IBM InfoSphere Guardium for Greenplum Database

Proactively address regulatory compliance requirements and protect sensitive data

Highlights

- Monitor and audit access to Greenplum Database data
- Build upon tested database monitoring technology
- Enforce separation of duties with a nonintrusive architecture
- Scale across the enterprise using a federated architecture
- Harden and secure your database using security best practices

The proliferation of data from endpoint devices, growing user volumes, and new computing models like cloud, social business and big data have created demands for data access and analytics that can effectively handle staggering amounts of data. Greenplum Database is one of the many new products to address the challenge of analytics on huge volumes of data.

Addressing Greenplum Database data security and protection challenges

Greenplum Database (DB) (which is based on the open source Postgres SQL) includes many of the typical built-in database features, such as role-based permissions and client authentication. However, audit and compliance requirements around the world require more robust accountability in terms of being able to log and verify who did what, and when for a database transaction. This information must be stored for a defined period of time, sometimes years. Relying on database logs is not a viable solution.

From an audit and compliance perspective, organizations still need to consider best practices such as:

- Continuous real-time monitoring to ensure data access is protected and audited.
- Policy-based controls based on access patterns to rapidly detect unauthorized or suspicious activity and alert key personnel.
- Protection of sensitive data repositories against new threats or other malicious activity.
- Demonstrated compliance to pass audits: It is not enough to develop a holistic approach to data security and privacy; organizations must also demonstrate and prove compliance to third-party auditors.

The hidden costs and security risks of custom security solutions

How are organizations handling the requirements for audit and compliance for Greenplum DB? It is likely that many organizations have not yet come to terms with the problem or are considering custom solutions based on aggregating and mining database log data. Custom solutions are problematic in many ways:

- Any approach that relies on log data does not comply with separation of duties (SOD) requirements as these can be tampered with by privileged Greenplum DB users.
- Real-time alerting is not supported; any compliance infraction or data breach could take weeks or months to discover using custom approaches.
- There are no capabilities for real-time prevention of data breaches, such as blocking or masking privileged user access to sensitive data.
- There is no automated way to return only the audit data required for compliance purposes, to distribute that data properly for review and signoffs, or to maintain the required audit trail of those signoffs.

Organizations would need to spend significant IT resources working around these issues; creating custom audit trails for compliance is not the best use of skilled IT resources.

Scalable enterprise-wide database security and compliance platform

IBM® InfoSphere® Guardium® has extended its market leading[1] data activity monitoring solution to include leading edge platforms such as Greenplum Database to help your organization meet compliance requirements while exploiting new innovations in data processing and analytics.

With its nonintrusive architecture (See Figure 1), InfoSphere Guardium provides full visibility into data activity and provides full separation of duties. This architecture requires no configuration changes to the Greenplum Database servers. Operating system software taps, called S-TAPs, are installed on the Master servers. The S-TAP unobtrusively streams the network packets to a hardened, tamper-proof hardware or software appliance known as a "collector" for parsing, analysis, and logging into its internal repository. Because processing of the network traffic occurs on the collector, overhead on the Greenplum DB cluster is very low.

The InfoSphere Guardium repository is the heart of the system and enables rich reporting, real-time alerting, and automated workflow management.

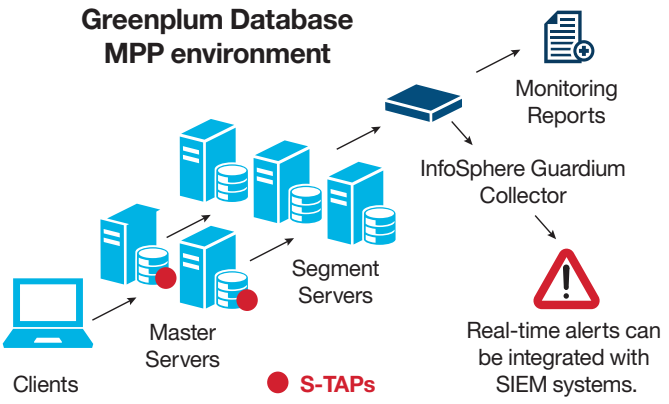

Figure 1: Architecture enforces separation of duties

Automated, policy-based monitoring and auditing streamline compliance validation

The InfoSphere Guardium web console provides centralized management of alerts, report definitions, compliance workflow processes, and settings (such as archiving schedules) without the involvement of Greenplum DBAs, thus providing the SOD required by auditors and streamlining compliance activities. A broad range of management functions can be executed across your entire database infrastructure, including:

- Defining granular security policies, using indicators of possible risk (appropriate for your particular environment), including the file or data object, type of access (reading, updating, deleting), user ID, source program, and more
- Defining actions in response to policy violations, such as generating alerts and logging full incident details (See Figure 2)

Access Rule Description	Client IP	Server IP	DB User Name	Full SQL String
Alert: Privileged User Access on sensitive data	9.70.145.179	9.70.145.179	GPADMIN (SUPER USER)	select * from creditcard;

Figure 2: Real-time alerts can be color coded based on severity level

- Blocking access to sensitive data from privileged users or hackers
- Automating compliance workflow for routine activities and incident responses, including steps such as sign-offs, commenting and escalation
- Ready-to-use reports for compliance and a rich customizable reporting capability (See Figure 3)
- Vulnerability assessment testing based on security standards best practices that can be scheduled to run once or on a regular basis to monitor progress over time

Server Type	DB User Name	Source Program	Full Sql	SQL Verb	Object Name
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	select * from creditcard ;	SELECT	creditcard
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	GRANT SELECT ON creditcard TO PUBLIC;	GRANT	creditcard
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	GRANT SELECT ON creditcard TO PUBLIC;	GRANT	PUBLIC
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	DROP EXTERNAL TABLE IF EXISTS mapreduce_4518_book CASCADE;	DROP EXTERNAL TABLE	mapreduce_4518_book
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	DROP FUNCTION IF EXISTS mapreduce_4518_wordsplit_python(text) CASCADE;	DROP FUNCTION	mapreduce_4518_wordsplit_python
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	DROP TYPE IF EXISTS mapreduce_4518_wordsplit_python_rtype CASCADE;	DROP TYPE	mapreduce_4518_wordsplit_python_rtype
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	DROP VIEW IF EXISTS book CASCADE;	DROP VIEW	book
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	DROP VIEW IF EXISTS mapreduce_4518_run_1 CASCADE;	DROP VIEW	mapreduce_4518_run_1
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	set gp_mapreduce_define=true	SET	gp_mapreduce_define
GREENPLUMDB	GPADMIN (SUPER USER)	GREENPLUM POSTGRESQL CLIENT PROGRAM	SELECT * FROM mapreduce_4518_run_1 ORDER BY key, value;	SELECT	mapreduce_4518_run_1

Figure 3: Monitoring reports show who, what, when, and where

With InfoSphere Guardium, you gain full visibility into Greenplum DB data activity, making it possible to identify unauthorized activities like data tampering or hacking, and address them in real time. Automation of the entire security and compliance lifecycle can help reduce labor costs, facilitate communication throughout the organization, and streamline audit preparation.

Why choose InfoSphere Guardium?

IBM InfoSphere Guardium provides the simplest, most robust solution for assuring the privacy and integrity of trusted information in your data center, and reduces costs by automating the entire compliance auditing process in heterogeneous environments.

Supported Greenplum Database releases and capabilities

Guardium capabilities	Greenplum DB 4.0, 4.1, 4.2*
Supports separation of duties	√
Activity monitoring, including detailed monitoring of super users or of any access to sensitive data	√
Integrate audit results with other monitored databases for enterprise wide reporting	√
Real time alerts and integration with SIEM systems	√
Policy-based security for consistency across heterogeneous environments	√
Ready-to-use and customizable reports	√
Blocking privileged user access to sensitive data	√
Federated architecture for scalability	√
Compliance workflow and automation to reduce total cost of ownership	√
Full set of administration APIs for automation and scripting	√
Vulnerability testing based on security best practices to help you harden your database against attacks and breaches	√

*InfoSphere Guardium also supports Greenplum HD activity monitoring. For an updated list of supported data platforms for monitoring, see ibm.com/support/docview.wss?&uid=swg27035836.

For more information

To learn more about IBM InfoSphere Guardium solutions, contact your IBM sales representative, IBM Business Partner® or visit: **ibm.com**/guardium.

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
June 2013

1 Gartner: "IBM InfoSphere Guardium is the market leader in terms of revenue and number of clients. Its offering has the widest platform coverage and the most robust set of features, and the company has demonstrated the ability to leverage the IBM sales model with its DAP offering."

IMD14441-USEN-00