I. SỬ DỤNG DATASUNRISE VỚI POSTGRESQL

A. Giới thiệu chung

1. Khái niệm

DataSunrise là một giải pháp bảo mật cơ sở dữ liệu toàn diện, được thiết kế để bảo vệ dữ liệu nhạy cảm, giám sát hoạt động truy cập, và hỗ trợ tuân thủ các tiêu chuẩn bảo mật quốc tế. Công cụ này hoạt động như một lớp bảo vệ bổ sung cho các hệ quản trị cơ sở dữ liệu, giúp tăng cường bảo mật mà không ảnh hưởng đến hiệu suất.

DataSunrise cung cấp các tính năng giám sát, phân tích và kiểm soát truy cập cơ sở dữ liệu, giúp ngăn chặn các rủi ro như tấn công SQL Injection, truy vấn trái phép và rò rỉ dữ liệu. Đồng thời, công cụ này cũng hỗ trợ phát hiện và quản lý dữ liệu nhạy cảm, giúp doanh nghiệp tuân thủ các quy định bảo mật nghiêm ngặt như GDPR, HIPAA, PCI DSS và nhiều tiêu chuẩn khác.

2. Tính năng

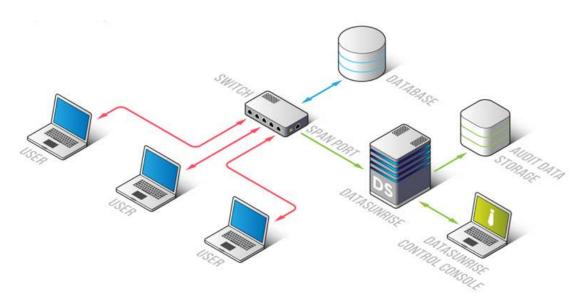
DataSunrise sở hữu nhiều tính năng mạnh mẽ để đảm bảo an toàn cho cơ sở dữ liệu:

- Kiểm toán (Audit):
 - Ghi lại tất cả các hành động của người dùng, truy vấn SQL và kết quả trả về từ cơ sở dữ liêu.
 - Hỗ trợ tích hợp với các hệ thống giám sát bên ngoài như SIEM để quản lý thông tin tập trung.
 - Theo dõi chi tiết các phiên làm việc của người dùng, bao gồm thời gian đăng nhập, truy vấn và đối tượng truy cập.
- Bảo mật thời gian thực (Security):
 - Phân tích lưu lượng cơ sở dữ liệu để phát hiện các hành vi đáng ngờ như truy vấn trái phép hoặc tấn công SQL Injection.
 - Ngăn chặn các truy vấn không hợp lệ trước khi chúng được thực hiện trên cơ sở dữ liệu.
 - Tự động gửi cảnh báo qua email hoặc các ứng dụng nhắn tin khi phát hiện mối đe dọa.
- Che giấu dữ liệu (Masking):
 - Dynamic Masking: Che giấu dữ liệu nhạy cảm trong kết quả truy vấn mà không làm thay đổi dữ liệu gốc trong cơ sở dữ liệu.
 - Static Masking: Tạo bản sao dữ liệu giả để sử dụng trong môi trường phát triển hoặc kiểm thử mà vẫn đảm bảo tính bảo mật.
- Khám phá dữ liệu nhạy cảm (Data Discovery):

- Xác định và phân loại dữ liệu nhạy cảm trong cơ sở dữ liệu dựa trên các từ khóa, kiểu dữ liệu hoặc tập từ vựng (lexicon).
- Sử dụng công nghệ NLP (Xử lý ngôn ngữ tự nhiên) để tìm kiếm dữ liệu nhạy cảm trong các cột dữ liệu không cấu trúc, như email hoặc số điện thoại trong văn bản.
- Liên kết các cột dữ liệu liên quan bằng tính năng Table Relations để tổ chức dữ liệu nhạy cảm một cách hiệu quả hơn.
- Tuân thủ bảo mật (Compliance):
 - Đáp ứng các yêu cầu bảo mật quốc tế như GDPR, HIPAA, PCI DSS, CCPA, ISO/IEC 27001 và nhiều tiêu chuẩn khác.
 - Tự động tìm kiếm và bảo vệ dữ liệu nhạy cảm mới được thêm vào cơ sở dữ liệu để duy trì tính bảo mật liên tục.
- Báo cáo (Reporting):
 - Tạo báo cáo chi tiết dưới định dạng PDF hoặc CSV về các sự kiện kiểm toán, bảo mật và khám phá dữ liệu nhạy cảm.
- 3. Các chế độ hoạt động

3.1 Sniffer Mode

Trong chế độ này, DataSunrise kết nối với cổng SPAN của một switch mạng, do đó, nó hoạt động như một công cụ phân tích lưu lượng mạng, có khả năng sao chép lưu lượng cơ sở dữ liệu từ cổng sao chép của switch mạng.



Chế độ Sniffer Mode

DataSunrise hoạt động ở chế độ này chủ yếu để bảo mật thụ động. Các tính năng bảo mật chủ động như tường lửa cơ sở dữ liệu hay che giấu dữ liệu không được hỗ trợ trong chế độ này.

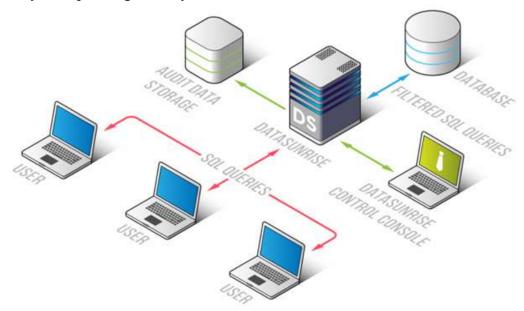
DataSunrise chỉ có thể thực hiện giám sát hoạt động cơ sở dữ liệu, vì không thể sửa đổi lưu lượng cơ sở dữ liệu trong cấu hình này

Chế độ Sniffer thích hợp để kiểm toán dữ liệu hoặc chạy DataSunrise trong chế độ học (Learning mode). Nó cũng hỗ trợ lưu lượng VXLAN cho AWS khi cấu hình chế độ sao chép (mirroring).

Lưu lượng cơ sở dữ liệu không nên được mã hóa, vì chế độ này không hỗ trợ kết nối mã hóa. Nếu sử dụng SQL Server, tránh sử dụng các cipher tạm thời (ephemeral ciphers). DataSunrise không hỗ trợ kết nối được chuyển hướng đến cổng ngẫu nhiên (ví dụ như Oracle).

3.2 Proxy Mode

Trong chế độ Proxy, DataSunrise hoạt động như một trung gian giữa máy chủ cơ sở dữ liệu và các ứng dụng khách. Công cụ sẽ xử lý tất cả các truy vấn đến trước khi chuyển tiếp chúng tới máy chủ cơ sở dữ liệu.



Chế độ Proxy Mode

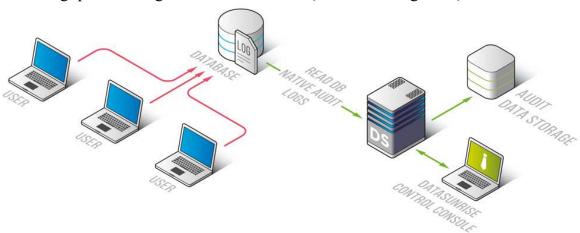
Đây là chế độ bảo vệ chủ động. DataSunrise sẽ chặn, kiểm tra, ghi nhận, hoặc sửa đổi các truy vấn SQL để đảm bảo chúng tuân thủ các chính sách bảo mật hiện tại.

DataSunrise cung cấp giám sát hoạt động cơ sở dữ liệu, tường lửa cơ sở dữ liệu, và che giấu dữ liệu (masking) động và tĩnh trong chế độ này. Độ trễ trung bình khi sử dụng chế độ Proxy dao động từ 3-10%.

Chế độ Proxy được khuyến nghị sử dụng vì nó cung cấp bảo vệ toàn diện và hỗ trợ xử lý lưu lượng mã hóa cũng như kết nối chuyển hướng (chẳng hạn như đối với Hana, Oracle, MS SQL).

3.3 Trailing DB audit logs

Chế độ này được sử dụng để kiểm toán các cơ sở dữ liệu như Oracle, Snowflake, Neo4J, PostgreSQL, MS SQL Server, MongoDB, và các cơ sở dữ liệu tương tự khác thông qua các công cụ kiểm toán nội bộ (native auditing tools).



Chế Độ Trailing DB Audit Logs

Cơ sở dữ liệu thực hiện kiểm toán qua các cơ chế tích hợp và lưu kết quả vào một bảng dữ liệu chuyên dụng hoặc tệp CSV/XML, sau đó DataSunrise tải và phân tích dữ liệu kiểm toán.

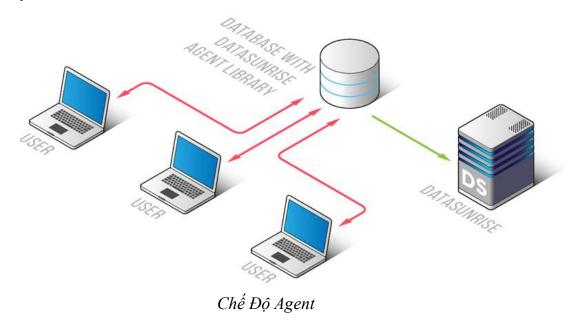
Chế độ này thích hợp khi cần ghi lại các hoạt động tại chỗ, ví dụ như hoạt động của quản trị viên cơ sở dữ liệu.

Một số hạn chế như Quản trị viên cơ sở dữ liệu có thể xóa các nhật ký kiểm toán nếu có quyền truy cập vào chúng , Kiểm toán nội bộ có thể ảnh hưởng đến hiệu suất của cơ sở dữ liêu.

Các cơ sở dữ liệu hỗ trợ: Amazon Aurora (MySQL và PostgreSQL), Amazon Redshift, MariaDB, MongoDB, MySQL, Oracle, PostgreSQL (EDB), và một số cơ sở dữ liêu khác.

3.4 Agent Mode

Trong chế độ này, DataSunrise sử dụng một tác nhân (Agent) để xử lý và phân tích các gói yêu cầu (request packets) và phản hồi (response packets) giữa máy khách và máy chủ cơ sở dữ liệu.

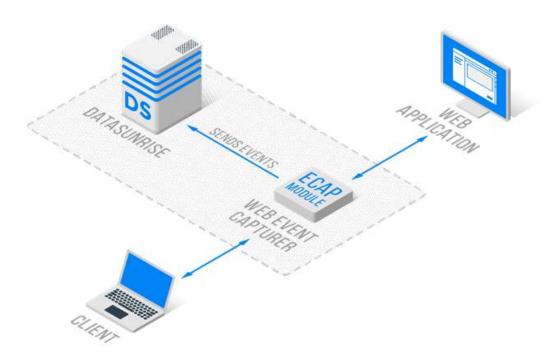


Agent là một thư viện động (DLL hoặc .so) được cài vào quá trình DBMS, cho phép chặn và phân tích các gói dữ liệu liên quan đến tương tác giữa máy khách và DBMS

Tác nhân sử dụng kết nối SSL với chứng chỉ xác thực hai chiều Chế độ này hiện hỗ trợ Oracle 11+ (64-bit) và PostgreSQL 12+ khi bật tính năng Giám sát Hoạt động.

3.5 Web Application Mode

Chế độ này được sử dụng để kiểm toán các ứng dụng web thông qua **Web Proxy** và các module đặc biệt như **squid-proxy** và **backend-proxy**.



Chế Độ Web Application

Các module như **squid-proxy** thu thập lưu lượng giữa máy khách và ứng dụng web, gửi chúng đến DataSunrise để phân tích.

Chế độ này hiện chỉ hỗ trợ **bảo mật thụ động** (passive security). Các tính năng như tường lửa cơ sở dữ liệu và che giấu dữ liệu không được hỗ trợ.

Chế độ Web Application hiện được hỗ trợ cho các ứng dụng như **Salesforce** và **ChatGPT**.

B. DEMO

- 1. Các yêu cầu chuẩn bị có trong bài
 - 1 máy Ubuntu có cài đặt Datasunrise, CSDL PostgreSQL, máy chủ web apache
 - 1 máy Windows có cài công cụ SQL Shell của PostgreSQL
- 2. Thực hiện
- 2.1 Cấu hình ban đầu

Trước hết, tạo 1 Database có tên "demo_login" để thuận tiện dùng cho máy chủ web và kết nối đến Datasunrise, sau đó tạo 1 bảng "users" chứa thông tin người

dùng. Kiểm tra trên máy chủ CSDL ta có:

```
b21dcat058@dat:~/Desktop$ psql -U postgres -d demo_login -h localhost -W
Password:
psql (12.20 (Ubuntu 12.20-0ubuntu0.20.04.1))
SSL connection (protocol: TLSv1.3, cipher: TLS_AES_256_GCM_SHA384, bits: 256, compression: off)
Type "help" for help.

demo_login=# \ \tau \
```

Thông tin về CSDL cần được quản lý

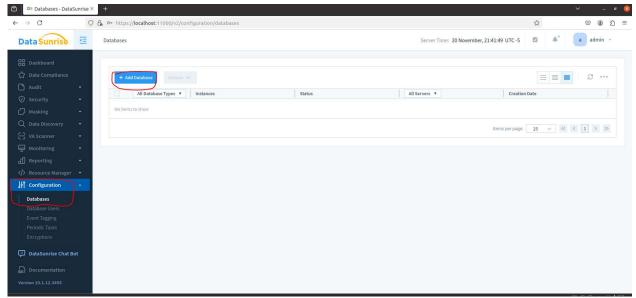
DataSunrise được cung cấp giao diện dựa trên web toàn diện (Bảng điều khiển web) được sử dụng để kiểm soát tất cả các chức năng của chương trình do đó sau khi cài đặt thành công , truy cập đến địa chỉ ở dưới để sử dụng :

https://<DataSunrise ip address>:11000

DataSunrise_ip_address: là ip máy chủ cài dặt giải pháp(IP máy ubuntu)
Sau đó đăng nhập với tài khoản Admin và mật khẩu mà bạn đã thiết lập trước đó.
Lần đầu tiên kết nối sẽ hiển thị các cài đặt kết nối có thể bỏ qua mà không ảnh hưởng gì.

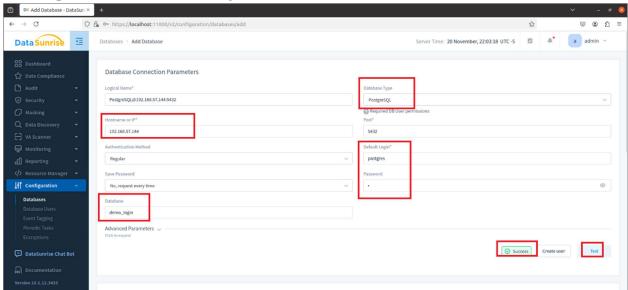
Sau khi đăng nhập , truy cập đến mục Configuration->Databases->Add Database

Để thiết lập kết nối tới cơ sở dữ liệu cần bảo vệ và giám sát.



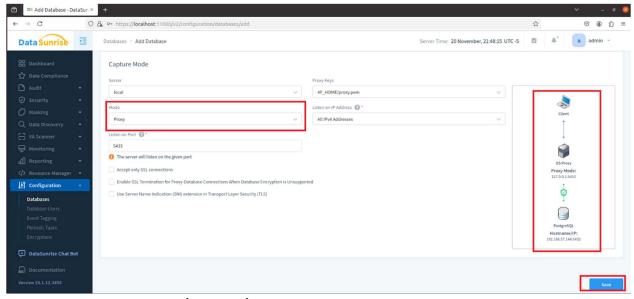
Truy cập đến chức năng thêm Database

Thiết lập và chỉnh sửa các trường dữ liệu đã được dánh dấu như ảnh dưới .



Thiết lập thông tin Database cần quản lý

Sau khi click Test và hiển thị Success, tiếp tục kéo xuống và cấu hình quản lý database bằng Proxy Mode như bên dưới đây:



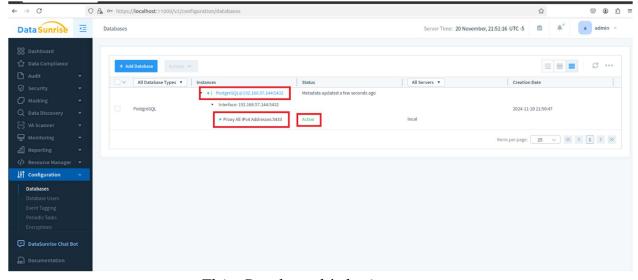
Thiết lập chế độ proxy của Datasunrise

Sau khi thêm thành công Database, có một số điều đáng chú ý như sau:

Cổng mặc định để truy cập PostgreSQL là 5432

Cổng để truy cập đến proxy quản lý là 5433

Khi thực hiện quản lý cơ sở dữ liệu, các kết nối đến máy chủ PostgreSQL cần phải để cổng Proxy là 5433 để thực hiện quản lý

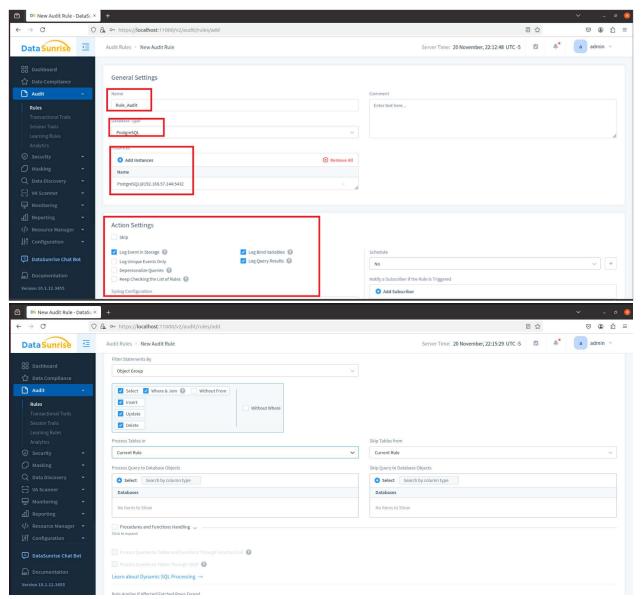


Thêm Database thành công

2.2 Chức năng Kiểm toán (Audit)

Trong chức năng này , ta cấu hình DataSunrise để kiểm tra tất cả các truy vấn được chuyển đến cơ sở dữ liệu đích.

Truy cập Audit -> Rules -> Add Rule



Cấu hình Rule audit

Sau khi chỉnh sửa các thông tin ta click Save rule .

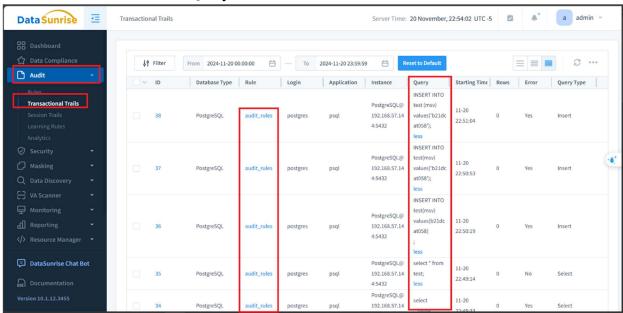
Trên máy windows sử dụng SQL Shell để kết nối đến máy chủ csdl từ bên ngoài và dùng port của proxy là 5433 . Sau đó sử dụng một số câu lệnh SQL như hình dưới:

```
Select SQL Shell (psql)
Schema | Name | Type | Owner
public | users | table | postgres
(1 row)
demo_login=# select * from users;
id | username | password
 1 | admin
                  admin123
 2 user
                  user123
(2 rows)
demo_login=# select * from x;
ERROR: relation "x" does not exist
LINE 1: select * from x;
demo_login=# create table test ;
ERROR: syntax error at or near
LINE 1: create table test;
demo_login=# create table test() ;
CREATE TABLE
demo_login=# select * from test;
(0 rows)
demo_login=# INSERT INTO test(msv)
demo_login-# values(b21dcat058)
demo_login-#;
ERROR: column "msv" of relation "test" does not exist
LINE 1: INSERT INTO test(msv)
demo_login=# INSERT INTO test(msv)
```

Kết nối đến CSDL từ máy windows

Tai Audit->Transactional Trails

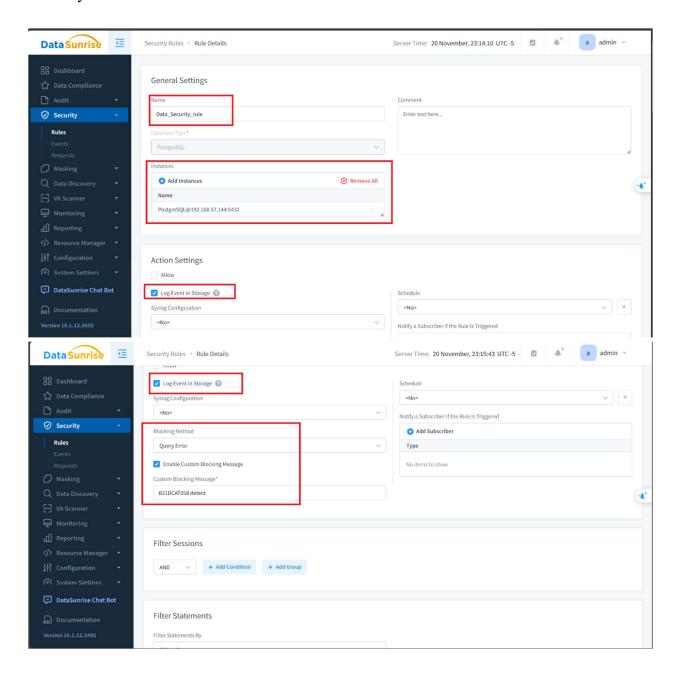
Ta thấy được các truy vấn tới máy chủ CSDL đã được bắt lại và hiển thị với tên bộ luật bắt được kèm các câu Query .

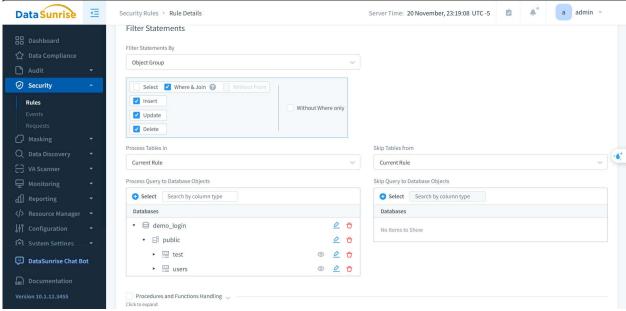


Lịch sử các câu truy vấn bắt được

2.3 Bảo mật thời gian thực (Security)

Với mong muốn ngăn chặn việc sửa đổi CSDL ta thực hiện thêm bộ luật sau tại Security -> rules -> Add Rule





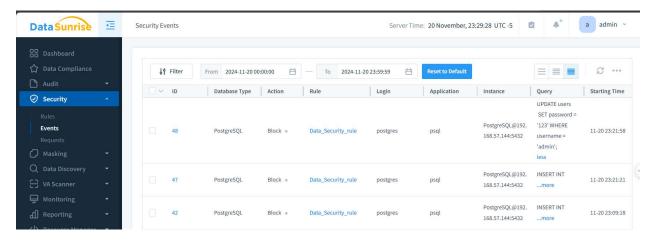
Cấu hình luật chặn các sửa đổi CSDL

Khi kích hoạt bộ luật, các câu SQL sửa đổi CSDL đã bị chặn hình dưới hiển thị chi tiết kèm theo dòng cảnh báo đã thiết lập sẵn "b21dcat058 detect"

```
SQL Shell (psql)
                                                                                                                    demo_login=# INSERT INTO users (username, password)
demo_login-# VALUES ('test_user', 'test123');
INSERT 0 1
demo_login=# UPDATE users
demo_login-# SET password = '1' WHERE username = 'admin';
UPDATE 1
demo_login=# DELETE FROM users
demo_login-# WHERE username = 'test_user';
DELETE 1
demo_login=# select * from users;
id | username | password
               user123
 2 user
 1 | admin
               1
(2 rows)
demo_login=# INSERT INTO users (username, password)
 demo_login-# VALUES ('test_user', 'test123');
RROR: B21DCAT058 detect
 demo_login=# UPDATE users
 demo_login-# SET password = '123' WHERE username = 'admin';
RROR: B21DCAT058 detect
 lemo_login=#
```

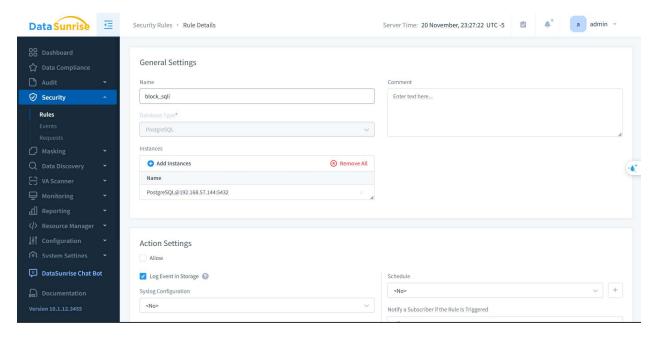
Phát hiện thao tác chỉnh sửa CSDL

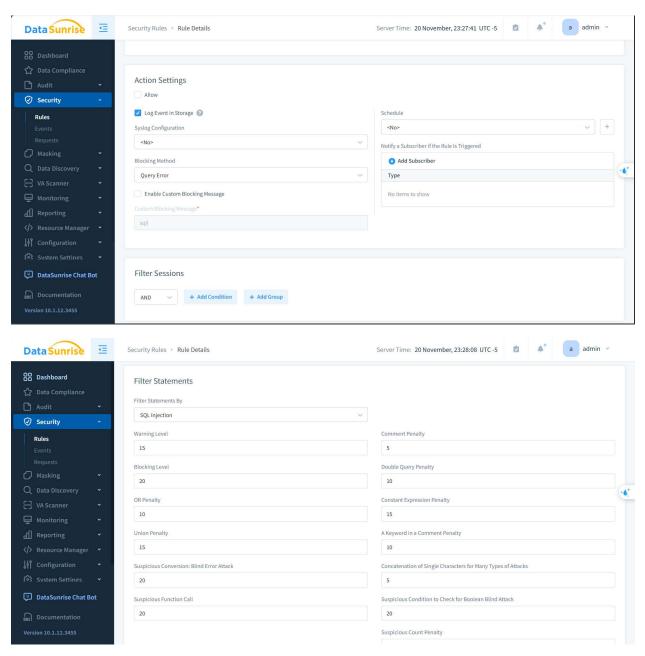
Ngoài ra ta cũng bắt được các thao tác chỉnh sửa được hiển thị cùng với bộ luật bắt được và chi tiết câu query



Lịch sử các câu query sửa đổi CSDL đã bị chặn

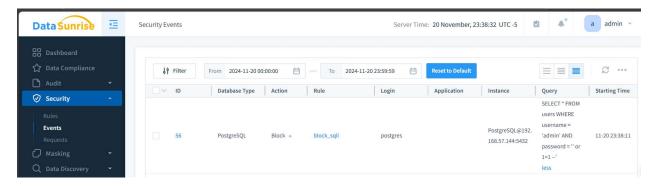
Bên cạnh đó với mong muốn chặn các truy vấn SQL injection độc hại từ web server ta cần tạo 1 trang web có chứa lỗ hồng SQLi trong chức năng đăng nhập sau đó thực hiện thêm bộ luật ngăn chặn SQLi sau:





Cấu hình luật chặn SQLi trên web server

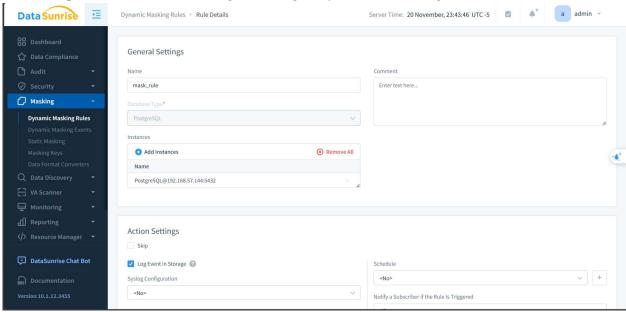
Tiến hành truy cập trang web và sử dụng payload SQLi để tấn công ta nhận thấy khi bật bộ luật lên các câu SQLi không còn tác dụng nữa mà đã bị cấm truy cập, bên cạnh đó ta cũng phát hiện ra lịch sử tấn công SQLi kèm theo bộ luật bắt được và câu truy vấn độc hại

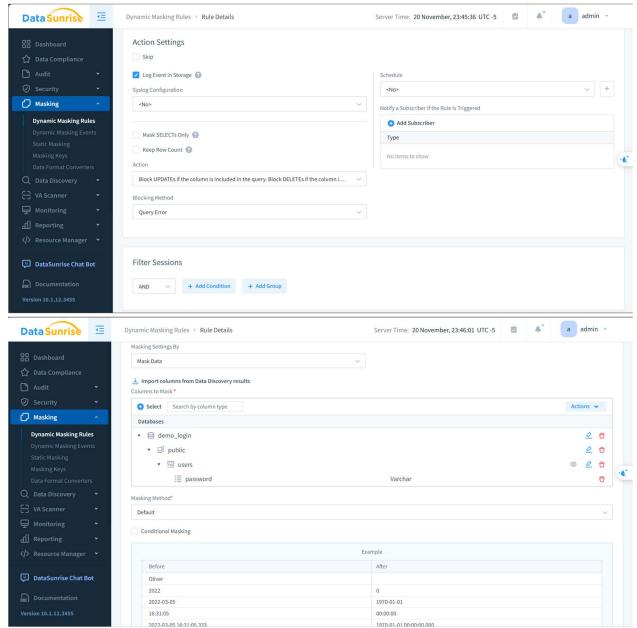


Phát hiện và chặn SQLi trên web server

2.4 Che giấu dữ liệu (Masking)

Với mong muốn ẩn hoặc xóa đầu ra của 1 trường dữ liệu khi có câu lệnh truy vấn , ta thiết lập thêm bộ luật mớitrong Masking->Dynamic Masking Rules





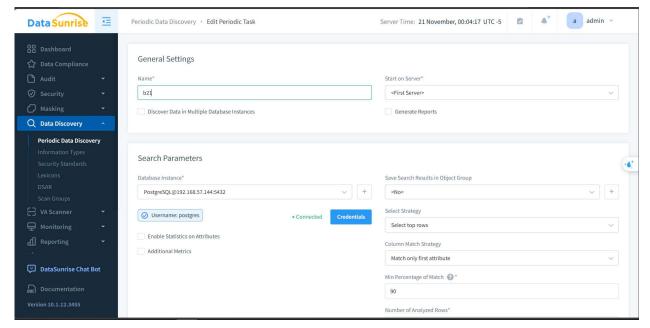
Luật che dấu dữ liệu

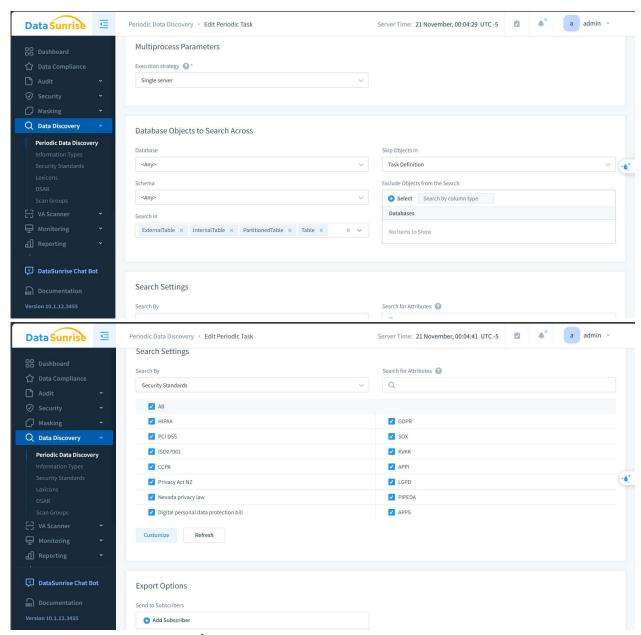
Sau khi thiết lập và chạy bộ luật, ta thực hiện câu truy vấn để xem thông tin bảng và thấy rằng thông tin trả về đã bị che đi.

Thông tin đã bị ẩn khi kích hoạt bộ luật

2.5 Khám phá dữ liệu nhạy cảm (Data Discovery)

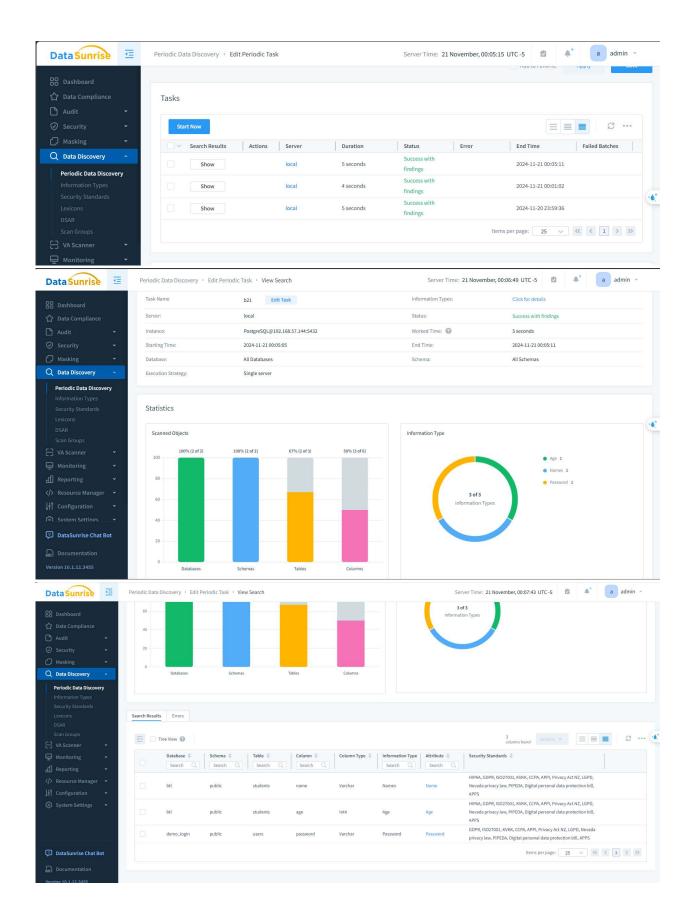
Với mong muốn tự động tìm kiếm và phân loại dữ liệu nhạy cảm trong cơ sở dữ liệu của mình, thiết lập cấu hình để tự động rà quét và phân loại các dữ liệu nhạy cảm sau:





Cấu hình cài đặt Data Discovery

Sau khi cấu hình thành công ta click startnow và show để xem các loại dữ liệu nhạy cảm có trên máy chủ CSDL



Phát hiện và phân loại các dữ liệu nhạy cảm