**TẬP ĐOÀN BƯU CHÍNH VIỄN THÔNG VIỆT NAM**

**CÔNG TY DỮ LIỆU VNPT – VNPT DATA**



**MẪU BÁO CÁO**

Hà Nội, 11/01/2016

Mục lục

[1. Giới thiệu chung 4](#_Toc440263742)

[2. Search 5](#_Toc440263743)

[3. Stream 8](#_Toc440263744)

[3.1. Tạo cảnh báo qua Slack hoặc Email với Stream 11](#_Toc440263745)

[4. Dashboard 18](#_Toc440263746)

[5. Sources 21](#_Toc440263747)

[6. System 21](#_Toc440263748)

[6.1. Input 22](#_Toc440263749)

[6.1.1. GELF TCP 22](#_Toc440263750)

[6.1.2. Syslog UDP 25](#_Toc440263751)

Danh mục các hình vẽ

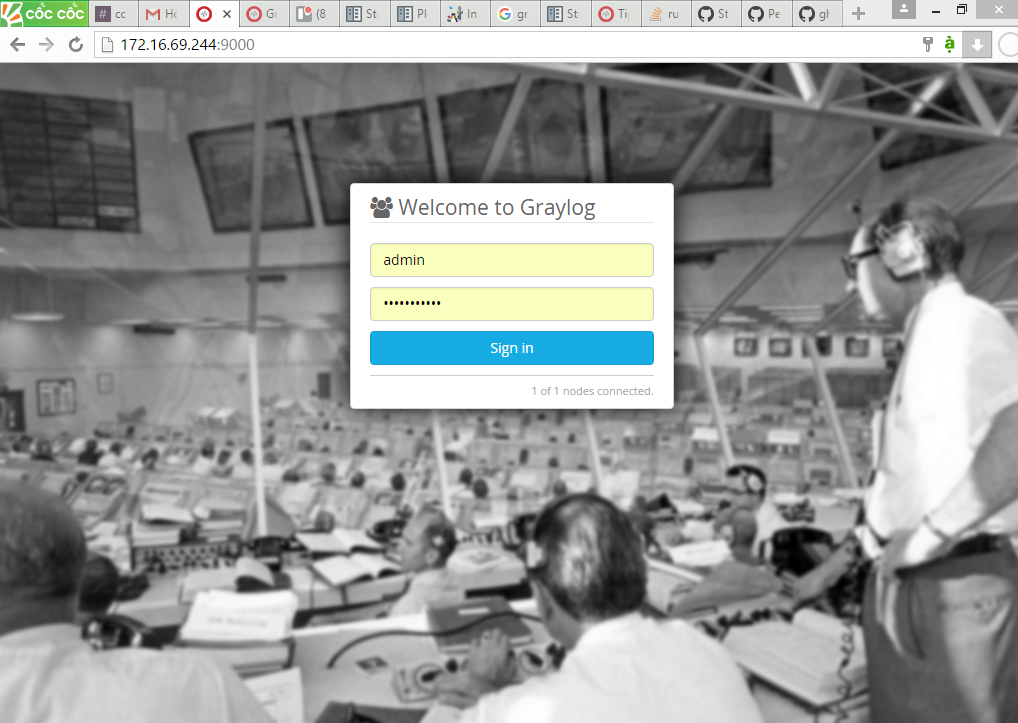
Danh mục các bảng biểu

No table of figures entries found.

# Giới thiệu chung

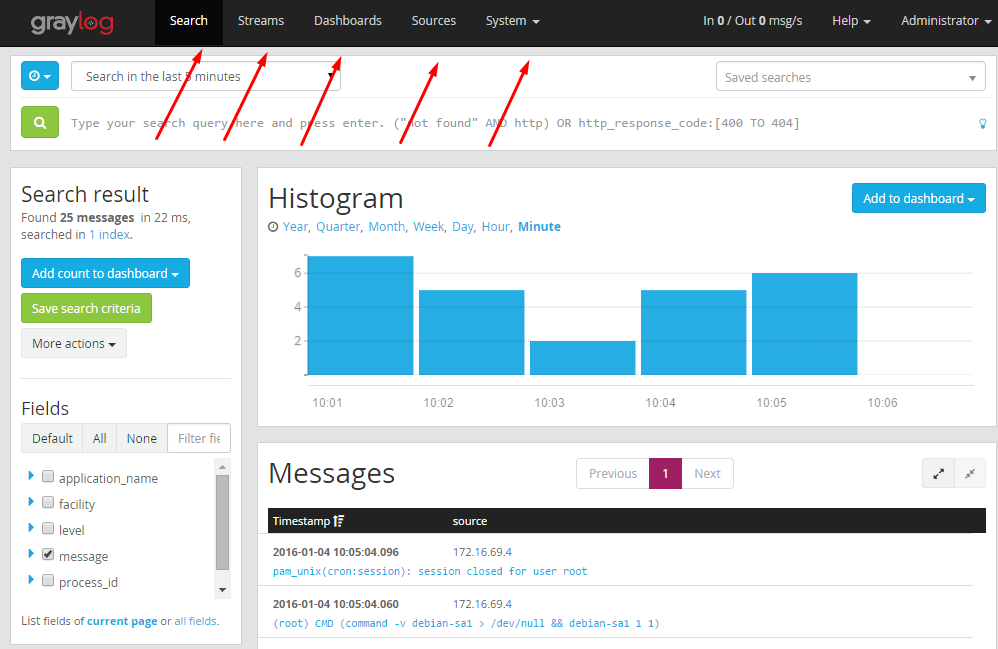
Đăng nhập Web-Interface trên Web-Browser với cú pháp :

<http://IP-GraylogServer:9000>



Trên dashboard của Web-Interface có 5 mục chính :

* Search
* Stream
* Dashboard
* Sources
* System



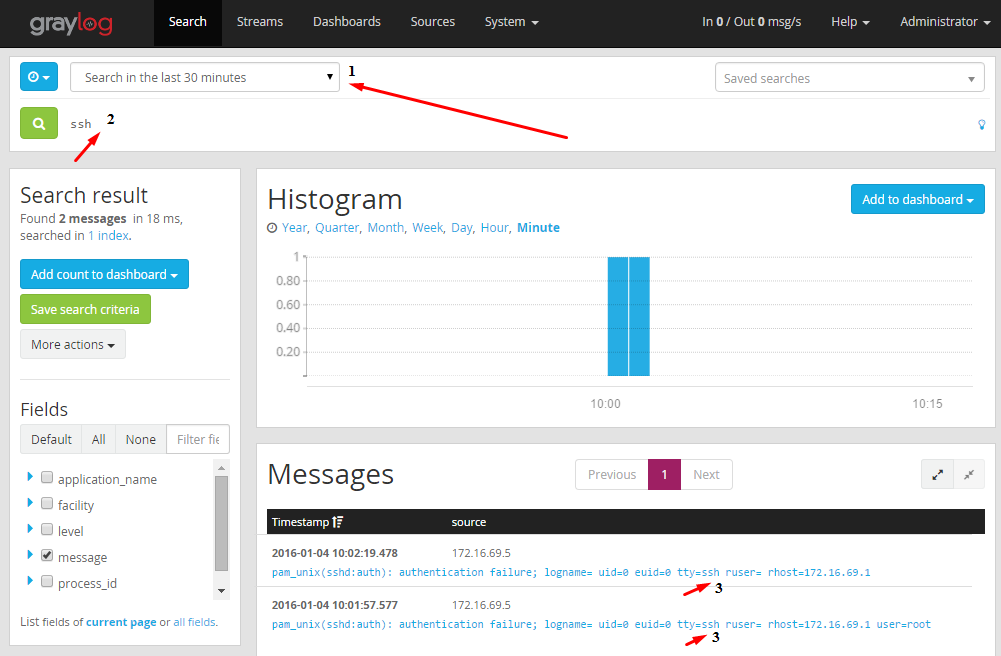
Tôi sẽ giới thiệu và hướng dẫn làm việc với từng content của Graylog. Kịch bản ở đây là : Chúng ta đã có sẵn 1 máy Graylog-Server hoàn chỉnh, và đã có sẵn những Input được đẩy về (Cách cài đặt Graylog-server đã được giới thiệu ở bài trước, và các kiểu cài đặt và cấu hình Input cụ thể sẽ được giới thiệu ở phần System => Input)

# Search

Graylog sử dụng Elasticsearch để lưu trữ log và message, vì vậy kỹ thuật searching trên Graylog cũng sử dụng kiểu search rất giống với Elasticsearch, đó là Lucene syntax. Ở đây tôi sẽ lấy một vài ví dụ về các kiểu search :

* Search các bản tin chứa ***ssh***  : ssh
* Search các bản tin chưa term ***ssh*** hoặc ***login*** : ssh login
* Search các bản tin chứa chính xác cụm từ ***ssh login*** : “ssh login”
* Search các bản tin mà trường *type* chứa ***ssh*** : type:ssh
* Search các bản tin mà trường *type* chứa ***ssh*** hoặc login : type:(ssh login)
* Search các biểu thức AND hoặc OR để kết hợp các trường cho việc tìm kiếm : “ssh login” AND 172.16.69.5

Ví dụ 1 : Search tất cả các bản tin chứa ***ssh*** :

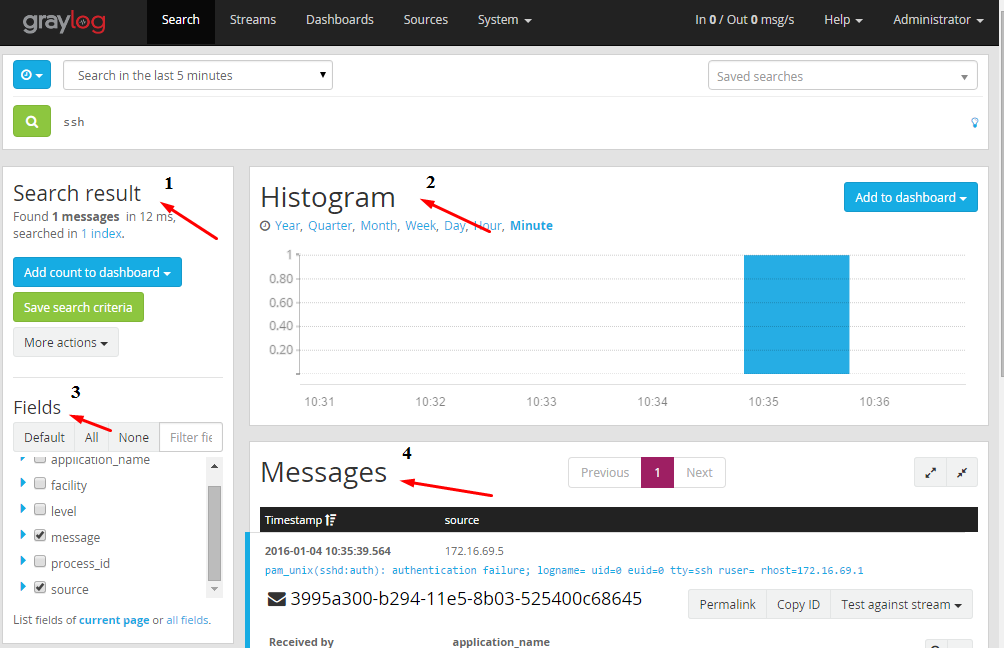


**1** : Tùy chỉnh thời gian search ( search các bản tin trong vòng 5, 15, 30 phút, 1h, 2h… 1 ngày, 2 ngày… ).

**2** : Ô tìm kiếm, nơi ta nhập term muốn tìm kiếm ( trong trường hợp này là ssh ).

**3** : Phần Message sẽ hiện các bản tin chứa chính xác những term ta tìm kiếm.

Các thành phần chính trong giao diện của phần Search



**1** : Hiện kết quả tìm kiếm, ví dụ kết quả ở đây báo là tìm kiếm thấy 1 bản tin chưa term “ ssh”, và chỉ tìm kiếm trên 1 index.

2 : Biểu đồ kết quả tìm kiếm. Graylog sẽ tạo 1 biểu đồ hiển thị số bản tin tìm kiếm theo thời gian ( Có thể tùy chỉnh thời gian tìm kiếm )

3 : Message : Hiển thị nội dung bản tin. Bản tin nhận về sẽ chưa tất cả những thông tin có liên quan trong bản tin log chưa term “ssh”

4 : Fields : Hiển thị những trường lọc được từ bản tin, giúp việc đọc bản tin dễ dàng hơn, và có thể tạo thông số và biểu đồ để thêm vào dashboard. Ta cũng có thể tạo thêm những field này với kỹ thuật tạo regex ( Regular Expression ) sẽ được giới thiệu ở phần sau. Ví dụ ở hình dưới ta show ra 2 nội dung là **Statistics** và **Quick value** cho 2 trường message và source, trên giao diện sẽ hiển thị như sau :

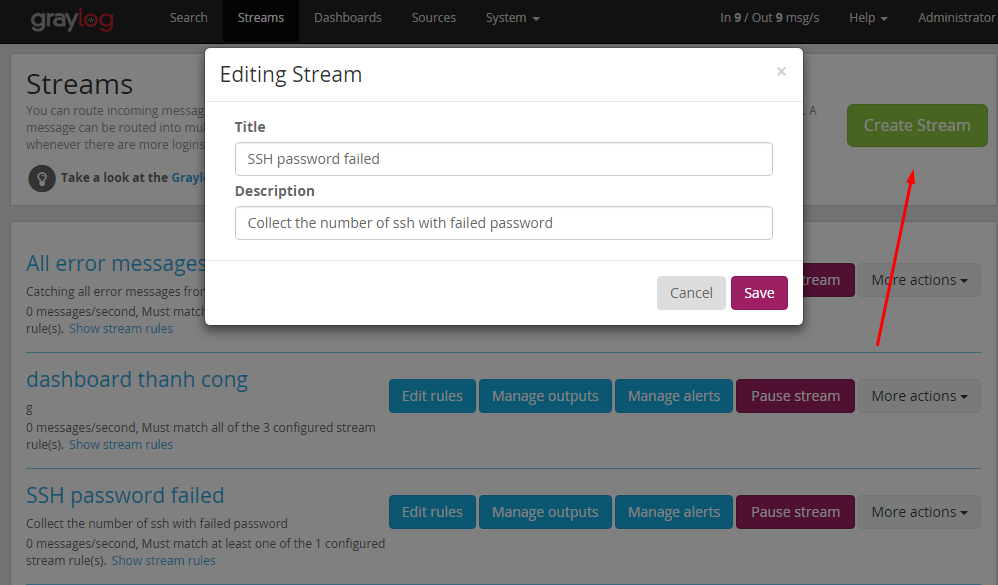


Các bảng và biểu đồ này ta có thể thêm vào dashboard, giúp việc đọc dữ liệu dễ dàng hơn.

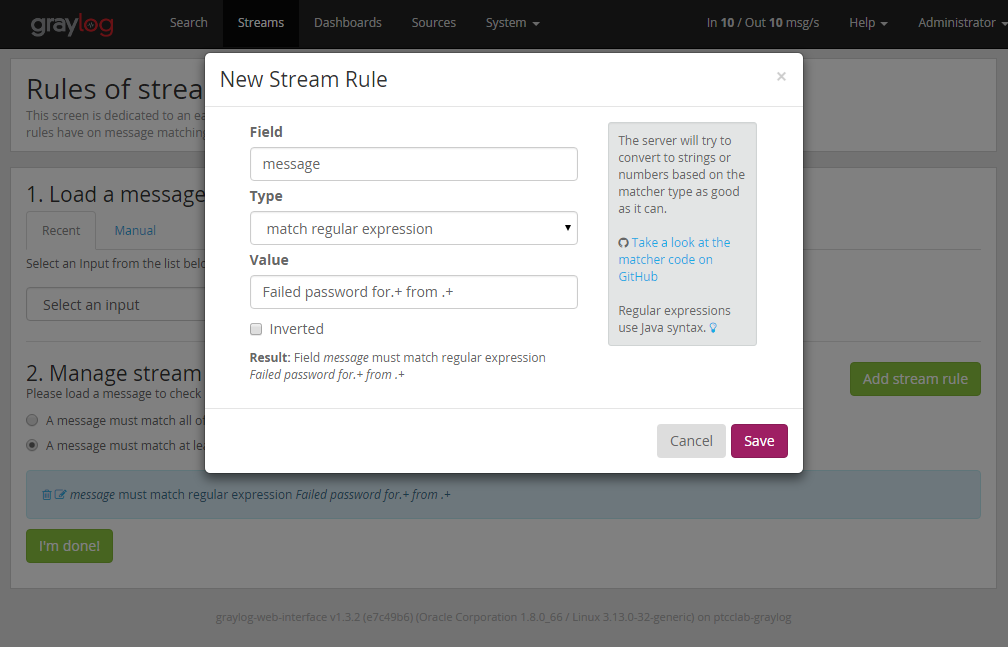
# Stream

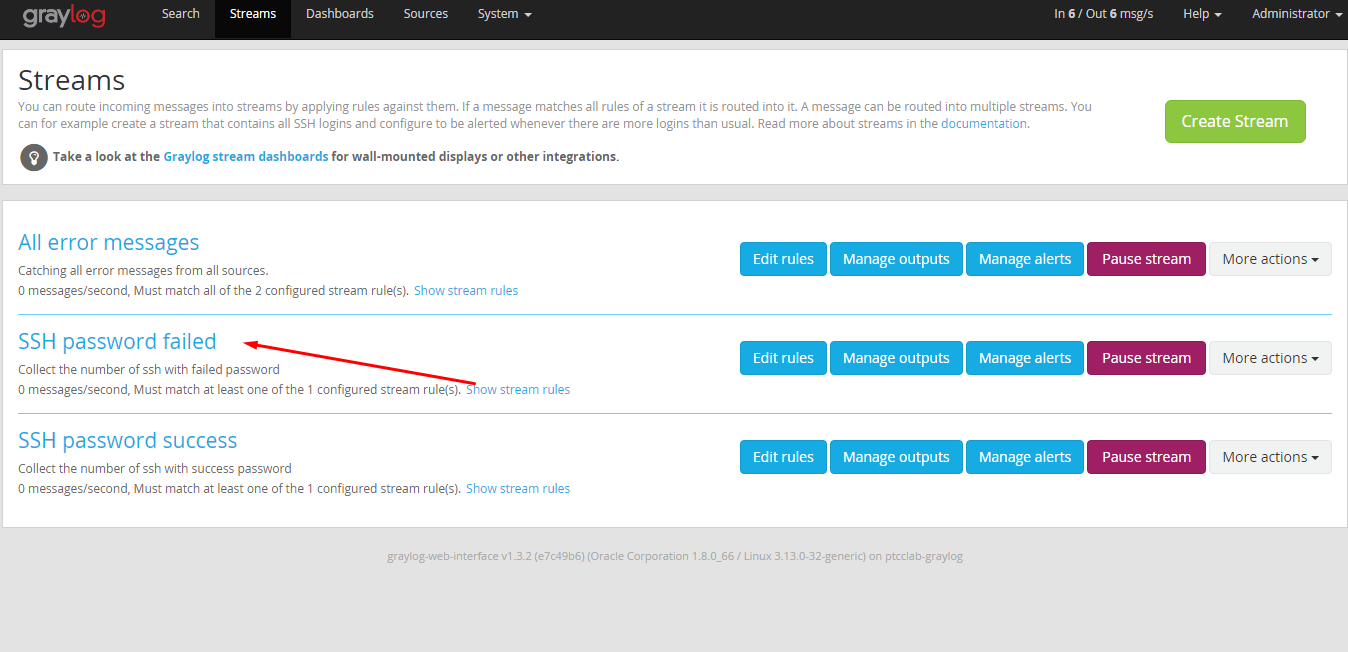
Stream là kỹ thuật định tuyến bản tin tới một chỉ mục nhất định trong thời gian thực. Trên Stream sẽ có những rule để lọc những bản tin mong muốn. Mục đích của Stream nhằm tổng hợp những bản tin cần thiết và từ đó có thể tạo cảnh báo cho người quản trị. Ví dụ như tạo stream SSH Failed để lọc lấy những bản tin chứa ssh thất bại trong vòng 30 phút, ta có thể đặt ngưỡng là trong vòng 30 phút nếu số lượng ssh thất bại vượt quá 50 lần thì sẽ gửi mail cảnh báo hoặc gửi tin nhắn đến slack.

Trong mục Stream, ta click vài Create Stream, sau đó điền Title và Description cho Stream.

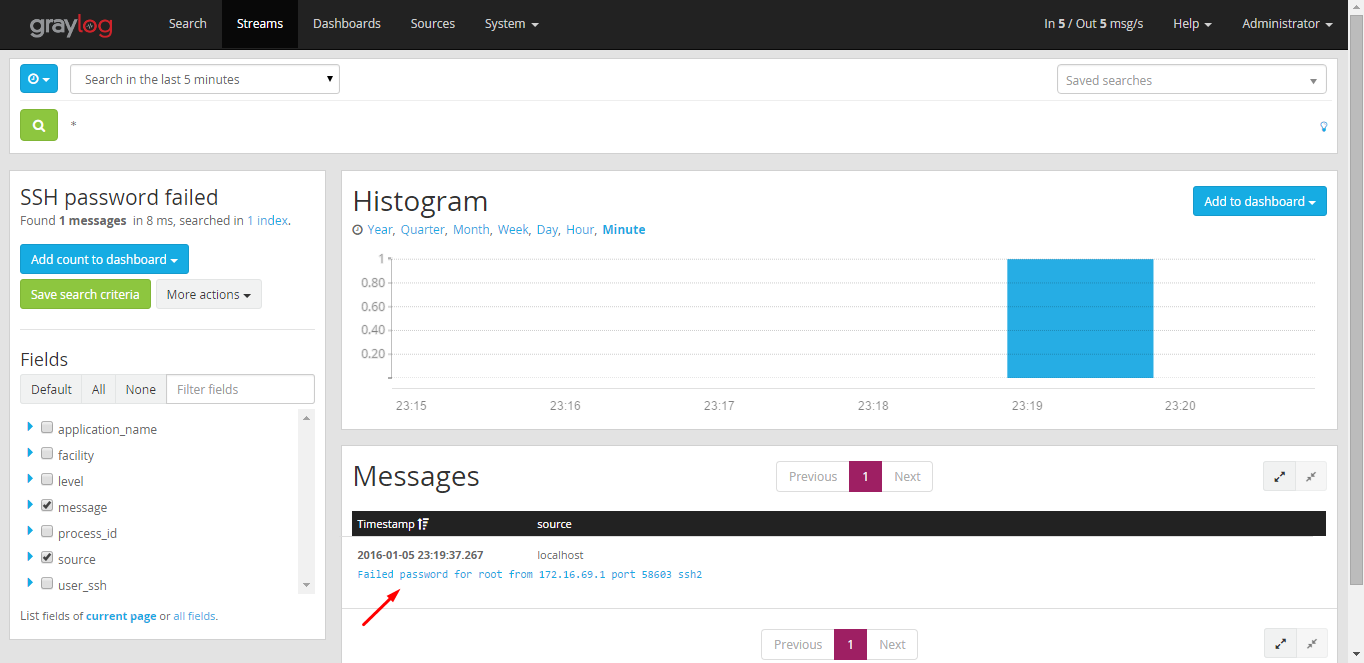


Click vào Edit Rule, sau đó thêm Rule mới cho Stream. Ví dụ ở đây ta sẽ tạo rule để thu thập các bản tin chứa regex là : Failed password for .+ from .+

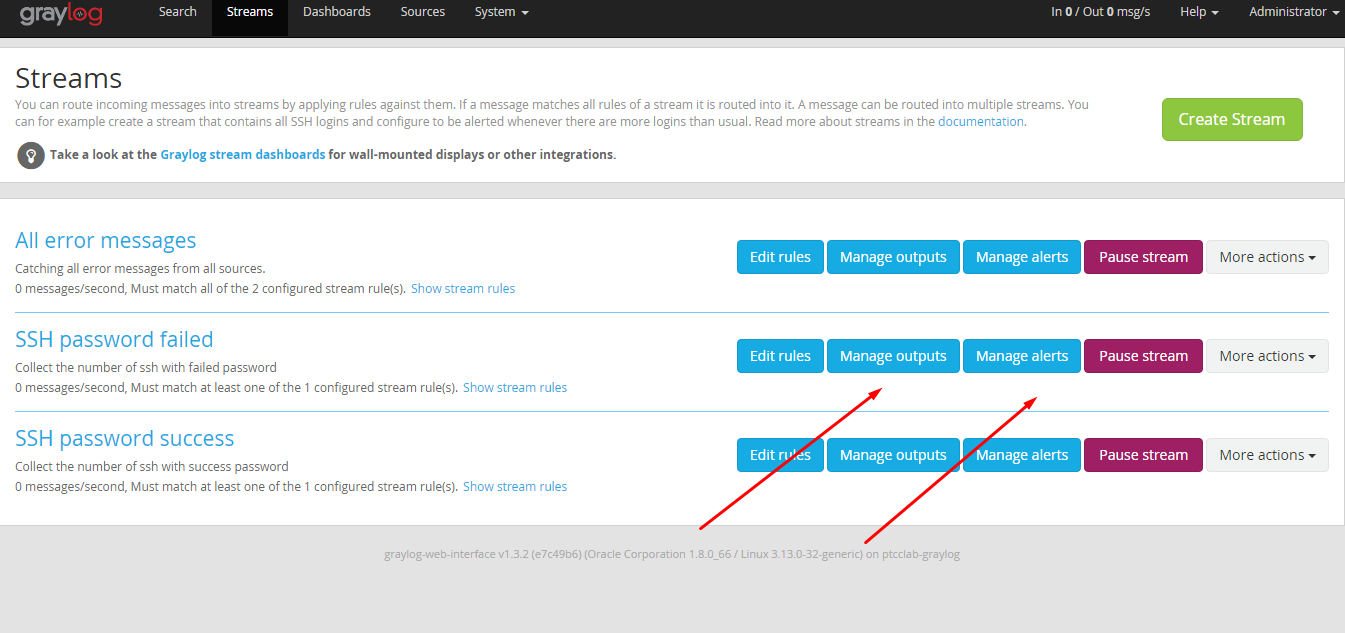




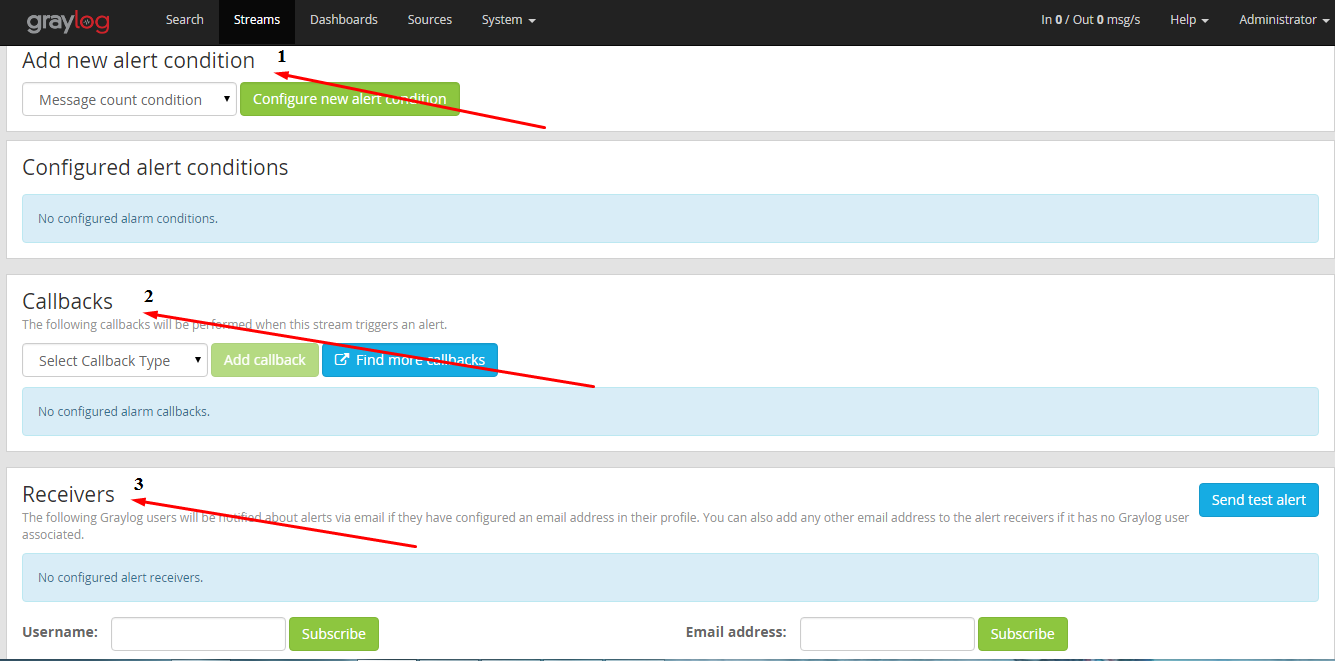
Stream SSH Fail sẽ bắt tất cả những bản tin chứa chuỗi :Failed password for… from …



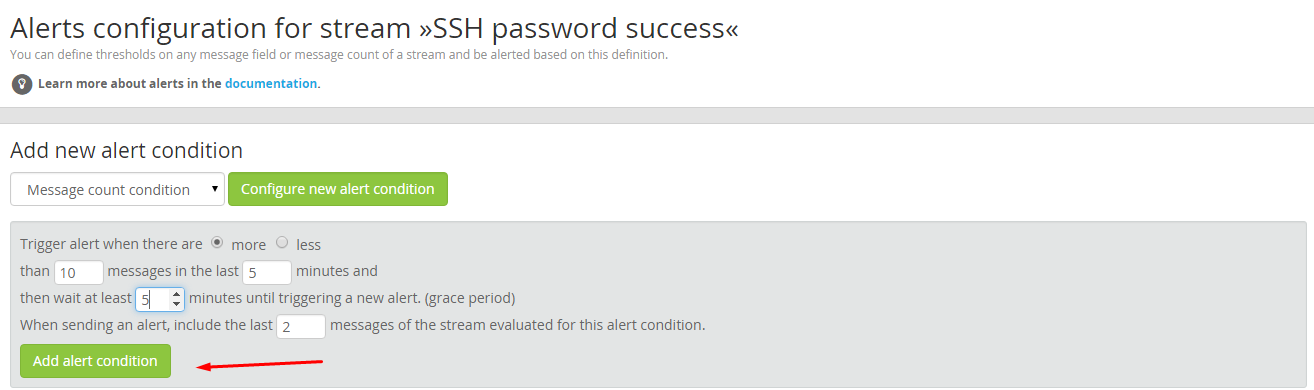
## Tạo cảnh báo qua Slack hoặc Email với Stream



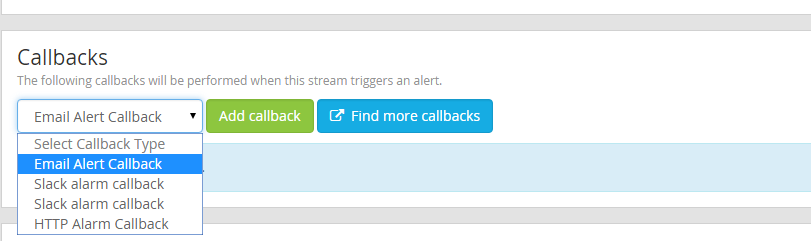
Tôi sẽ giới thiệu tạo cảnh báo gửi qua slack với chức năng Manage Alert của Stream, phần Manage Output sẽ được đề cập ở phần sau.



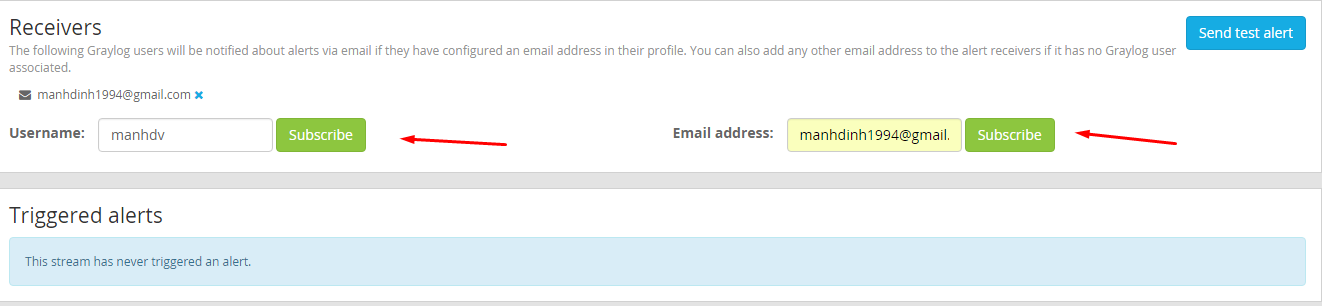
**1**: Tạo điều kiện cho cảnh báo, ví dụ muốn tạo 1 cảnh báo là khi Stream nhận nhiều hơn 10 bản tin về ssh fail trong vòng 5 phút cuối và đợi ít nhất 5 phút cho đến khi gửi một cảnh báo mới. Khi gửi cảnh báo thì chứa 2 bản tin cuối cùng của stream.



**2** : Chọn loại Alert Callback và Add

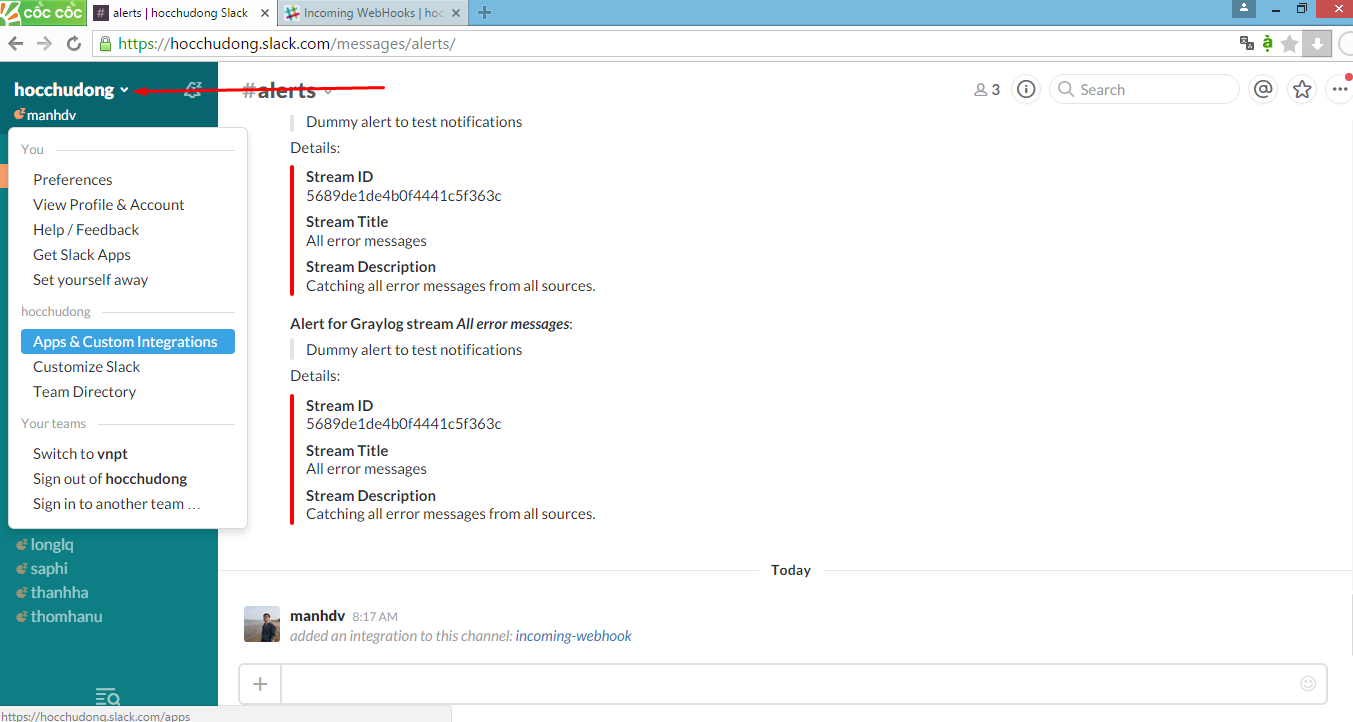


**3 :** Nhập User hoặc địa chỉ email mà bạn muốn cảnh báo được gửi đến.

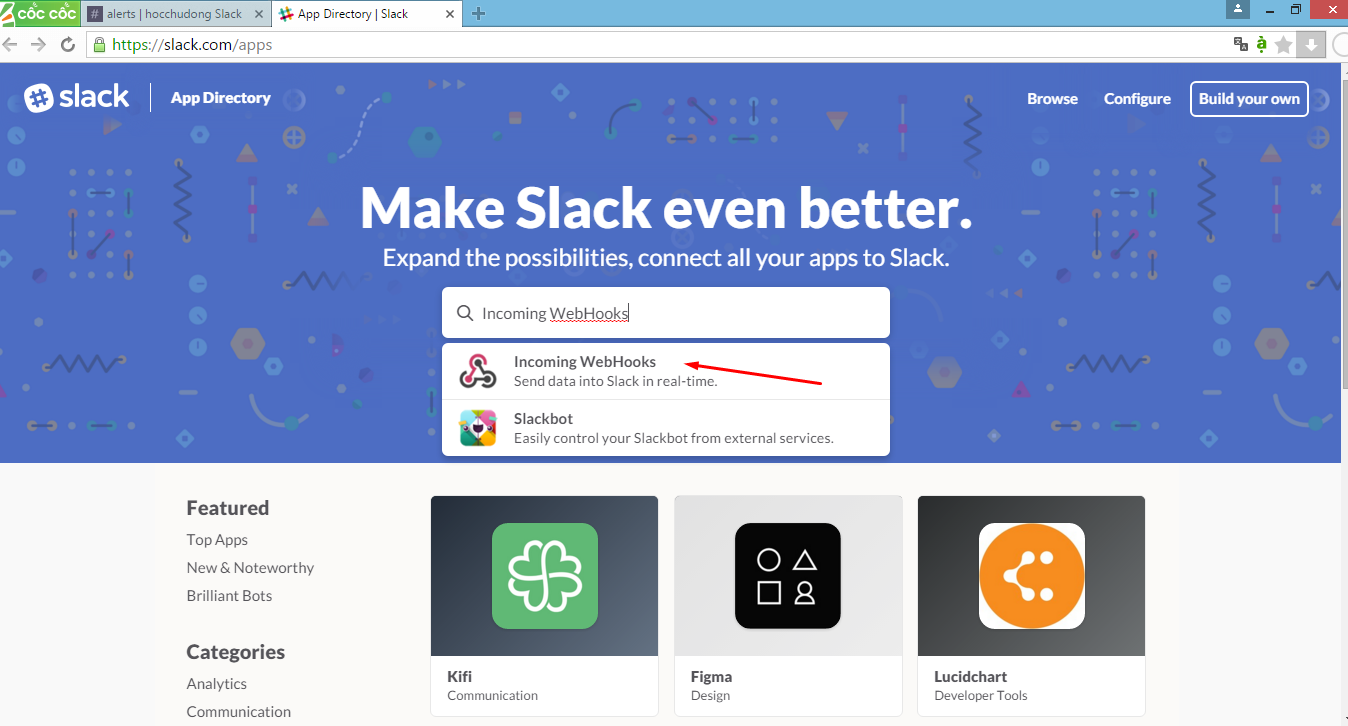


\*Với việc gửi cảnh báo qua kênh Slack, bạn cần phải cấu hình Incoming Webhook trên Slack theo hướng dẫn sau :

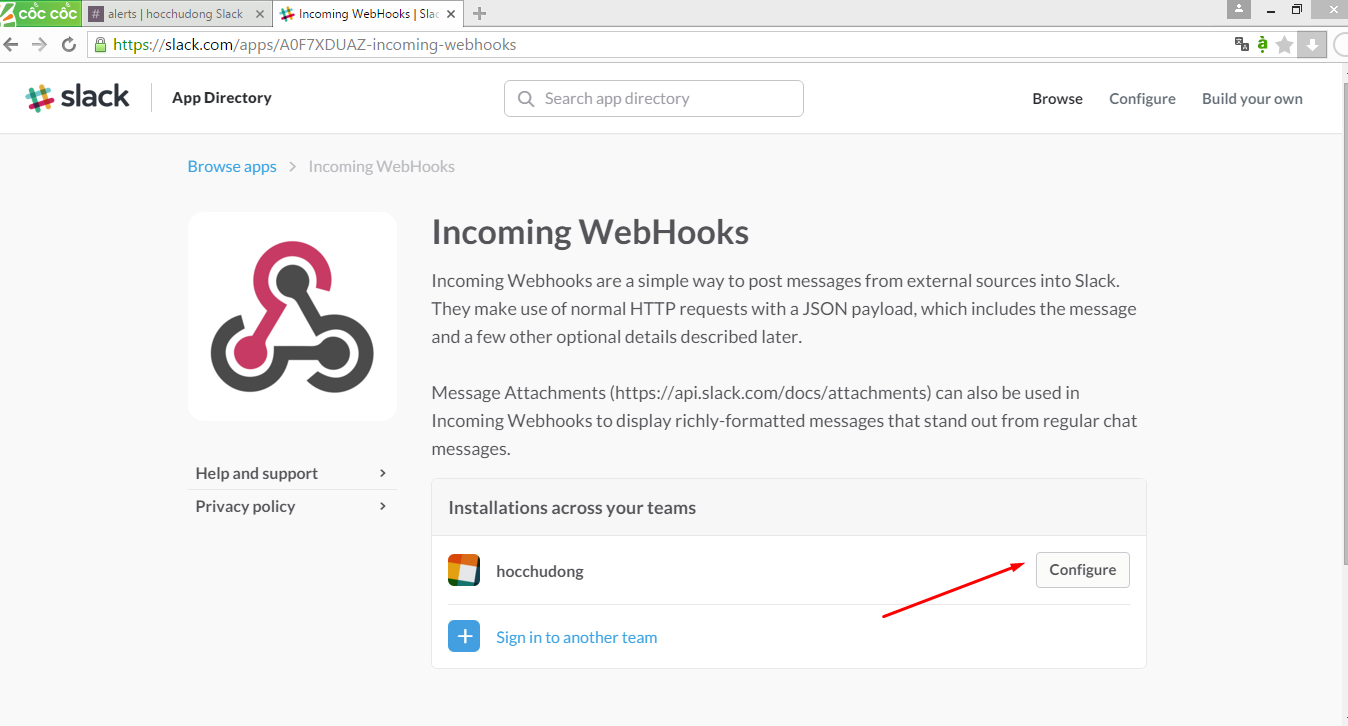
Step 1 : Vào Slack, ấn vào hình có mũi tên đỏ, chọn phần Apps & Custom Intergration



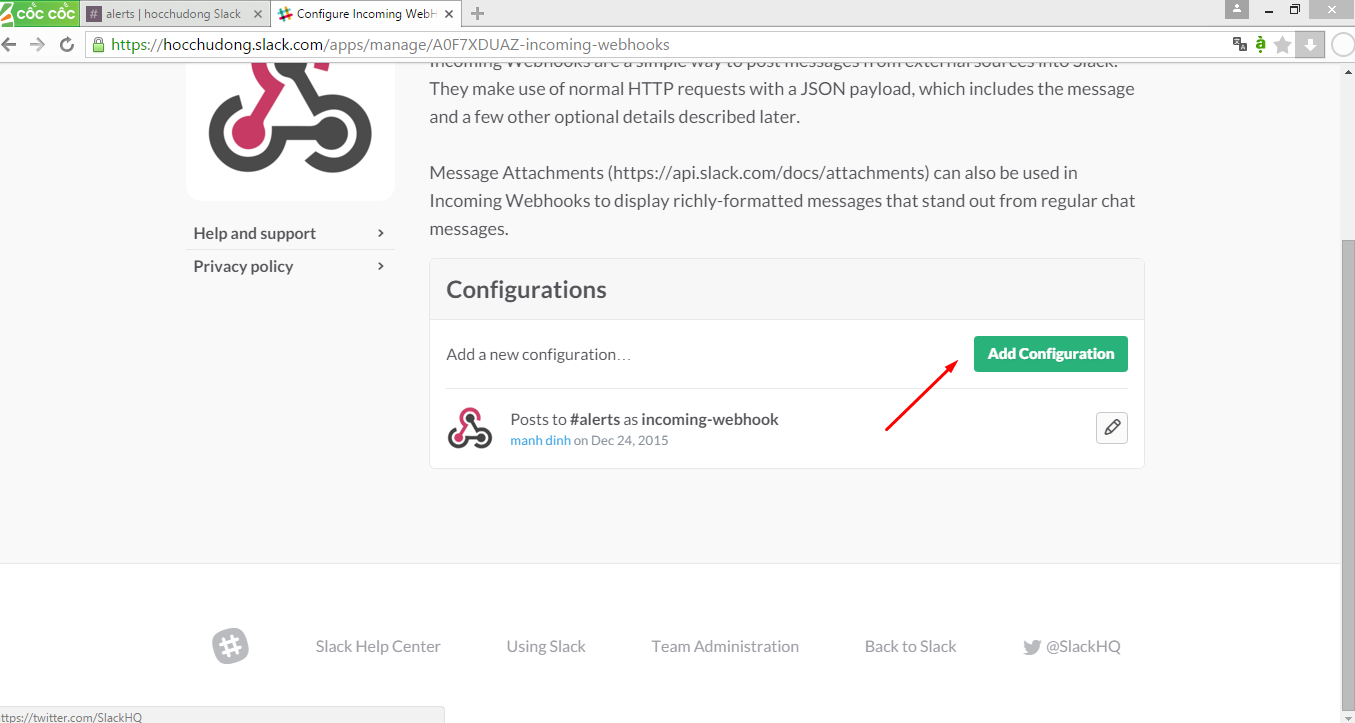
Step 2 : Tại cửa sổ Apps & Custom Intergration, tìm kiếm dịch vụ Web Incoming Webhooks



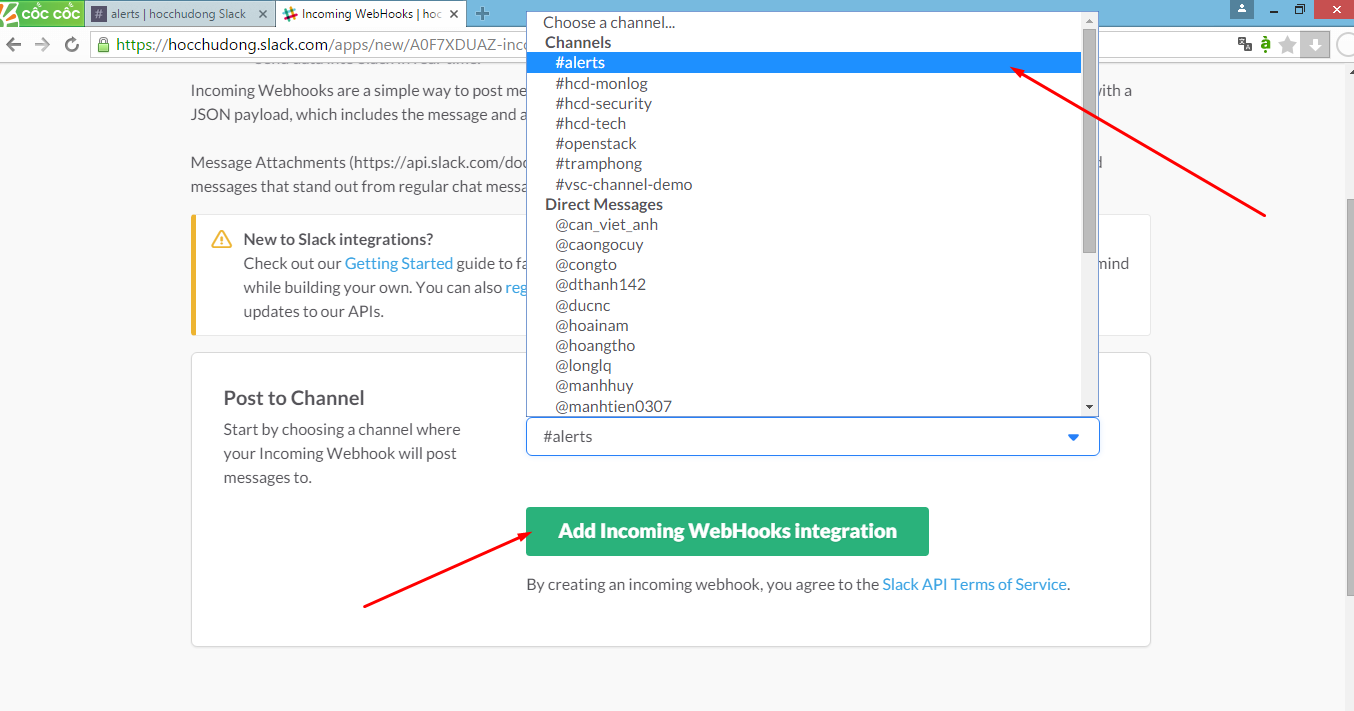
Step 3 : Chọn Domain Slack của bạn, ấn configure



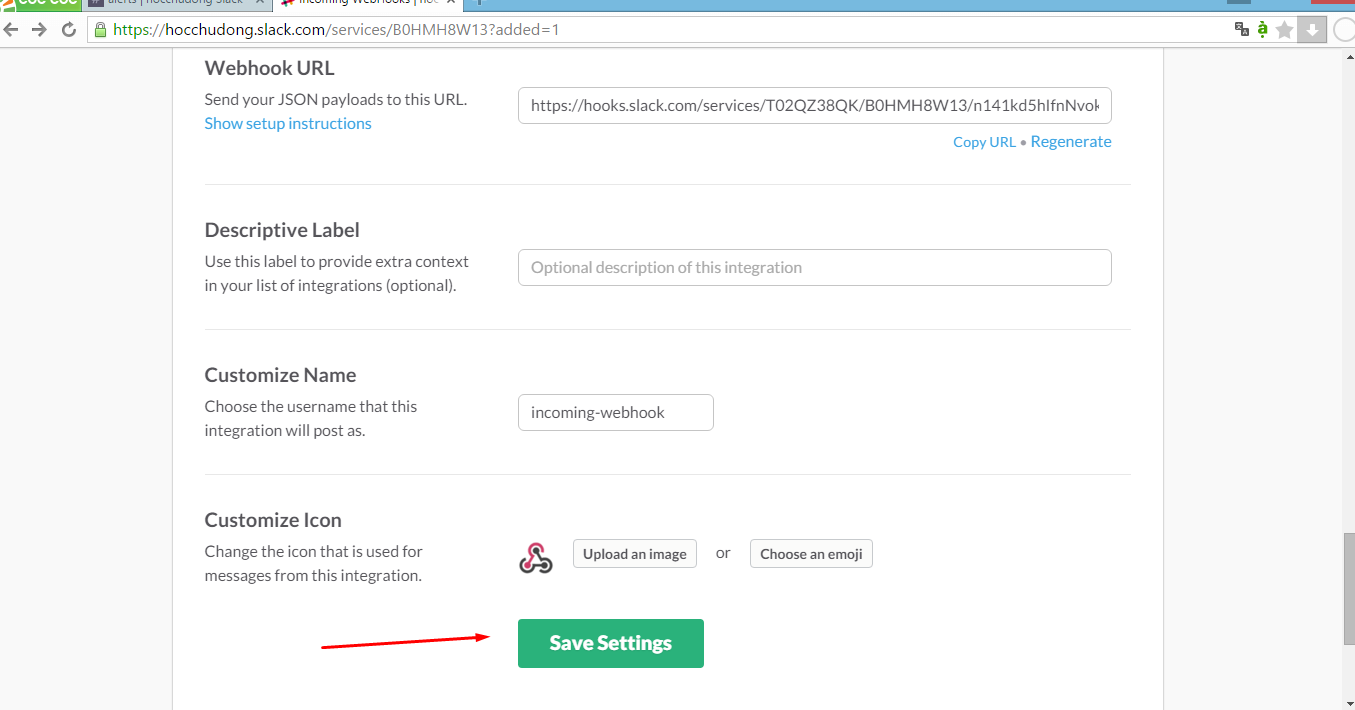
Step 4 : Add một configuration mới :



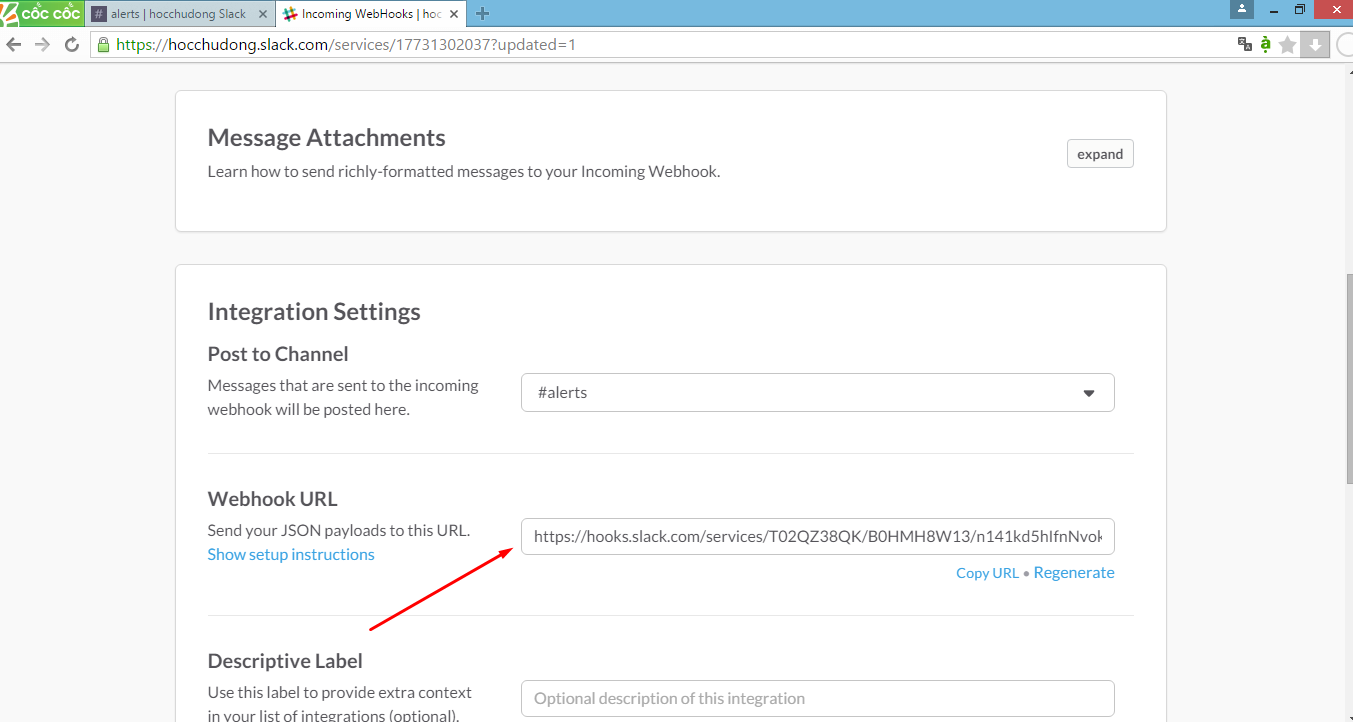
Step 5 : Chọn kênh bạn muốn thiết lập để gửi cảnh báo về (Nên tạo 1 kênh riêng biệt)



Step 6 : Trong trang thiết lập, có thể tùy chỉnh thông số ( thêm image, icon… ) hoặc giữ nguyên như mặc định.

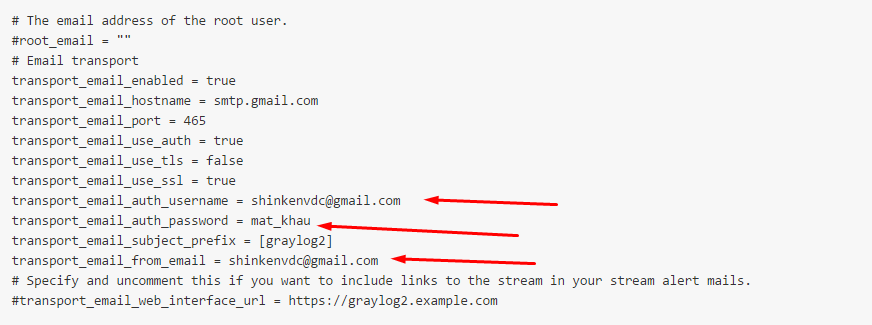


Step 7 : Sau khi cài đặt xong, Webhook mới sẽ có 1 URL riêng, các bạn chú ý dùng URL này cho các cài đặt của dịch vụ khác ( Graylog, Zabbix… ) :

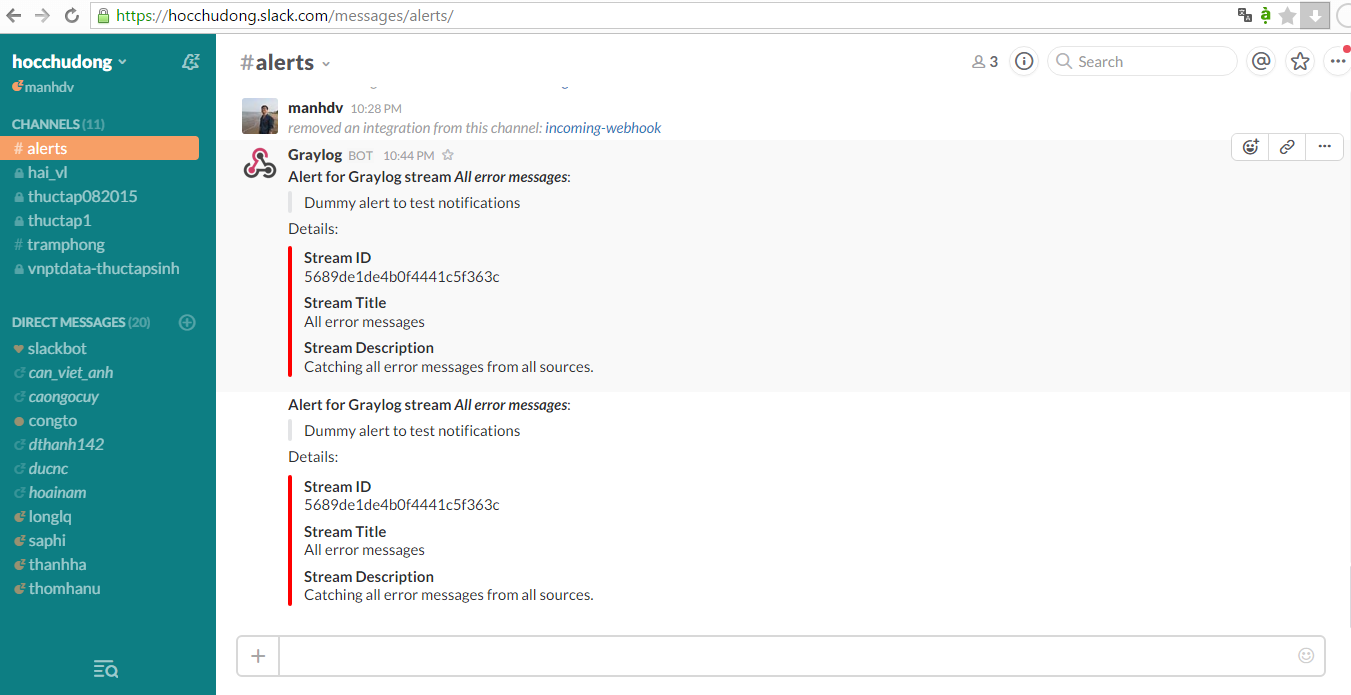


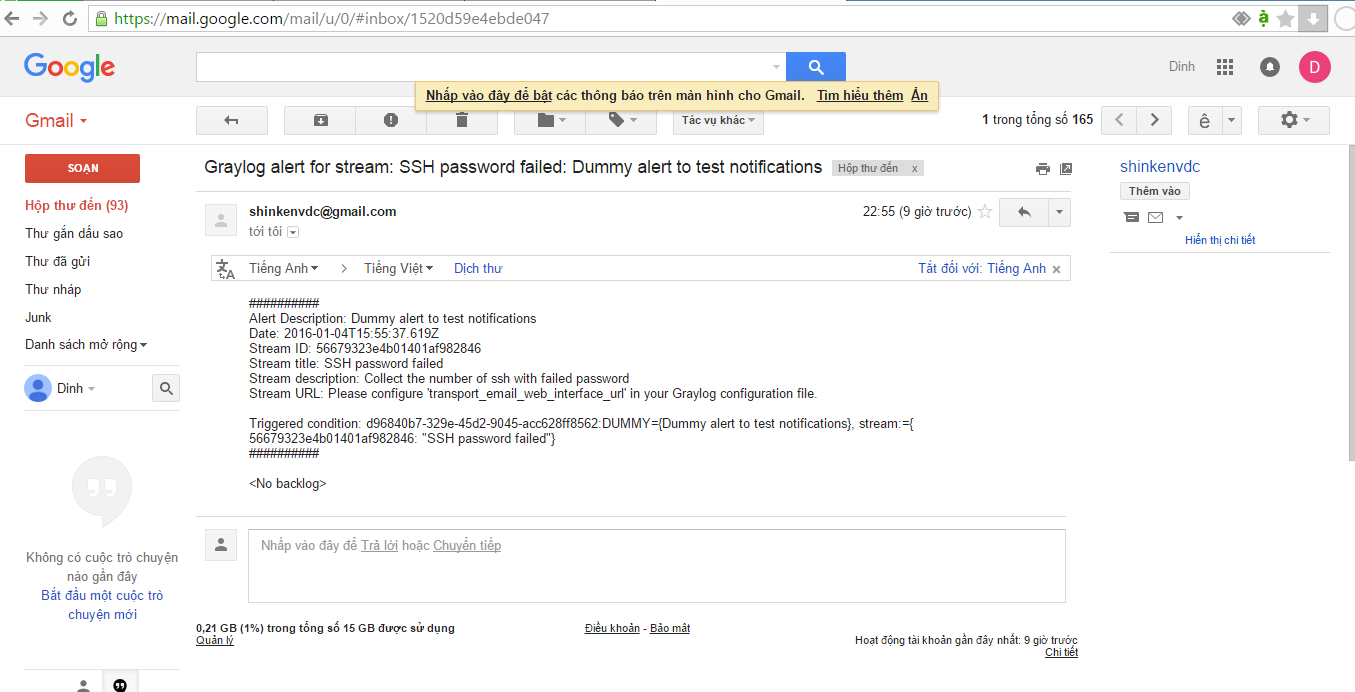
\*Với việc gửi cảnh báo qua email, thiết lập ở file config của Graylog như sau :

Thêm các dòng sau vào file conf của Graylog Server (thay các dòng có mũi tên đỏ với địa chỉ email và mật khẩu tương ứng) :



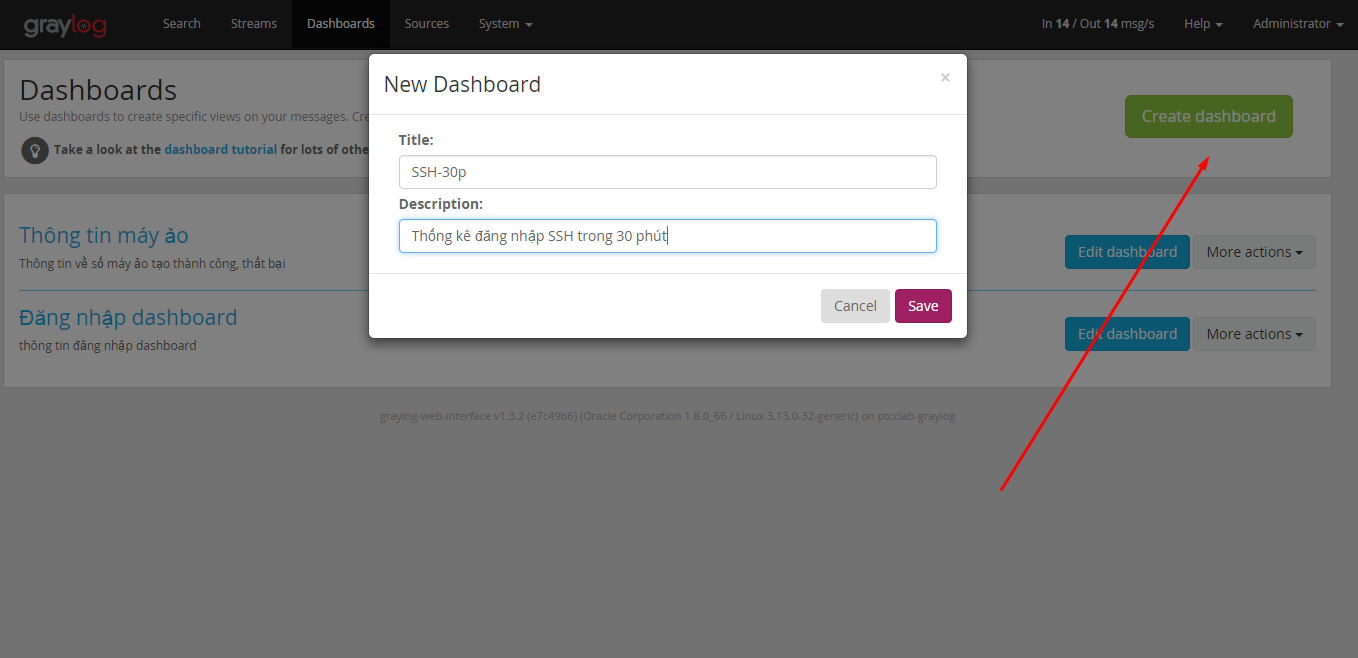
Ấn gửi test alert để kiểm tra xem đã cấu hình đúng hay chưa. Kết quả :



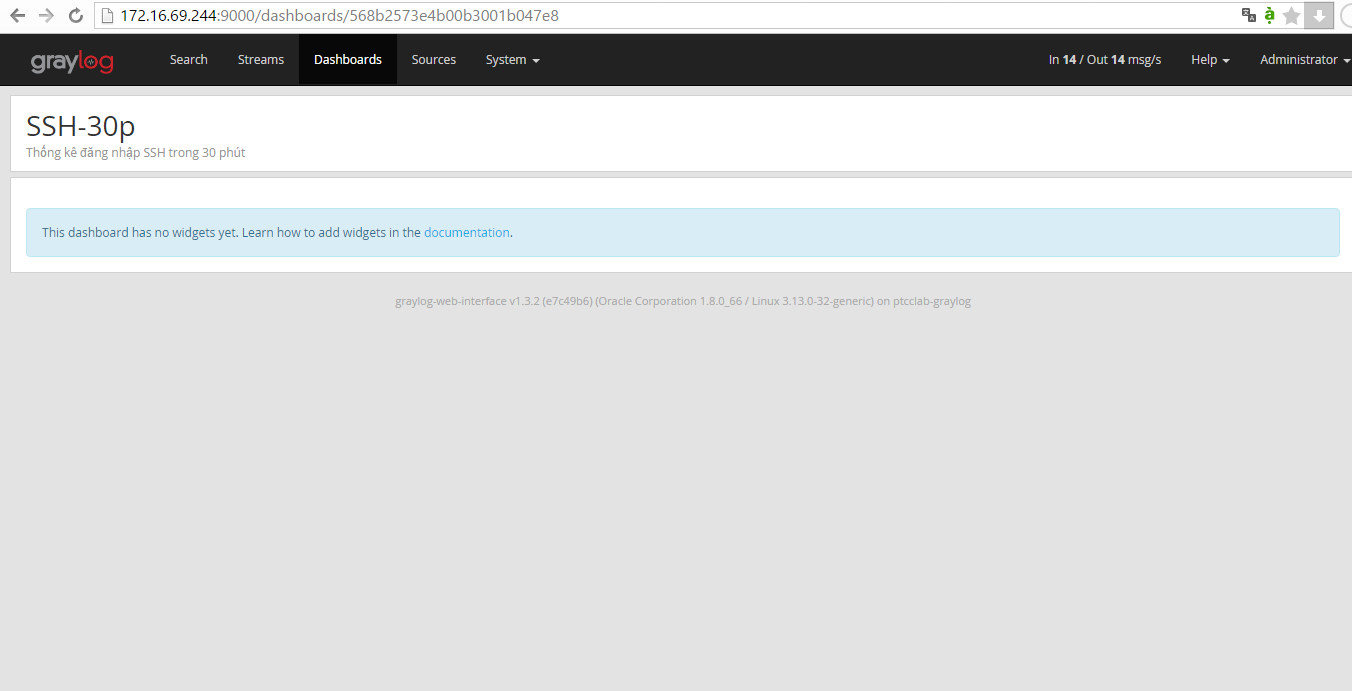


# Dashboard

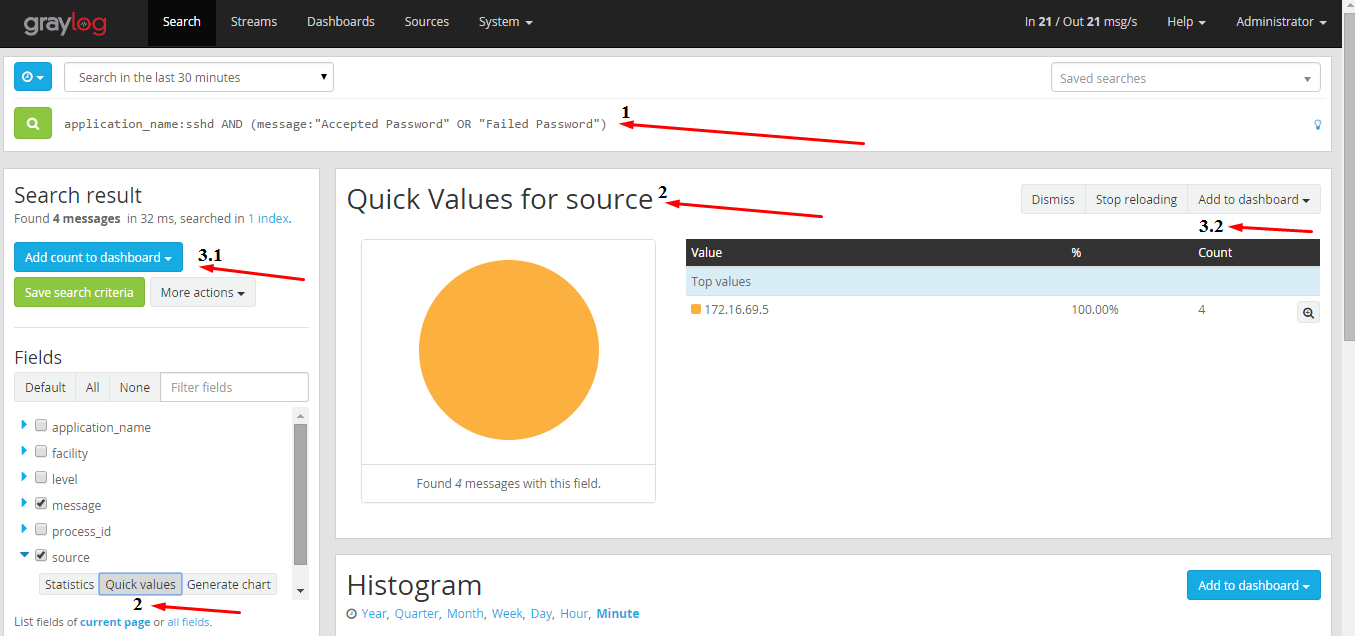
Trên Dashboard, tạo một Dashboard mới, thêm Title và Description cho Dashboard



Sau khi tạo xong, Dashboard hiện giờ vẫn đang rỗng, ta cần thêm thông số vào Dashboard thông qua Search hoặc Stream.



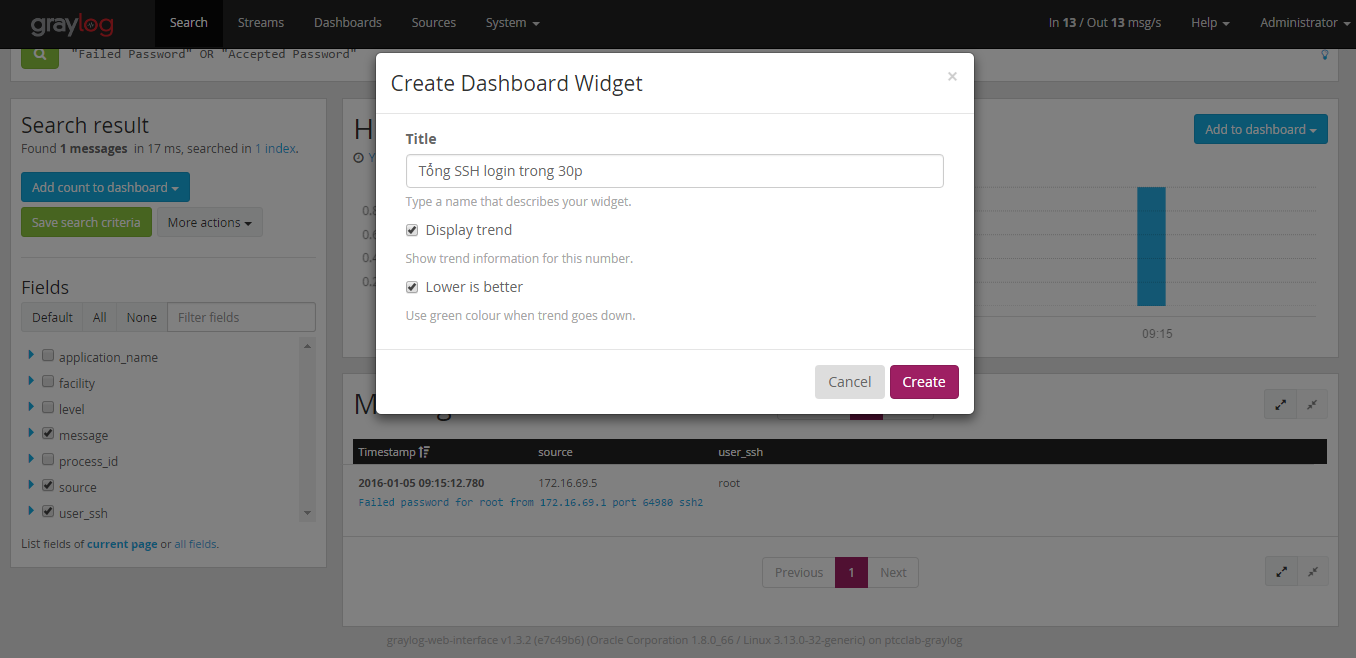
Ví dụ 1 : Dashboard thống kê số lần đăng nhập SSH thành công, thất bại, và tổng số đăng nhập trong 30 phút



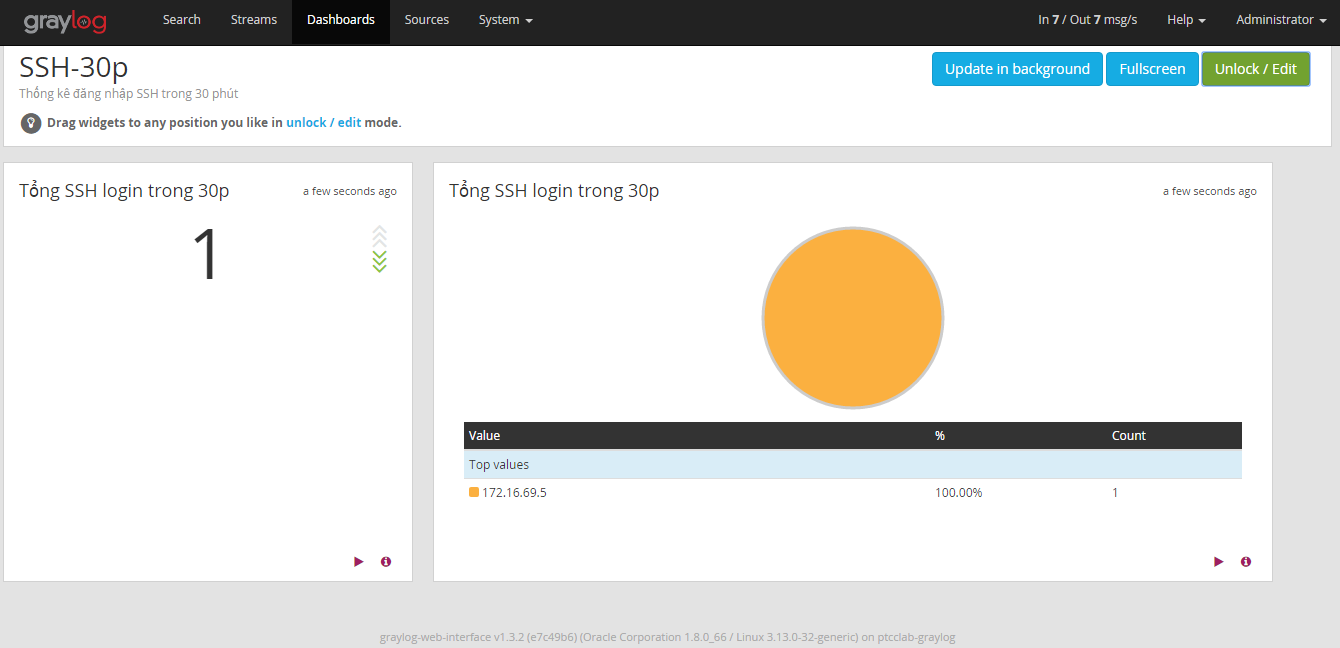
**1** : Nhập term tìm kiếm ở mục search, biểu thức tìm kiếm ở mục 1 có ý nghĩa là : Tìm kiếm các bản tin các chính xác 2 cụm từ : Failed Password hoặc Accepted Password.

**2** : Hiện lên thông số và biểu đồ thống kê cho từng field, ví dụ ở đây khi chọn filed là source và chọn mục Quick Value, sẽ hiện biểu đồ và bảng thống kê những source nào có xuất hiện các bản tin trên.

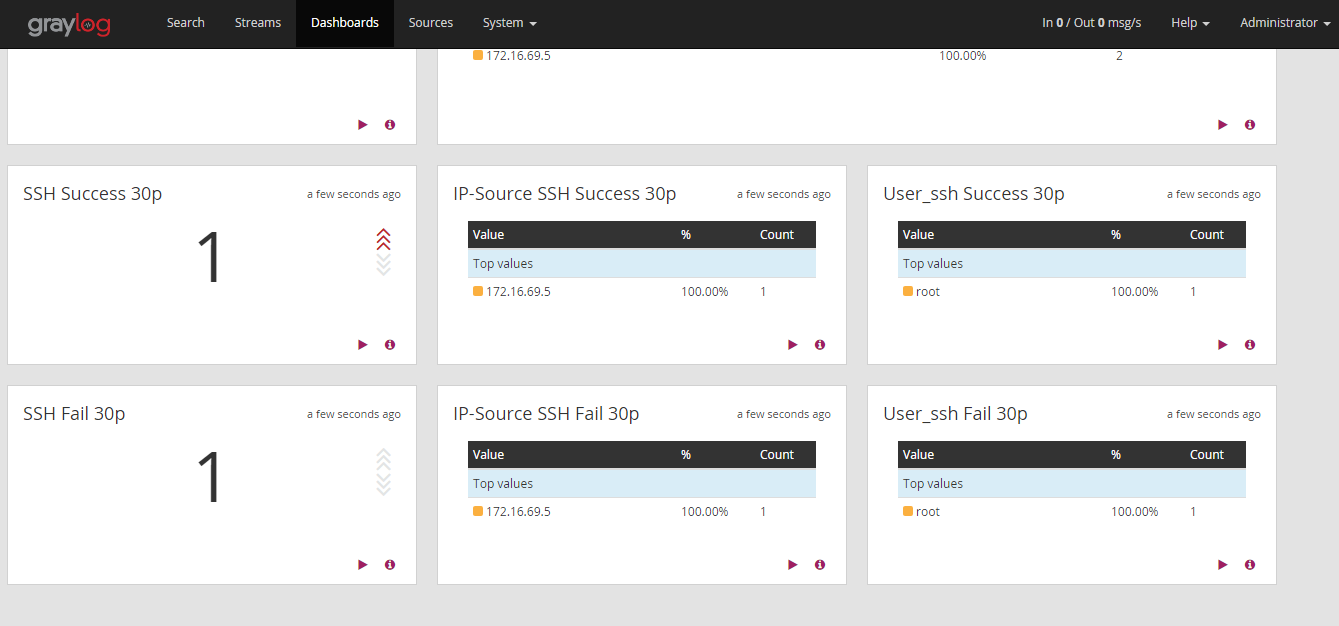
3 : Chọn các mục muốn thêm vào Dashboard, ví dụ muốn thêm tổng số lần đăng nhập SSH, ta thêm theo 3.1 ( Thực chất ở đây là đếm số bản tin log nhận về )



Làm tương tự với 3.2. Ta có Dashboard sau

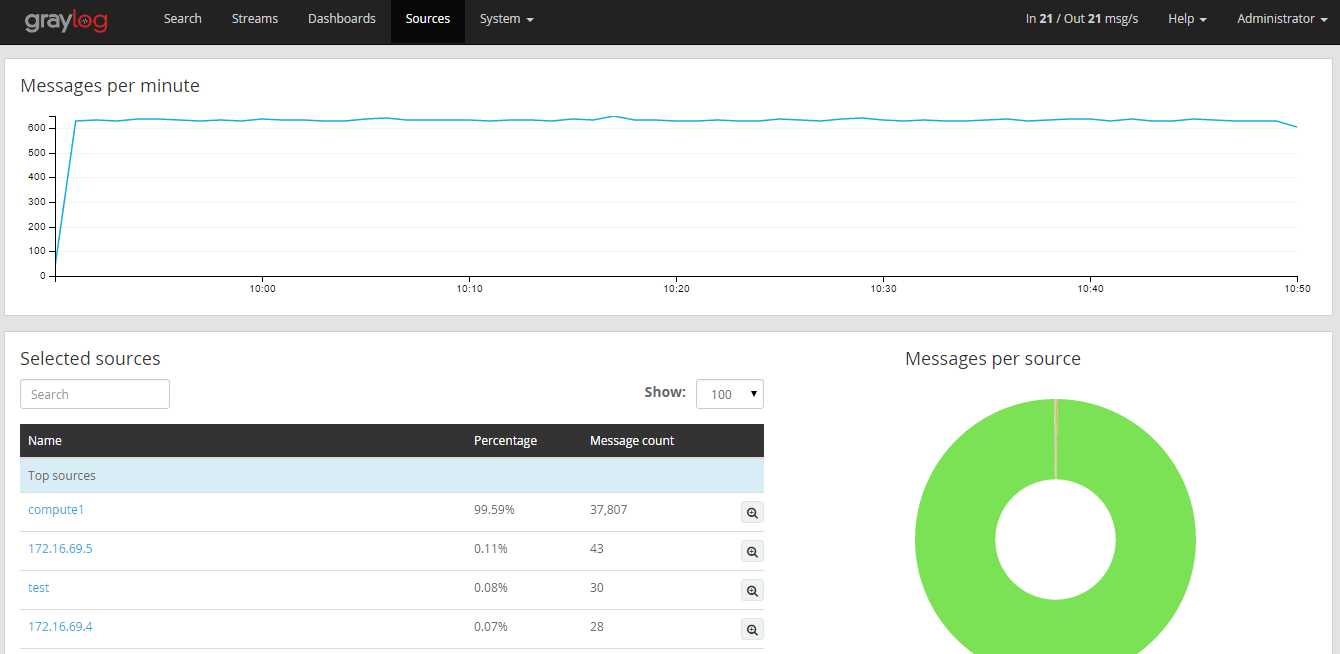


Lấy thêm các thông tin về IP, User… đăng nhập SSH thành công thất bại…



# Sources

Sources thống kê số bản tin nhận về theo thời gian dưới dạng biểu đồ, và ip đang đẩy log về Graylog



# System

Trong phần chúng ta sẽ chỉ tập trung vào một số phần quan trọng của System như : Input, Output, Collector, Content Packs.

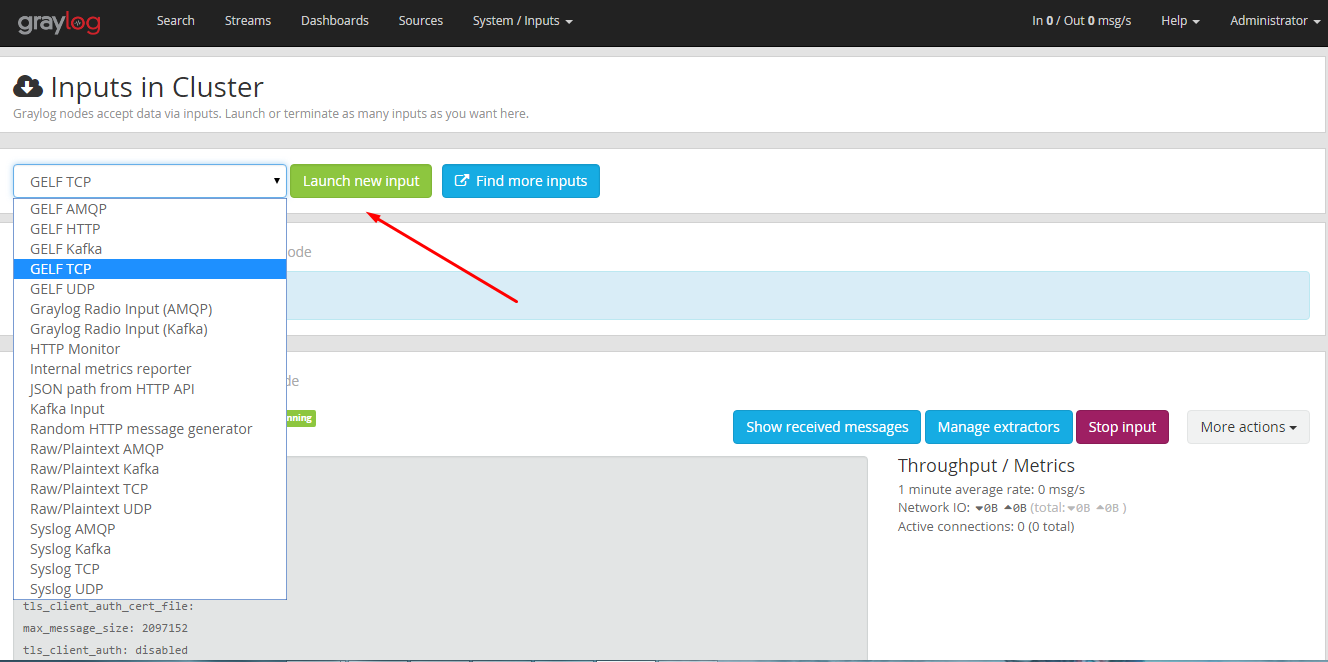
## Input

Input giống như một địa chỉ nhà, các bản tin từ máy client sẽ được cung cấp thông tin về địa chỉ đó để có thể đẩy được log về cho Graylog Server. Có rất nhiều dạng Input nhưng ở đây chúng ta sẽ chỉ đề cập đến 2 loại input là GELF TCP và Syslog để sử dụng với 2 kỹ thuật đẩy log là đẩy bằng Syslog thuần túy và đẩy bằng Graylog Collector.

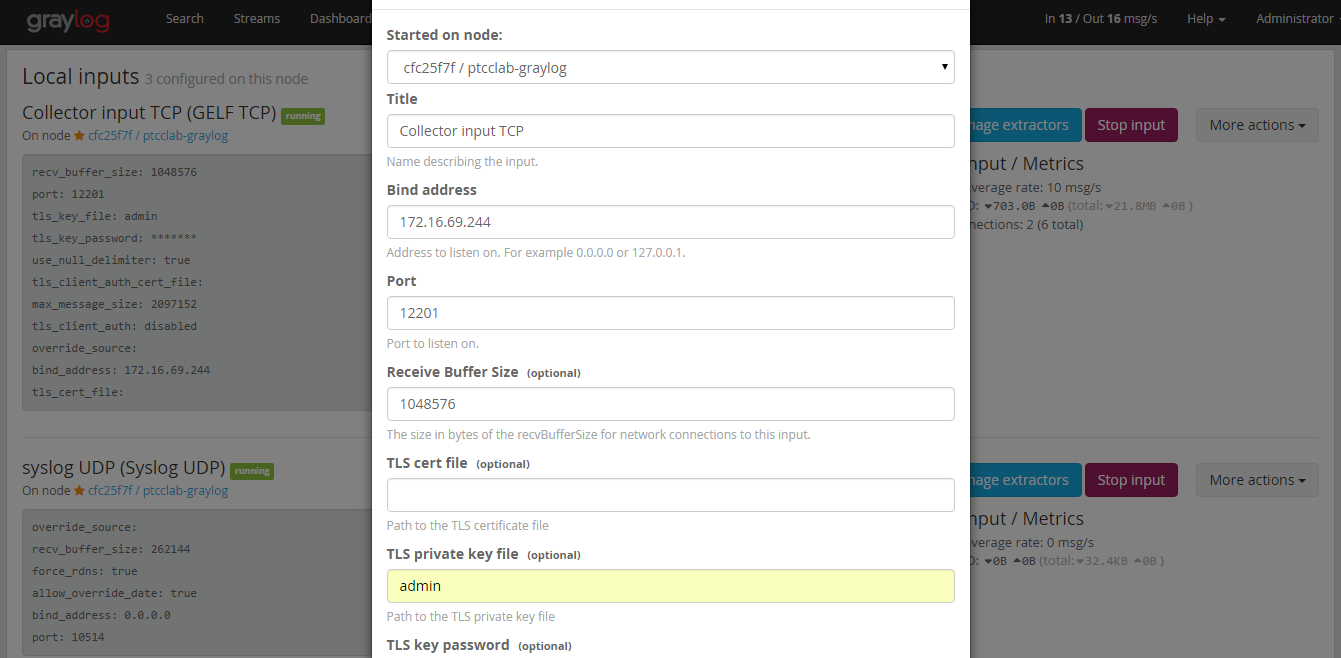
### GELF TCP

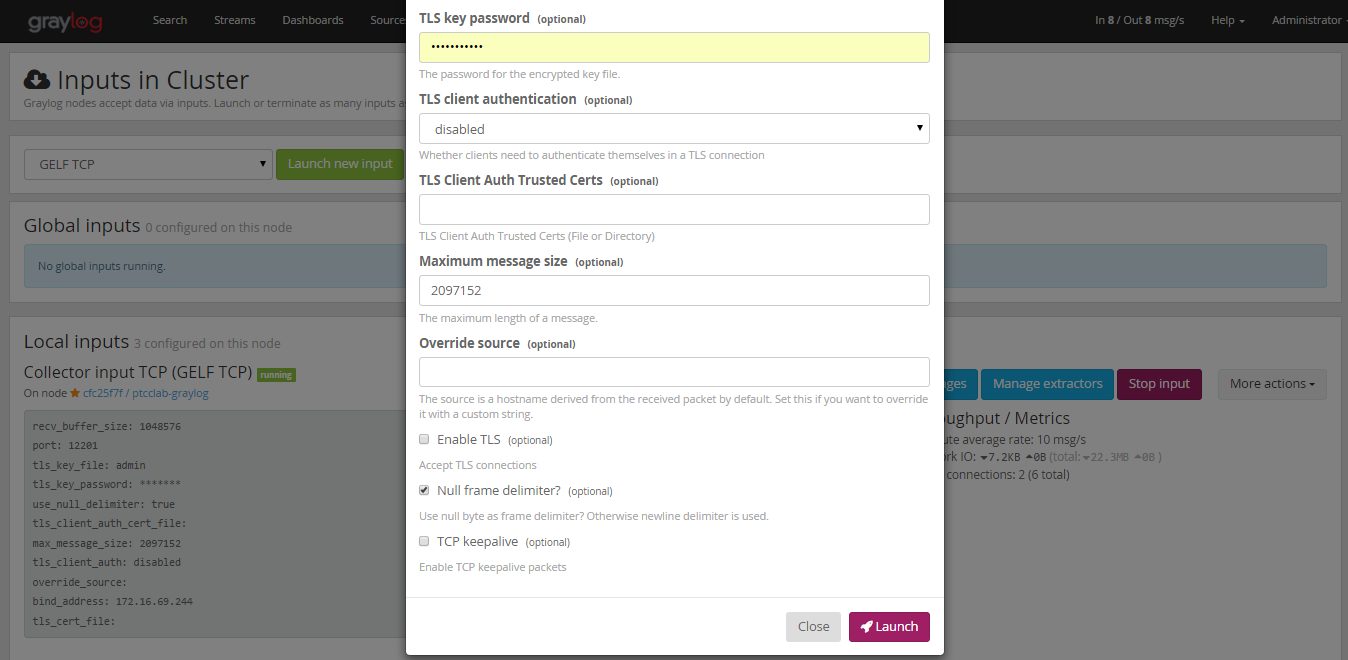
GELF TCP là input chuyên dùng cho Graylog collector, xem hướng dẫn về Graylog-Collector theo link sau :

https://github.com/manhdinh/ghichep-graylog/tree/master/graylog-collector



Sau khi launch input mới, ta nhập các thông tin cần thiết vào bảng

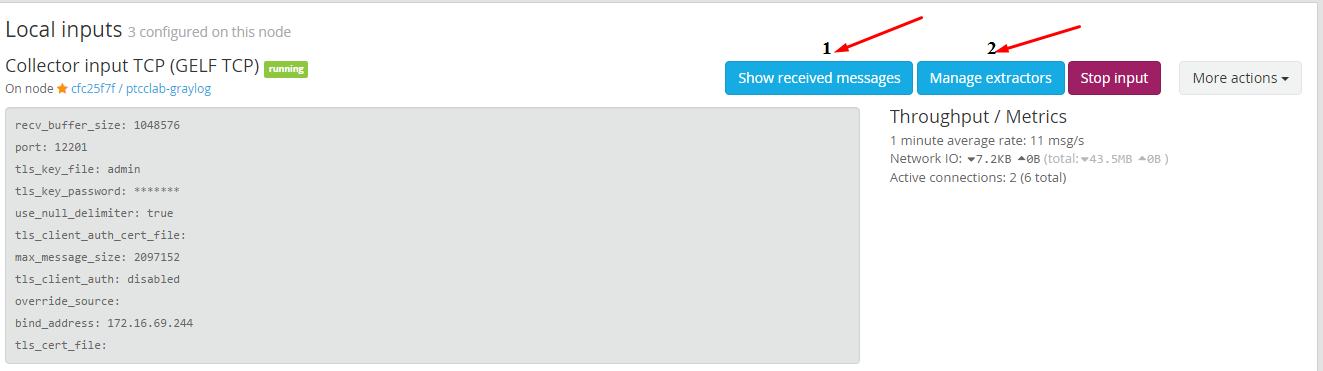




Một số mục cần lưu ý khi nhập thông tin :

* Bind IP : Nhập IP của Graylog-Server hoặc 0.0.0.0 ( Nếu đặt 0.0.0.0 Graylog server sẽ lắng nghe tất cả các bản tin trả về, chỉ đặt nếu đã thiết lập IPTables)
* Port : Chú ý đặt trùng với port thiết lập trong file config của máy Collector Client ( mặc định của cả 2 là 12201 )
* Bạn có thể tìm hiểu thêm về cơ chế đẩy log với TLS để sử dụng TLS với Graylog-Colector để bảo mật tốt hơn khi truyền các bản tin log.

Sau khi launch xong input, cần có 2 phần của Input cần lưu ý



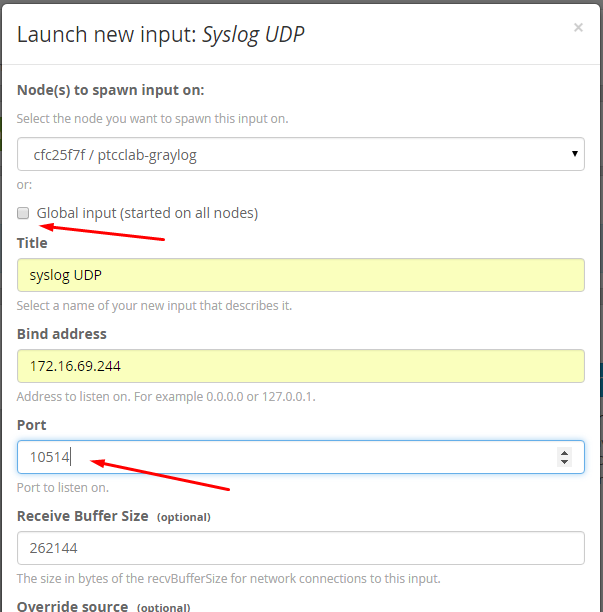
**1** : Kiểm tra xem các bản tin log đã được đẩy về hay chưa. ( Các bản tin sẽ bắt đầu đẩy về sau khi graylog-collector service khởi động )

**2** : Quản lý các extractor được tạo ở phần searching.

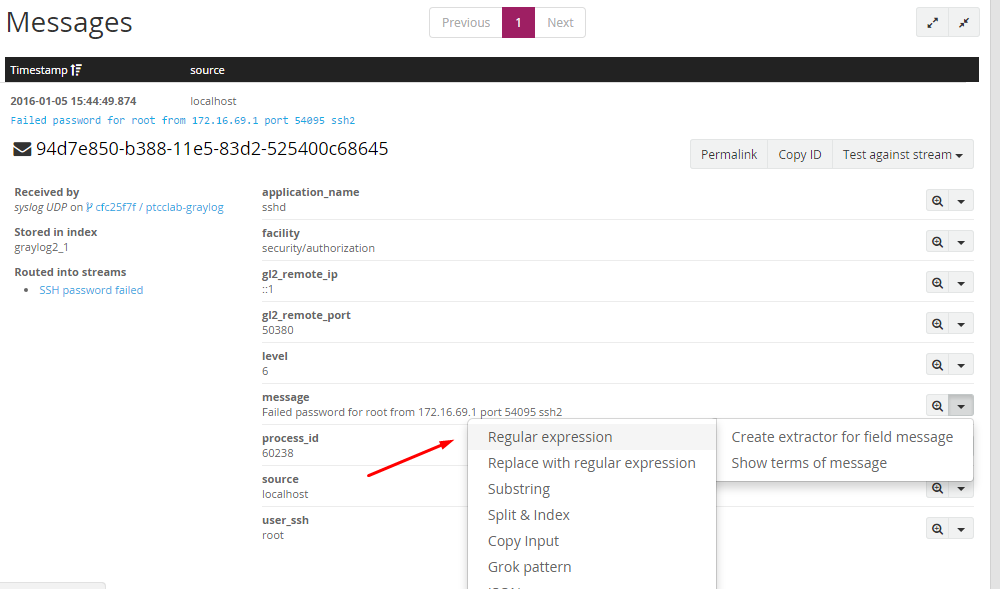
### Syslog UDP

Nhận những bản tin được đẩy bằng syslog thuần túy.

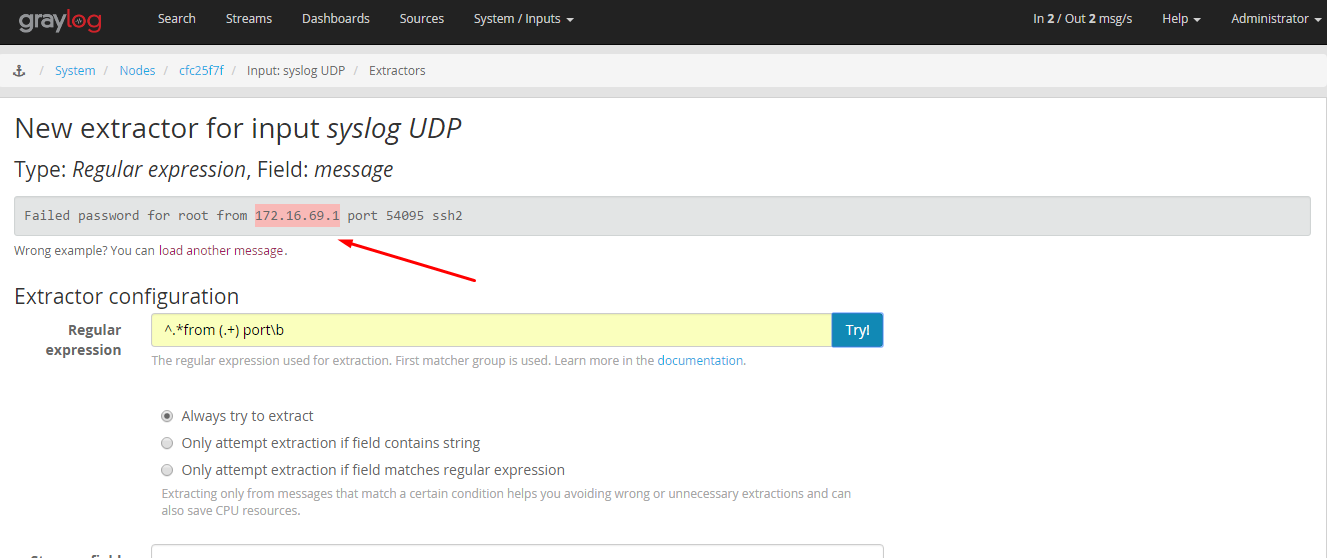
Các thông số không khác với GELF TCP, Syslog Input hỗ trợ



Ở đây tôi sẽ lấy ví dụ về cách cắt lọc nội dung của bản tin với regex :



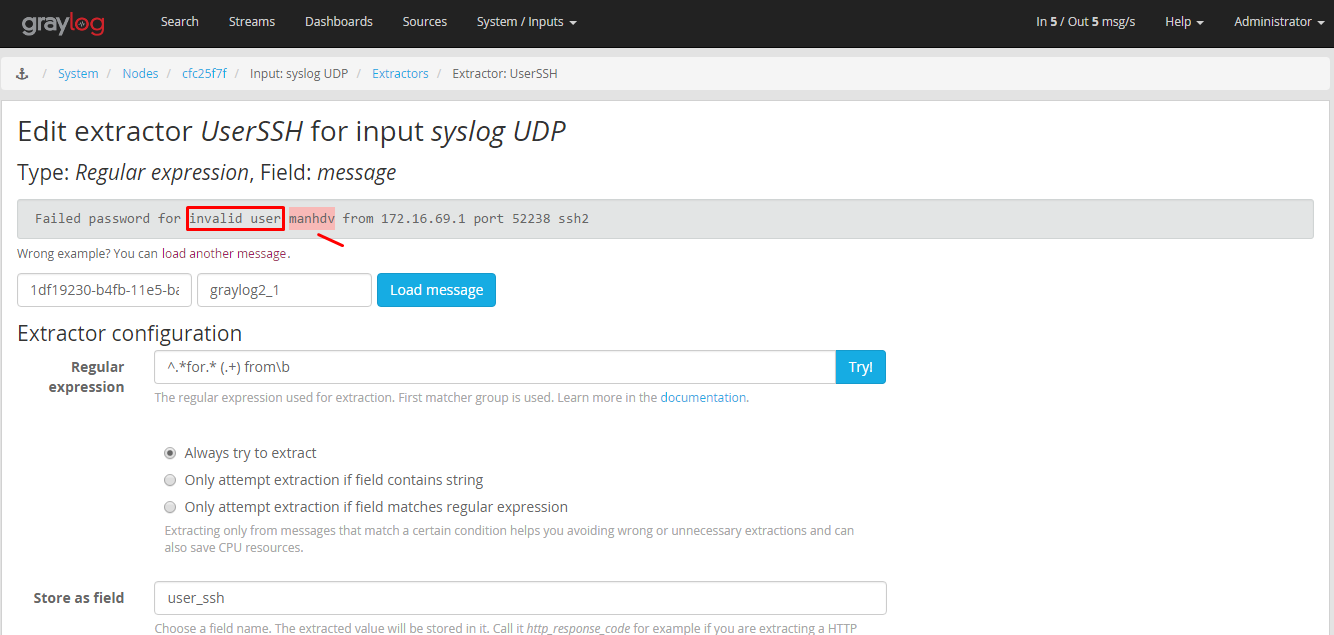
Ví dụ ta muốn lọc thông tin về ip của source, dùng cú pháp sau : **^.\*from (.+) port\b**



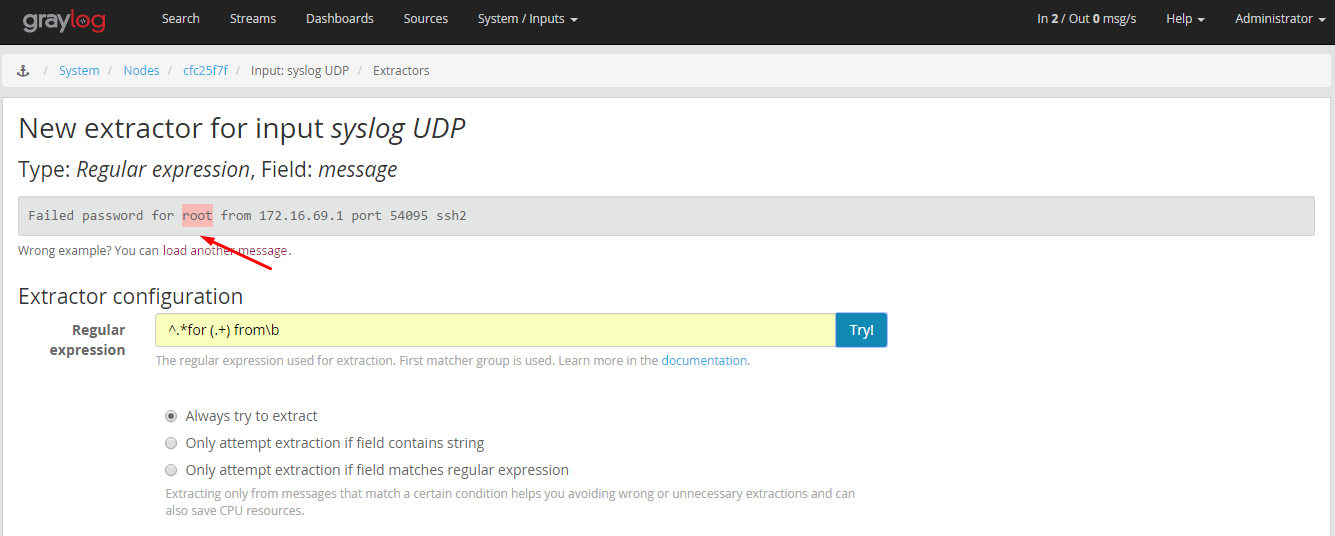
Hoặc muốn lọc user nào đanh sử dụng để ssh vào hệ thống, dùng cú pháp :

**^.\*for.\* (.+) from\b** ( Sẽ chỉ lấy từ đúng sau for và đứng trước from, bỏ qua các từ còn lại ). Ví dụ khi ssh với user **manhdv** ( user này không tồn tại trong hệ thống )

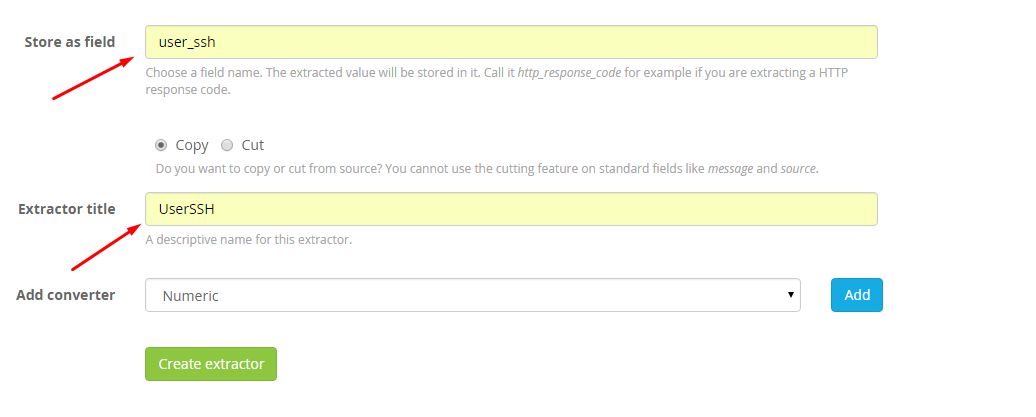
Field user\_ssh sẽ bỏ qua cụm từ **invalid user**, chỉ lấy từ **manhdv.**



Hoặc đăng nhập với user root

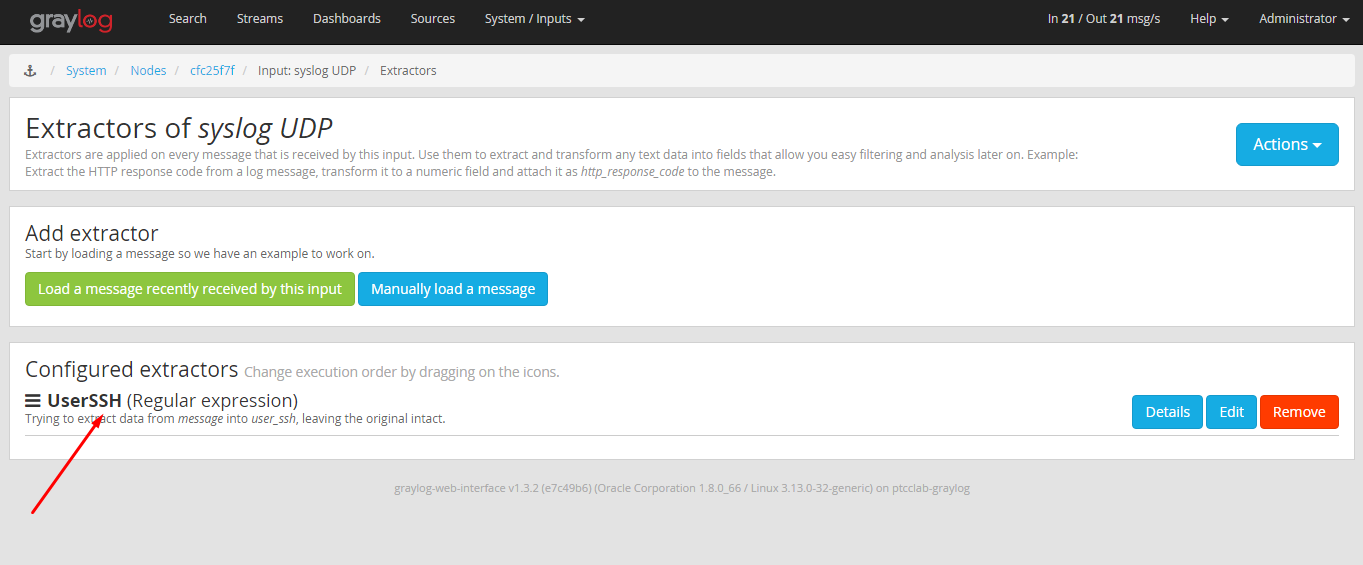


Nhập thông tin cho vùng field và Extractor title

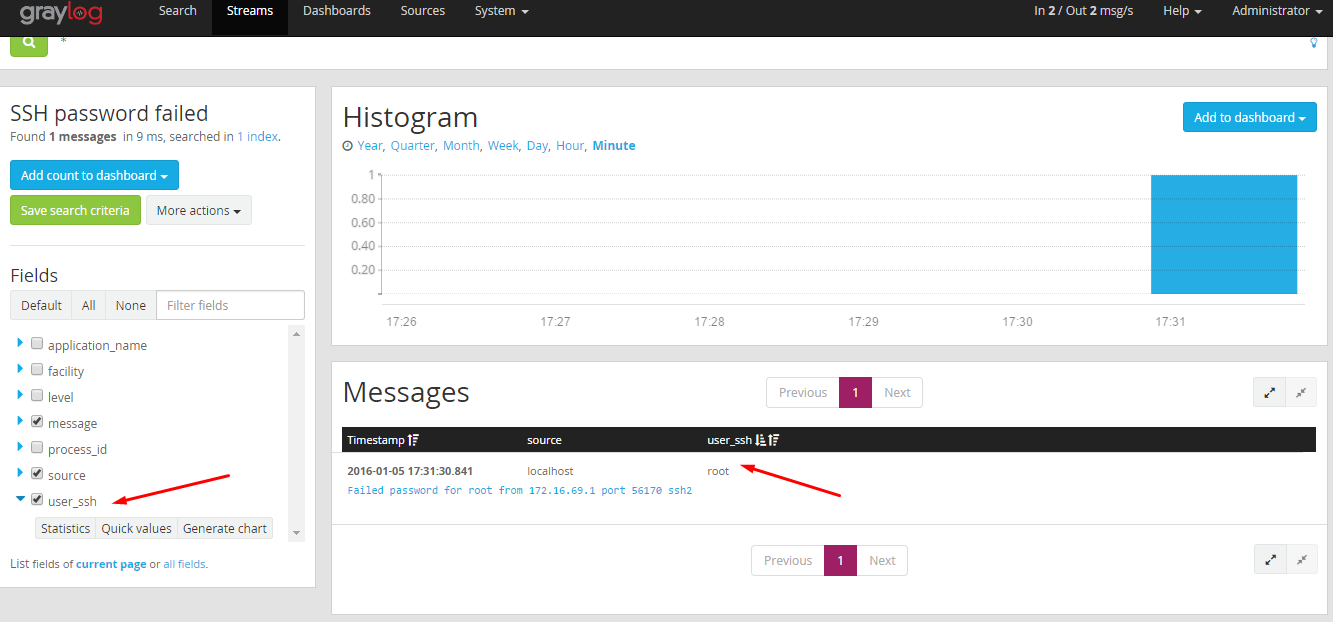


Sau khi Create extractor, kiểm tra lại phần Manage Extractors trên Input ( khi tạo Extractor từ bản tin thuộc Input nào thì Extractor sẽ được tạo ra trên Input đó ).

Regex UserSSH được tạo trên bản tin từ Input Syslog



Đã xuất hiện thêm field mới của những bản tin thuộc Input Syslog



Việc Extractor sẽ giúp người quản trị dễ dàng biết được user nào đang đăng nhập vào hệ thống khi đã tạo bảng trên Dashboard như đã hướng dẫn ở phần 4 .Dashboard

