both AWS and third-party security analytics tools without vendor lock-in

Service Integration

Security Lake deeply integrates with core AWS services for comprehensive functionality. Primary integrations include Amazon S3 for data storage, AWS Glue for data cataloging and crawling, AWS Lake Formation for access control, and Amazon EventBridge for event notifications:

- The service collects data from AWS CloudTrail, Amazon VPC Flow Logs, AWS Security Hub, Amazon EKS, Route 53 resolver logs, and AWS WAF
- Subscriber services include Amazon Athena for querying, Amazon OpenSearch Service for analytics, Amazon SageMaker for ML insights, and Amazon Detective for investigations
- Additional integrations with AWS IAM, AWS KMS, Amazon SQS, and AWS Lambda provide security, encryption, messaging, and serverless processing capabilities

Onboarding and Offboarding

Customers can get started with Security Lake through the AWS Management Console with streamlined setup that automatically creates necessary IAM roles, or programmatically using APIs/CLI requiring manual IAM role creation. Prerequisites include an AWS account and administrative user access:

- The service offers a 15-day free trial in supported regions
- Onboarding involves enabling Security Lake in desired regions, configuring data sources from AWS services or custom sources, and setting up subscribers for data consumption
- For organizations, customers can use AWS Organizations integration with delegated administrator capabilities
- The setup process automatically provisions S3 buckets, Glue databases, and Lake Formation tables required for operation

Getting Started Guide: https://docs.aws.amazon.com/security-lake/latest/userguide/getting-started.html

Data Removal

Offboarding involves disabling Security Lake through console, API, or CLI, which stops log collection while retaining existing settings, resources, and data. For custom sources, integrations must be disabled outside Security Lake console. Data removal requires manual deletion of S3 buckets and associated resources. When re-enabling after disabling, a new S3 bucket is created requiring manual data migration from old bucket using S3 Sync command. Complete cleanup includes deleting Glue databases, Lake Formation resources, and IAM roles created during setup.

Backup/Restore and Disaster Recovery

Security Lake leverages AWS global infrastructure with high availability across Regions and Availability Zones. The service uses Amazon S3 for data storage with built-in durability and resilience features including lifecycle configurations, versioning, and multiple storage classes. S3 provides automatic backup capabilities with cross-region replication options. However, Security Lake itself does not provide specific backup/restore functionality beyond standard S3 capabilities. Data resiliency depends on underlying S3 features and customer-configured lifecycle policies for archiving and retention management.

Backup Capabilities

Security Lake does not have native backup capabilities but relies on Amazon S3 underlying infrastructure for data durability and availability. The service does not integrate with AWS Backup service:

- Data protection is achieved through S3 features including versioning, lifecycle management, and storage class transitions
- Customers can configure S3 Cross-Region Replication for additional data protection and implement custom backup strategies using S3 capabilities and lifecycle policies for long-term archival to Glacier storage classes

Disaster Recovery

Security Lake does not directly support AWS Elastic Disaster Recovery integration. Disaster recovery capabilities depend on underlying Amazon S3 infrastructure and multi-region deployment architecture:

- The service supports rollup regions for centralizing data from multiple contributing regions, providing some disaster recovery benefits
- Recovery relies on S3 cross-region replication, multi-region data distribution, and the inherent high availability of AWS infrastructure rather than specific disaster recovery tooling

Recovery Objectives

Security Lake helps meet RPO and RTO objectives through its multi-region architecture and S3-backed storage durability. The service enables low RPO through continuous data ingestion and S3's 99.999999999% durability. RTO improvements come from rollup region capabilities allowing quick access to centralized data and the ability to query data using multiple analytics tools like Athena. However, specific RPO/RTO metrics depend on customer configuration of data replication, retention policies, and subscriber access patterns rather than service guarantees.

Service Constraints

Service Quotas: https://docs.aws.amazon.com/general/latest/gr/security-lake.html

FAQ: https://aws.amazon.com/security-lake/faqs/

Technical Requirements

Service Documentation: https://docs.aws.amazon.com/security-lake/latest/userguide/what-is-security-lake.html

User Guide: https://docs.aws.amazon.com/security-lake/latest/userguide/

API Reference: https://docs.aws.amazon.com/security-lake/latest/APIReference/

Pricing Overview

Please see the AWS UK G Cloud 15 Pricing Document accompanying this service in the Digital Marketplace.

Public Pricing: https://aws.amazon.com/security-lake/pricing/

Cost Optimization

Security Lake offers several unique cost optimization options including a 15-day free trial, data lifecycle management with S3 storage class transitions to reduce long-term storage costs, and no charges for bringing third-party or custom data to the service. Customers can optimize costs by configuring appropriate data retention periods, leveraging S3 Intelligent Tiering for automatic cost optimization, and using lifecycle policies to transition data to cheaper storage classes like Glacier:

- The service pricing is based on data ingestion and normalization volumes, allowing customers to control costs by managing data source selection and retention policies
- Regional data processing helps minimize data transfer costs

Savings Considerations

Main cost savings come from eliminating the need to build and maintain custom security data infrastructure, reducing operational overhead through managed service benefits, and avoiding vendor lock-in with data stored in customer-owned S3 buckets. The service eliminates costs associated with separate SIEM licensing and infrastructure while providing automatic data normalization and transformation:

- Customers save on data engineering resources typically required for security data pipeline management
- The OCSF standardization reduces integration costs with multiple security tools
- Usage-based pricing model allows customers to pay only for data processed, and the ability to query data using existing AWS analytics services like Athena provides additional cost efficiencies compared to proprietary analytics platforms