**Tìm hiểu và triển khai hệ thống thống kê AWStats**

**1. Danh sách thành viên và công việc**

|  |  |  |  |
| --- | --- | --- | --- |
| Họ và Tên | Mã sinh viên | Công việc | Tiến độ |
| Nguyễn Kỳ Anh | 175A071284 | Cài đặt demo AWStast  bao gồm:  +Hướng dẫn sử dụng  +Quản trị  +Thiết lập cron để cập nhập nhật ký | -Cài đặt demo AWStats |
| Nguyễn Đức Trung | 175A071327 | -Tìm hiểu về AWStasts  -Phân tích ưu , nhược điểm của AWStast  -So sánh AWStats với những hệ thống thông kê khác | -Tìm hiểu về AWStats  -Phân tích ưu, nhược điểm  -So sánh Awstast với Webalizer |

**2. Nội dung nghiên cứu**

**a. Tổng quan về AWStats**

AWStats là một công cụ báo cáo phân tích Web có mã nguồn mở, phù hợp để phân tích dữ liệu từ các dịch vụ trên mạng như web, các phương tiện truyền thông, mail và các máy chủ FTP.

**b. Chức năng tổng thể**

AWStats phân tách và nhận diện các logfile máy chủ, tạo báo cáo HTML. Dữ liệu được trình bày trực quan trong các báo cáo bằng bảng và biểu đồ thanh. Báo cáo tĩnh có thể được tạo thông qua giao diện dòng lệnh và báo cáo theo yêu cầu được hỗ trợ thông qua chương trình CGI của trình duyệt Web.

AWStats hỗ trợ hầu hết các định dạng tệp log máy chủ web chính bao gồm định dạng log Apache (NCSA kết hợp / XLF / ELF hoặc Định dạng log chung (CLF)), Webstar, IIS (định dạng log W3C) và nhiều định dạng log máy chủ web phổ biến khác.  
  
Sự phát triển công cụ đã được chuyển từ SourceForge sang GitHub vào năm 2014.

**c. Ưu nhược điểm**

Ưu điểm

* Giao diện dễ nhìn nhờ vào việc hiển thị dữ liệu đơn giản và theo kiểu biểu đồ.
* Phân tích được nhiều loại web log: Apache, IIS, Web, Proxy, Mail log....
* Có thể mở rộng các module bằng plug-in.
* Cơ sở dữ liệu có thể lưu dưới dạng XML do dố có khả năng sử dụng làm input cho các ứng dụng phân tích khác.
* Chạy trên tất cả các nền tảng hỗ trợ Perl

Nhược điểm

* Chỉ phân tích dữ liệu theo các tháng (Các phiên bản mới hơn đã khác phục được).
* Vì AWStats truy cập được vào server log file nên có khả năng xuất hiện vấn đề về bảo mật (Lupper Worm tấn công được web server thông qua lỗ hổng bảo mật trên AWStats)
* Tốn tài nguyên khi chạy liên tục, vì mỗi lần chạy AWStats yêu cầu lấy dữ liệu từ web server, trong trường hợp server có nhiều website và dữ liệu quá trình này sẽ tốn rất nhiều tài nguyên.
* Dữ liệu hiển thị là tĩnh, người dùng không thể thay đổi hay tùy chỉnh đối tượng so sánh
* Không hỗ trợ Export

**d. So sánh AWStats với những hệ thống thông kê khác**

**-So sánh Awstats với Webalizer**

|  |  |  |
| --- | --- | --- |
| Tính năng/Phần mềm | AWStats | Webalizer |
| Ngôn ngữ | Perl | C |
| Tương thích với tất cả các nền tảng | Có | Có |
| Có thể đọc mã nguồn | Có | Có |
| Giá | Free | Free |
| Hoạt động với Apache dạng  XLF/ELF | Có | Có |
| Hoạt động với Apache dạng CLF | Tất cả các tính năng có sẵn với định dạng nhật ký (b) | Tất cả các tính năng có sẵn với định dạng nhật ký (b) |
| Hoạt động với IIS (W3C) | Có | Cần một bản vá |
| Hoạt động với Log format cá nhân hóa | Có | Không |
| Phân tích Web/FTP/Mail log files | Có/Có/Có | Có/Không/Không |
| Báo cáo và cập nhập số liệu thống kê từ | CLI/CGI | CLI |
| Báo cáo truy cập đặc biệt | Có | Không |
| Báo cáo thời lượng phiên làm việc | Có | Không |
| Báo cáo quốc gia | Từ vị trí IP hoặc tên miền | Tên miên |
| Báo cáo thành phố và các quốc gia lớn | Cần CSDL Maxmind Cities | Không |
| Report ISP | Cần CSDL Maxmind ISP | Không |
| Báo cáo kích thước màn hình | Có | Không |
| Báo cáo loại files | Có | Không |
| Báo cáo kích thước files | Không | Không |

**AWStats**  
AWStats không phải là một công cụ tiếp thị và sử dụng các tệp nhật ký máy chủ để giải thích dữ liệu. Nó cung cấp một ý tưởng tốt hơn về lượng sử dụng băng thông và thông tin về khách truy cập. Nó chủ yếu được sử dụng bởi các quản trị viên mạng.  
  
AWStats xác định số lượng khách truy cập như sau:

1. AWStats tạo các tệp nhật ký máy chủ và sử dụng địa chỉ IP và tác nhân người dùng để xác định xem khách truy cập là người hay robot.

2. Nó ghi lại các lượt truy cập mới trong khung thời gian là 60 phút.

3.  Ngay cả khi AWStats giữ một bản ghi liên quan đến các robot đã biết, một số bot không thể được xác định được.

4.   AWStats có thể được sử dụng để kiểm tra băng thông được sử dụng bởi nhiều robot và trình thu thập thông tin khác nhau và sử dụng dữ liệu này để xác định xem khách truy cập có phải là người hay không.  
  
**Webalizer**  
  
Webalizer không phải là một công cụ tiếp thị và chức năng để nắm bắt lưu lượng truy cập mà không phân biệt giữa người và robot. Nó cung cấp một cửa sổ khách truy cập trong 30 phút.  
  
Webalizer xác định số lượng khách truy cập như sau:

1.Webalizer diễn giải các tệp nhật ký máy chủ và tạo các báo cáo tương ứng.

2.Nó cung cấp một cửa sổ 30 phút cho khách truy cập và ghi lại các lượt truy cập trong khoảng thời gian này.

3.Theo mặc định, nó đối xử với robot và khách truy cập của con người như nhau, nhưng nó có thể được thiết lập để bỏ qua các bot nhất định. Nhưng, ngay cả sau khi thiết lập như vậy, tất cả các bot đều có thể được phát hiện, và do đó tốt hơn là bỏ qua nó.  
  
**Webalizer so với AWStats**  
Webalizer cung cấp một cửa sổ khách truy cập trong 30 phút, trong khi cửa sổ là 60 phút trong trường hợp AWStats. Do đó, các báo cáo được tạo bởi Webalizer cao hơn.  
  
Webalizer theo mặc định không phân biệt giữa khách truy cập người và robot, trong khi các báo cáo do AWStats tạo sẽ cung cấp sự hiểu biết tốt giữa khách truy cập người và robot.  
  
AWStats hiển thị số lượng cao hơn và trình bày dữ liệu tốt hơn so với Webalizer.

**e. Hướng dẫn cài đặt AWStats trên Ubuntu**

-Vì Awstast đã có trong kho ứng dụng của Ubuntu nên ta sẽ cài đặt bằng cách gõ lệnh

sudo apt-get install awstast



Dựa vào số liệu trên có thể thấy phiên bản hiện tại của Awstats là (7.6+dfsg-2)

-Sau khi cài đặt xong Awstats ta cần định cấu hình cho tiên miền mình muốn thống kê , ví dụ là "[www.kyanh.com](http://www.kyanh.com/)"

-Thực hiện việc này bằng cách sao chép cấu hình mặc định của Awstats vào tệp có tên miền

''[www.kyanh.com](http://www.kyanh.com/)"

-Tệp cấu hình mặc định của Awstats nằm trong /etc/Awstats

-Đầu tiên ta đi vào tệp chứa cấu hình mặc định của Awstats bằng lệnh

cd /etc/awstats



-Sử dụng lệnh ls ta có thể thấy được tệp awstats.conf là tệp cấu hình mặc định của Awstats



-Tiếp theo ta sẽ tạo ra 1 tệp có nội dung giống với tệp có cấu hình mặc định của Awstast bằng lệnh cp

sudo cp awstats.conf awstats.kyanh.com.conf



-Sau khi tạo xong awstats.kyanh.com.conf ta sẽ vào tệp tin đó để định lại cấu hình bằng lệnh

sudo nano awstats.kyanh.com.conf

-Sau khi thực hiện lệnh nano đây chính là nội dung của tệp cấu hình Awstats



-Chúng ta sẽ chỉnh lại một số nội dung như là :

a.SiteDowmain



Tên miền của trang web sẽ là [www.kyanh.com](http://www.kyanh.com/)

b.LogFormat



LogFormat=1 tức là chỉ cần 1 file log cho toàn bộ trang web

c.AllowToUpdateStatsFromBrowser



Cho phép cập nhập thông tin từ duyệt trình

-Sau khi chỉnh lại một số nội dung quan trong ta sẽ nhấn Ctrl+O để lưu file và Ctrl+X để thoát

-Cài đặt apache2

Tương tự như Awstats, apache2 cũng nằm trong kho ứng dụng của Awstats nên ta chỉ cần gõ lệnh

sudo apt-get install apache2



-Kiểm tra xem tình trạng của apache2 bằng lệnh

systemctl status apache2



Chúng ta chỉ cần chú ý đến chữ running nghĩa là apache2 đang hoạt động

-Kích hoạt mô đun cgi trong apache2 bằng lệnh

sudo a2enmod cgi



Mô đun cgi cho phép khởi tạo giao tiếp giữa server và chương trình

-Sau khi kích hoạt xong mô dun cgi ta khởi động lại apache2 bằng lệnh

systemctl restart apache2



-Định cấu hình Apache2 để hiển thị số liệu thống kê

để làm được việc này cần sao chép nội dung thư mục "cgi-bin" vào thư mục gốc của tài liệu trong tiến trình cài đặt apache2 bằng lệnh

sudo cp -r awstast.pl /var/www/html/



-Sau khi copy xong ta sẽ thay đổi quyền sở hữu



-Đổi quyền sở hữu xong ta sẽ đổi quyền thao tác trên tệp tin



-Định cấu hình apache2 xong ta sẽ cập nhập lại awstast bằng lệnh

/usr/lib/cgi-bin/awstats.pl -config = kyanh.com-update



đây là giao diện thông báo sau khi update thành công

-Cập nhập xong ta sẽ kiểm tra Awstast trên duyệt trình bằng đường dẫn

<http://192.168.1.189/cgi-bin/awstats.pl?config=kyanh.com>

-192.168.1.189 chính là ip của server



-Đây là giao diện của website khi truy cập vào đường dẫn bên trên

**f. Hướng dẫn Sử dụng/Quản trị**