ClearDATA
SECURE • HEALTHCARE • CLOUD

Automated Safeguards for Redshift

Amazon Redshift is a fast, fully managed data warehouse that makes it simple and cost-effective to analyze all your data using standard SQL and your existing Business Intelligence (BI) tools. It allows you to run complex analytic queries against petabytes of structured data, using sophisticated query optimization, columnar storage on high-performance local disks, and massively parallel query execution. Most results come back in seconds. Additionally, Amazon Redshift Spectrum enables you to run SQL queries directly against all of your data, out to exabytes, in Amazon S3 - you simply pay for the number of bytes scanned.

Amazon Redshift is priced hourly based on the type and number of nodes in your cluster.

Amazon Redshift Spectrum bills you at $5 per terabyte of data scanned, rounded up to the next megabyte, with a 10 megabyte minimum per query. For example, if you scan 10 gigabytes of data, you will be charged $0.05. If you scan 1 terabyte of data, you will be charged $5.

Backup Storage

Increasing your backup retention period or taking additional snapshots increases the backup storage consumed by your data warehouse. There is no additional charge for backup storage up to 100% of your provisioned storage for an active data warehouse cluster. For example, if you have an active single XL node cluster with 2TB of storage, we will provide up to 2TB-Month of backup storage at no additional charge. Backup storage beyond the provisioned storage size and backups stored after your cluster is terminated are billed at standard Amazon S3 rates.

Data transfer

There is no charge for data transferred between Amazon Redshift and Amazon S3 within the same AWS Region for backup, restore, load, and unload operations. For all other data transfers into and out of Amazon Redshift, you will be billed at standard AWS data transfer rates. In particular, if you run your Amazon Redshift cluster in Amazon VPC, you will see standard AWS data transfer charges for data transfers over JDBC/ODBC to your Amazon Redshift cluster endpoint.

See the Amazon Redshift pricing page for details.

Amazon Redshift is a petabyte scale data warehouse service that can provide exceptional performance at scale with a fully managed database. Redshift also has a supported feature called Spectrum that allows queries to be run against data that resides in Amazon S3 without loading it directly into a database.

Redshift also features powerful and flexible encryption and permission controls that can further secure the data in the data warehouse. See Amazon Redshift Security Overview for details.

For a full and updated list of Amazon Redshift features please visit Amazon Redshift Features and Amazon Redshift Resources for a full list of documentation.

ClearDATA's Automated Safeguards for Redshift ensure that each database cluster is properly configured to meet ClearDATA's defined controls required to host and process PHI. ClearDATA reviews newly created database clusters to ensure the cluster storage is encrypted, the Parameter Group enforces an encrypted database connection and necessary audit logging, backups are enabled, and the Redshift cluster is not publicly available. If these settings are not properly configured, ClearDATA will modify the settings to meet the controls listed below.

> During this modification period the Redshift cluster will not be accessible until the modifications are complete. See the Compliance Guidance below for information on the impact each Automated Safeguard check has on the cluster.

> It may take as long as 30 minutes for a non-compliant Redshift cluster to become fully compliant and available. This is due to the time it takes for the cluster to both become initially available, and available again after the Automated Safeguard remediation has occurred. Customers can use the AWS Config console and the `cleardata-configrule-AllResources` rule to view the compliance of the cluster.

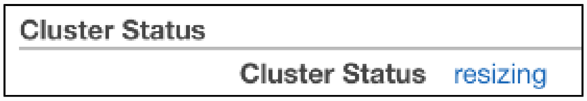
Compliance Guidance

Encrypted Storage

HIPAA Technical Safeguard 45 CFR 164.312(a)(2)(iv) requires encryption and decryption addressable standard strongly suggests that you implement a mechanism to encrypt and decrypt electronic protected health information (ePHI). ClearDATA's interpretation of this regulation is that all storage must be encrypted. ClearDATA Automated Safeguards will enable encryption at rest if it is not selected when a cluster is created. Amazon Redshift provides database encryption for its clusters to help protect data at rest. When customers enable encryption for a cluster, Amazon Redshift encrypts all data, including backups, by using hardware-accelerated Advanced Encryption Standard (AES)-256 symmetric keys. Amazon Redshift uses a four tier, key-based architecture for encryption. These keys consist of data encryption keys, a database key, a cluster key, and a master key. The cluster key encrypts the database key for the Amazon Redshift cluster.

Remediation

If the cluster is not encrypted during creation the ClearDATA Automated Safeguard will configure encryption at rest.

> This process is performed while the database is online, but it can take a significant amount of time for the encryption to be enabled. Enabling encryption at rest while the database is online is done by the Redshift service adding a second node and copying the database to a newly encrypted node. During this time the database will be in read-only mode. The cluster status will be listed as `resizing`, as shown below, until the encryption process is complete.
>
> 
>
> Cluster Status
>
> Cluster Status resizing
>
> **ClearDATA strongly recommends that customers enable encryption during the creation process to avoid any delay.**

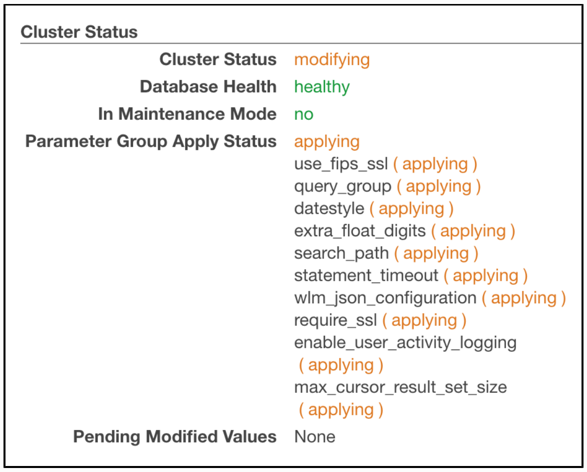
Encrypted Connections

HIPAA Technical Safeguard 45 CFR 164.312(a)(2)(iv) requires encryption and decryption addressable standard strongly suggests that you implement a mechanism to encrypt and decrypt electronic protected health information (ePHI). ClearDATA's interpretation of this regulation is that all connections to the Redshift databases must use a TLS encrypted connection, ensuring the data transmitted from client to and from database is encrypted. Redshift uses a Parameter Group setting called require_ssl to require that all connections are encrypted.

Remediation

If a Parameter Group that settings require_ssl to *true* is not assigned to the Redshift cluster, ClearDATA's Automated Safeguards will create a new Parameter Group and assign it to the cluster. During the time that the Parameter Group is syncing with the cluster, the database will be in Read Only mode.

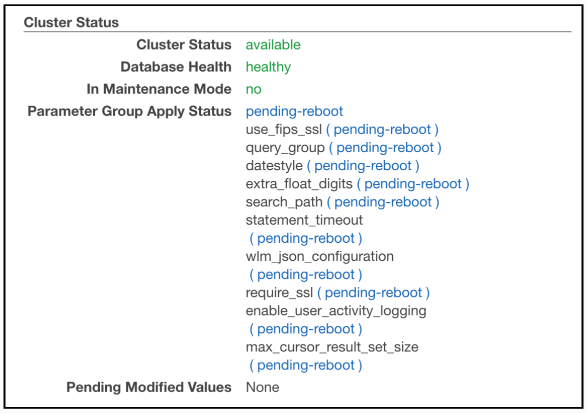
The assignment of a new Parameter Group, either by the ClearDATA Automated Safeguards or any other method, requires a reboot of the cluster to take effect. If necessary, the ClearDATA Automated Safeguards will apply the Parameter Group shortly after the cluster is online. During that time the cluster status will be `applying`, as seen below



Cluster Status

Cluster Status	modifying
Database Health	healthy
In Maintenance Mode	no
Parameter Group Apply Status	applying use_fips_ssl (applying) query_group (applying) datestyle (applying) extra_float_digits (applying) search_path (applying) statement_timeout (applying) wlm_json_configuration (applying) require_ssl (applying) enable_user_activity_logging (applying) max_cursor_result_set_size (applying)
Pending Modified Values	None

Once the Parameter Group has been applied to the cluster it will require a reboot. ClearDATA will initiate the cluster reboot shortly after the Parameter Group is applied. The cluster status will be listed as `pending-reboot` as seen below



Cluster Status

Cluster Status	available
Database Health	healthy
In Maintenance Mode	no
Parameter Group Apply Status	pending-reboot use_fips_ssl (pending-reboot) query_group (pending-reboot) datestyle (pending-reboot) extra_float_digits (pending-reboot) search_path (pending-reboot) statement_timeout (pending-reboot) wlm_json_configuration (pending-reboot) require_ssl (pending-reboot) enable_user_activity_logging (pending-reboot) max_cursor_result_set_size (pending-reboot)
Pending Modified Values	None

ClearDATA strongly recommends that customers configure the Parameter Group during the creation of the cluster to avoid any delay.

Automated Backups

HIPAA Technical Safeguard 45 CFR §164.308(a)(7)(ii)(B) requires a Covered Entity to "establish (and implement as needed) procedures to restore any loss of data." The automated backup feature of Amazon Redshift enables point-in-time recovery of your DB instance. Amazon Redshift replicates all your data within your data warehouse cluster when it is loaded and also continuously backs up your data to S3. Amazon Redshift always attempts to maintain at least three copies of your data (the original and replica on the compute nodes and a backup in Amazon S3). Customers can also enable Redshift to asynchronously replicate snapshots to S3 in another region for disaster recovery. By default, Amazon Redshift enables automated backups of your data warehouse cluster with a 1-day retention period and the ClearDATA Automated Safeguards will set that retention period to 14 days in order to meet our controls.

Remediation

ClearDATA automatically, and transparently, enables backups for all Redshift clusters. There is no interruption to the cluster access to enable backups.

Audit Logging

HIPAA Technical Safeguard 45 CFR § 164.312(b) requires a Covered Entity to "Implement hardware, software, and/or procedural mechanisms that record and examine activity in information systems that contain or use electronic protected health information." Most information systems provide some level of audit controls with a reporting method, such as audit reports. These controls are useful for recording and examining information system activity, especially when determining if a security violation occurred. ClearDATA recommends our customers utilize auditing in the application that will record the activity against PHI records. Database audit logs are available for all Redshift clusters are enabled by the ClearDATA Automated Safeguards. For more information on how to access the audit logs please see Database Audit Logging.

Remediation

ClearDATA automatically, and transparently, enables audit logs for all Redshift clusters. There is no interruption to the cluster access to enable audit logs.

Public Access

HIPAA Technical Safeguard 45 CFR § 164.312(e)(1) requires implementation technical security measures to guard against unauthorized access to electronic protected health information that is being transmitted over an electronic communications network. ClearDATA monitors provisioned Redshift clusters to ensure they are not publicly accessible, which could lead to unintended data access.

Remediation

ClearDATA will modify the publicly accessible settings to prevent the cluster from being publicly available.

> ⓘ Any Redshift cluster that has the Publicly Accessible setting enabled will require a reboot in order to remediate the setting. ClearDATA's Automated Safeguards will initiate the reboot shortly after the cluster is created.
>
> **ClearDATA strongly recommends that customers do not set the cluster to be publicly accessible during the creation of the cluster to avoid any delays.**

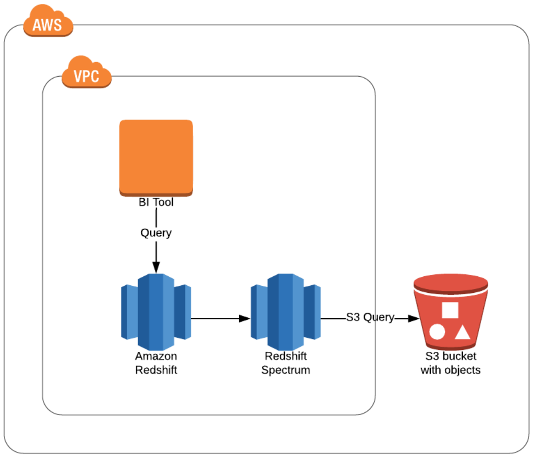
Shared Responsibility

ClearDATA will ensure that all Redshift clusters created in accounts with Automated Safeguards meet the requirements outlined above under Compliance Guidance. If a Redshift cluster is created and found to violate any of the items listed, the Automated Safeguards will remediate the cluster, during which any database will enter Read Only mode, and an alert will be sent to the ClearDATA SNS topic with details of the violation.

Please contact your ClearDATA team for a copy of the Responsibilities Matrix.

Exclusion

Disabling automated remediation at done the cluster level. Please contact ClearDATA Support to request that an exclusion be placed to allow for the instance to be created.



Users can be added to the **Safeguards-Redshift** group to gain access to create and manage RDS instances.

Item	ClearDATA	Customer
Enforcement of Automated Safeguards	RA	IC
Deployment of Redshift cluster	IC	RA
Database design, configuration, and administration	IC	RA
Database restore	IC	RA
Database cross region replication	IC	RA
All other Redshift and database tasks	IC	RA