

**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP. HCM**  
**KHOA CÔNG NGHỆ THÔNG TIN**

-----🙞🙜🕮🙞🙜-----

**ĐỒ ÁN 1**

**XÂY DỰNG HỆ THỐNG IPS**

**DÙNG HIDS-OSSEC**



GVHD: Nguyễn Hoà

Sinh viên thực hiện:  
1. Phạm Văn Duẩn (13037321)

2. Võ Thành Luân()

3. Mai Quốc Quy ()

4. …Vinh ()

*TP. Hồ Chí Minh, tháng 5 năm 201**6*

Mục lục

[I. Khái niệm về log 3](#_Toc450133349)

[II. Syslog 4](#_Toc450133350)

[**1.** **Giới thiệu syslog** 4](#_Toc450133351)

[**1.1.** **Nguồn sinh ra log** 5](#_Toc450133352)

[**1.2.** **Mức độ cảnh báo** 6](#_Toc450133353)

[**1.3.** **Định dạng chung của một gói tin syslog** 6](#_Toc450133354)

[**2.** **Cách thức lưu trữ** 7](#_Toc450133355)

[III. Log tập trung 9](#_Toc450133356)

[**1.** **Tại sao sử dụng log tập trung?** 9](#_Toc450133357)

[**2.** **Lợi ích** 9](#_Toc450133358)

[IV. Kết luận 10](#_Toc450133359)

1. **Khái niệm về log**

*Log là gì?Log để làm gì?*

Trước hết Bạn là người quản trị mạng của một doanh nghiệp, trong hệ thống mạng của bạn có một máy chủ chứa dữ liệu rất quan trọng. Một buổi tối bạn để máy chủ đó chạy suốt đêm nhưng khi về đến nhà bạn truy cập vào máy chủ thì báo lỗi từ chối dịch vụ do không thể kết nối, buổi sáng bạn vội vã đến xem xét tình hình thì thấy một số dữ liệu đã bị mất và vấn đề lúc này là xem ai đã gây ra vấn đề trên.

Vậy phải làm thế nào để điều tra xử lý, hay đơn giản là tìm nguyên nhân để khắc phục hậu quả vừa xảy ra. Log sẽ giúp bạn làm việc này.

*Vậy nên tác dụng của log là:*

* *Log ghi lại liên tục các thông báo về hoạt động của cả hệ thống hoặc của các dịch vụ được triển khai trên hệ thống và file tương ứng. Log file thường là các file văn bản thông thường dưới dạng “clear text” tức là bạn có thể dễ dàng đọc được nó, vì thế có thể sử dụng các trình soạn thảo văn bản (vi, vim, nano...) hoặc các trình xem văn bản thông thường (cat, tailf, head...) là có thể xem được file log.*
* *Các file log có thể nói cho bạn bất cứ thứ gì bạn cần biết, để giải quyết các rắc rối mà bạn gặp phải miễn là bạn biết ứng dụng nào, tiến trình nào được ghi vào log nào cụ thể.*
* *Trong hầu hết hệ thống Linux thì* ***/var/log*** *là nơi lưu lại tất cả các log.*

Như đã nói ở trên, tác dụng của log là vô cùng to lớn, nó có thể giúp quản trị viên **theo dõi hệ thống của mình tôt hơn, hoặc giải quyết các vấn đề gặp phải** với hệ thống hoặc service. Điều này đặc biệt quan trọng với các hệ thống cần phải online 24/24 để phục vụ nhu cầu của mọi người dùng.

1. **Syslog**
2. **Giới thiệu syslog**

Syslog là một giao thức client/server, là giao thức dùng để chuyển log và thông điệp đến máy nhận log. Máy nhận log thường được gọi là syslogd, syslog daemon hoặc syslog server.

Syslog có thể gửi qua UDP hoặc TCP. Các dữ liệu được gửi dạng cleartext. Syslog dùng cổng 514.

Syslog được phát triển năm 1980 bởi Eric Allman, nó là một phần của dự án Sendmail, và ban đầu chỉ được sử dụng duy nhất cho Sendmail. Nó đã thể hiện giá trị của mình và các ứng dụng khác cũng bắt đầu sử dụng nó. Syslog hiện nay trở thành giải pháp khai thác log tiêu chuẩn trên Unix-Linux cũng như trên hàng loạt các hệ điều hành khác và thường được tìm thấy trong các thiết bị mạng như router. Trong năm 2009, Internet Engineering Task Forec (IETF) đưa ra chuẩn syslog trong RFC 5424.

Syslog ban đầu sử dụng UDP, điều này là không đảm bảo cho việc truyền tin. Tuy nhiên sau đó IETF đã ban hành RFC 3195 Reliable Delivery cho syslog - đảm bảo tin cậy cho syslog và RFC 6587 Transmission of Syslog Messages over TCP - Truyền tải thông báo syslog qua TCP. Điều này có nghĩa là ngoài UDP thì giờ đây syslog cũng đã sử dụng TCP để đảm bảo an toàn cho quá trình truyền tin.

Trong chuẩn syslog, mỗi thông báo đều được dán nhãn và được gán các mức độ nghiêm trọng khác nhau. Các loại phần mềm sau có thể sinh ra thông báo: auth , authPriv , daemon , cron , ftp , dhcp , kern , mail, syslog, user, ... Với các mức độ nghiêm trọng từ cao nhất trở xuống Emergency, Alert, Critical, Error, Warning, Notice, Info, and Debug.

* 1. **Nguồn sinh ra log**

|  |  |  |
| --- | --- | --- |
| **Facility Number** | **Nguồn tạo log** | **Ý nghĩa** |
| 0 | Kernel | Những log mà do kernel sinh ra |
| 1 | User | Log ghi lại cấp độ người dùng |
| 2 | Mail | Log của hệ thống mail |
| 3 | daemon | Log của các tiến trình trên hệ thống |
| 4 | Autho | Log từ quá trình đăng nhập hệ hoặc xác thực hệ thống |
| 5 | syslog | Log từ chương trình syslogd |
| 6 | Lpr | Log từ quá trình in ấn |
| 7 | News | Thông tin từ hệ thống |
| 8 | Uucp | Log UUCP subsystem |
| 9 |  | Clock deamon |
| 10 | authpriv | Quá trình đăng nhập hoặc xác thực hệ thống |
| 11 | ftp | Log của FTP deamon |
| 12 |  | Log từ dịch vụ NTP của các subserver |
| 13 |  | Kiểm tra đăng nhập |
| 14 |  | Log cảnh báo hệ thống |
| 15 | cron | Log từ clock daemon |
| 16 | local 0 – local 7 | Log dự trữ cho sử dụng nội bộ |

* 1. **Mức độ cảnh báo**

|  |  |  |
| --- | --- | --- |
| **Code** | **Mức cảnh báo** | **Ý nghĩa** |
| 0 | emerg | Thông báo tình trạng khẩn cấp |
| 1 | alert | Hệ thống cần can thiệp ngay |
| 2 | crit | Tình trạng nguy kịch |
| 3 | error | Thông báo lỗi đối với hệ thống |
| 4 | warn | Mức cảnh báo đối với hệ thống |
| 5 | notice | Chú ý đối với hệ thống |
| 6 | info | Thông tin của hệ thống |
| 7 | debug | Quá trình kiểm tra hệ thống |

* 1. **Định dạng chung của một gói tin syslog**

Định dạng hoàn chỉnh của một thông báo syslog gồm có 3 phần chính như sau:

<PRI> HEADER MSG

Độ dài một thông báo không được vượt quá 1024 bytes.

* PRI

Phần PRI là một số được đặt trong ngoặc nhọn, thể hiện cơ sở sinh ra log hoặc mức độ nghiêm trọng. là 1 số 8bit. 3 bit đầu tiên thể hiện cho tính nghiêm trọng của thông báo. 5bit còn lại đại diện cho sơ sở sinh ra thông báo.

Giá trị Priority được tính như sau: Cơ sở sinh ra log x 8 + Mức độ nghiêm trọng. Ví dụ,thông báo từ kernel (Facility = 0) với mức độ nghiêm trọng (Severity =0) thì giá trị Priority = 0x8 + 0 = 0. Trường hợp khác, với "local use 4" (Facility =20) mức độ nghiêm trọng (Severity =5) thì số Priority là 20 x 8 + 5 = 165.

Vậy biết một số Priority thì làm thế nào để biết nguồn sinh log và mức độ nghiêm trọngcủa nó.

Ta xét 1 ví dụ sau: Priority = 191

Lấy 191:8 = 23.875 -> Facility = 23 ("local 7") -> Severity = 191 - (23\*8) = 7 (debug)

* HEADER

Phần Header thì gồm các phần chính sau

* Time stamp -- Thời gian mà thông báo được tạo ra. Thời gian này được lấy từ thời gian hệ thống ( Chú ý nếu như thời gian của server và thời gian của client khác nhau thì thông báo ghi trên log được gửi lên server là thời gian của máy client)
* Hostname hoặc IP
* Message

Phần MSG chứa một số thông tin về quá trình tạo ra thông điệp đó. Gồm 2 phần chính

* Tag field là tên chương trình tạo ra thông báo
* Content field chứa các chi tiết của thông báo

1. **Cách thức lưu trữ**

Phần lớn các bản phân phối Linux sẽ cài đặt một cấu hình syslog mặc định, bao gồm *logging to messages* và các *file log* khác trong /var/log. Để ngăn cản những file này ngày càng trở nên cồng kềnh và khó kiểm soát, một hệ thống quay vòng file log (a log file rotation scheme) nên được cài đặt. Hệ thống *cron* đưa ra các lệnh để thiết lập những file log mới, những file cũ được đổi tên bằng cách thay một con số ở hậu tố.

Với loại quay vòng này, */var/log/messages* của ngày hôm qua sẽ trở thành *messages.1* của ngày hôm nay và một *messages* mới được tạo. Sự luân phiên này được cấu hình cho một số lượng lớn các file, và các file log cũ nhất sẽ được xoá khi sự luân phiên bắt đầu chạy. Ví dụ trong /var/log của có các messages sau: messages, messages.1, messages-20160411, messages-20160418, ...

Tiện ích thi hành rotation là logrotate. Lệnh này được cấu hình sử dụng cho một hoặc nhiều files - được xác định bởi các tham số đi cùng. File cấu hình mặc định là **/etc/logrotate.conf**. Ví dụ với Ubuntu 14.04:

# see "man logrotate" for details

# rotate log files weekly

weekly

# use the syslog group by default, since this is the owning group

# of /var/log/syslog.

su root syslog

# keep 4 weeks worth of backlogs

rotate 4

# create new (empty) log files after rotating old ones

create

# uncomment this if you want your log files compressed

#compress

# packages drop log rotation information into this directory

include /etc/logrotate.d

# no packages own wtmp, or btmp -- we'll rotate them here

/var/log/wtmp {

missingok

monthly

create 0664 root utmp

rotate 1

}

/var/log/btmp {

missingok

monthly

create 0660 root utmp

rotate 1

}

Trong ví dụ này:

* Hệ thống sẽ quay vòng file log hàng tuần
* Lưu lại những thông tin log đáng giá trong 4 tuần
* Sử dụng định dạng Ngày tháng thêm vào để làm hậu tố của log files (20160411, 20160418, ...)
* Thông tin về sự quay vòng log của các gói nằm trong **/etc/logrotate.d**
* Rotation được thiết lập cho 2 files: **/var/log/wtmp** và **/var/log/btmp**

1. **Log tập trung**

Để quản lý log một cách tốt hơn, xu thế hiện nay sẽ sử dụng log tập trung.

Hiểu một cách đơn giản : Log tâp trung là quá trình tập trung, thu thập, phân tích... các log cần thiết từ nhiều nguồn khác nhau về một nơi an toàn để thuận lợi cho việc phân tích, theo dõi hệ thống.

1. **Tại sao sử dụng log tập trung?**

* Do có nhiều nguồn sinh log:
* Có nhiều nguồn sinh ra log, log nằm trên nhiều máy chủ khác nhau nên khó quản lý.
* Nội dung log không đồng nhất (Giả sử log từ nguồn 1 có có ghi thông tin về ip mà không ghi thông tin về user name đăng nhập mà log từ nguồn 2 lại có) -> khó khăn trong việc kết hợp các log với nhau để xử lý vấn đề gặp phải.
* Định dạng log cũng không đồng nhất -> khó khăn trong việc chuẩn hóa
* Đảm bảo tính toàn vẹn, bí mật, sẵn sàng của log.
* Do có nhiều các rootkit được thiết kế để xóa bỏ logs.
* Do log mới được ghi đè lên log cũ. -> Log phải được lưu trữ ở một nơi an toàn và phải có kênh truyền đủ đảm bảo tính an toàn và sẵn sàng sử dụng để phân tích hệ thống.

1. **Lợi ích**

* Giúp quản trị viên có cái nhìn chi tiết về hệ thống -> có định hướng tốt hơn về hướng giải quyết.
* Mọi hoạt động của hệ thống được ghi lại và lưu trữ ở một nơi an toàn (log server) -> đảm bảo tính toàn vẹn phục vụ cho quá trình phân tích điều tra các cuộc tấn công vào hệ thống
* Log tập trung kết hợp với các ứng dụng thu thập và phân tích log khác nữa giúp cho việc phân tích log trở nên thuận lợi hơn -> giảm thiểu nguồn nhân lực.

1. **Kết luận**

Syslog là một công cụ mạnh mẽ làm cho việc quản trị mạng dễ dàng hơn đối với quản trị viên

Thách thức lớn nhất là khối lượng dữ liệu

**Tài liệu tham khảo**

1. Một vài hiểu biết về log

<https://github.com/hocchudong/Mot-vai-hieu-biet-ve-log>

1. Cấu hình và sử dụng các file log của hệ thống

<http://congdonglinux.vn/forum/showthread.php?580-C%E1%BA%A5u-h%C3%ACnh-v%C3%A0-s%E1%BB%AD-d%E1%BB%A5ng-c%C3%A1c-log-files-c%E1%BB%A7a-h%E1%BB%87-th%E1%BB%91ng>

1. What is syslog?

<http://www.networkmanagementsoftware.com/what-is-syslog/>