

How to build detailed Network Usage Reports using RRDTool, flow-tools, FlowScan, and CUFlow

Preface

What does this document cover?

This document is aimed at providing step by step instructions to build useful documentation and reports from NetFlow "flows" on Cisco routers. Other vendors have their own implementations of NetFlow. You should be able to use this document to build reports from those devices, but I am a Cisco expert, so this document assumes that you are using Cisco products.

Here's a list of what you will be able to do with this application:

- Build graphs showing network utilization, including break downs by router, protocol, service, and network/host groups.
- Build "top talker" reports for your network. These are EXTREMELY useful for just about everything.
- Track down users and computers that are abusing your network, including virus infections and DOS attacks.
- Impress the boss with colorful graphs! :-)

Assumptions and Support

This document assumes that you are familiar with getting around a Unix/Linux system and a Cisco Router. I used to be a RedHat Linux person, but they decided to go a different direction so I did too. Currently I use either SuSE or CentOS. CentOS is a rebuild from source of RedHat Enterprise Linux. (Links are in the appendix.) This document will assume that you are using CentOS or another RHEL3 compliant Linux build. I do not claim to be all-knowing when it comes to Linux. I am sure that the way I do things is not the only way they can be done. This is simply the way I have done things and it works for me. You don't need to be a CCIE or a Linux guru to make this work, but you will be compiling packages and modifying configuration files from the command line, so be prepared. Just to warn you, you should also be prepared to spend at least a couple of hours to complete this. However, once it's up and running, you shouldn't have to touch it except to update the CUFlow configuration.

These instructions come with no warranty or guarantee. If you blow something up and lose business because of it, that's your problem. I do not provide support for these

packages or instructions. There are mailing lists for each one of the packages I use, please use them if you need additional help. I participate on several of them, so please do not e-mail me directly. I cannot promise that I will get back to you if you do. A list of these mailing lists is provided in the appendix.

I am not aware of any commercial support offered for RRDTool, flow-tools, FlowScan or CUFlow. If you know of any, please let me know at netflow and I will update this document.

Acknowledgements

This document is primarily an abridgement of the documentation provided by the authors of the packages we will be using. It is not meant to replace the documentation that comes with those packages. I have given credit to those individuals who have contributed to this document. If you feel that I have not given appropriate credit to someone, please let me know and I will correct my error.

Getting Started

Before proceeding with any of the installs or configuration steps, I suggest that you read at least that section. It's really best to read through this entire document before you start, but I too am known to be impatient from time to time. At least read each section before starting on the steps included in that section. Ready? Set? Let's go!

Routers Support

Most Cisco routers, all the way down to the 806, support NetFlow exporting. To see if your router and IOS version supports NetFlow, please visit <http://www.cisco.com/go/fn> and search for "NetFlow". Again, many other networking vendors have their own implementation of NetFlow. Please visit each vendor's respective site for more detailed information. If they support NetFlow version 5, this setup should be compatible.

For the purposes of this setup, you will need to use NetFlow version 5. The flow-tools collector is capable of handling newer version of NetFlow and storing them in its single unified file format that FlowScan can understand. This is useful, for instance, when collecting flows from a Cisco Catalyst 65xx, which uses a different version by default. (The 65xx can also be setup for version 5. See the appendix for the suggested configuration.) Please consult the FlowScan and flow-tools mailing lists listed in the appendix for more information about non-version 5 setups. Setting up FlowScan to use versions of NetFlow other than version 5 is outside of the scope of this document.

The Flow Collector and Reporting System

The Flow Collector can be any Unix/Linux based system. As I mentioned previously, I prefer CentOS for the operating system. I have found that it is best to have a dedicated server because the report generation can take a lot of processing time. You can use just about anything, but **IT MUST** process the flow files in under five minutes. I suggest that you aim for less than 2.5 to 3 minutes to better handle unusual traffic patterns such as virus infections and DOS attacks. Most of the virulent viruses these days use networks to spread. One computer with a virus can open thousands of connections per second. The flow reporting from the router will show these connections, even if they are not successful. (Use the TopTalkers reports to track down the IPs.)

I currently have several NetFlow Reporting systems in production. The "big bad" system is at our Primary Network Operations Center. Others include remote locations that support dialup clients, and my house. This section should give you some idea of how this will scale.

The main NOC system produces reports for two Cisco 7513 routers and one 7206 router. Each router has one Internet DS-3. During peak hours of traffic, we push about 80-100 Megabits (Mbps) per second total in each direction (to and from the Internet). This creates about 1.7 million flows every five minute for a file size of 23.7 MB. The collector is a Dual Processor AMD Athlon MP 2800+ with 1024 MB of RAM and an IDE disk drive. It takes about 160 seconds to process the flows and create the reports. ***The graphs are produced on the fly when someone accesses them. This is a feature of RRDTTool.*** Although this version of FlowScan does not use both processors to process the flow files, it is still useful to have the second processor to handle everything else.

One of the other locations produces reports for a single Cisco 3660 router that has a fractional DS-3 Internet backbone. During peak hours, we can reach up to the full 9 Mbps inbound and about 5 Mbps outbound. This creates about 175,000 flows every five minute for a file size of about 2-3 MB. The collector is a Pentium III 500 with 256 MB of RAM and an IDE disk drive. It takes about 70 seconds to process the flows and create the reports.

Software Packages

Here is the list of packages you will need. I suggest that you download them all before proceeding any further. It does not matter where you download them to, just don't misplace them. I usually create a folder at "/var/netflow/install" to keep the source packages. Also, please take note the version numbers may change from what is listed here. Be sure to grab the latest stable versions when you download the packages.

- **Apache** - You will need a web server to view the graphs and reports from this application. Any web server that supports CGI scripting will be fine. If you do not already have a web server installed, I suggest Apache. It can be downloaded from

- <http://httpd.apache.org>. The default install will work just fine. For detailed instructions, please see the documentation on the Apache site. If you did not install the Web Server packages when you installed CentOS, you can install them using the yum utility. To install apache on CentOS use:
- yum install httpd
 - **Perl5** - This is installed by default in most builds of Linux. If you don't have it, please visit www.perl.com or www.cpan.org, or simply rebuild the machine and make sure that you install the Perl packages. (Installing the packages during install is the preferred method.) If you don't already have Perl, or don't know what it is, this application is probably not for you.
 - **RRDTool** - This package can be downloaded from www.rrdtool.org. It is recommended that you install from the source tarball. When you configure and compile the package be sure to use the --enable-shared option. Most of the pre-compiled packages (RPMs) do not have this enabled, so if you get RRDTool errors and are using an RPM version, try manually compiling as described here. I install it using these commands:
 - tar -zxvf rrdtool-1.0.49.tar.gz
 - cd rrdtool-1.0.49
 - ./configure --enable-shared --prefix=/usr/local/rrdtool
 - make
 - make install
 - make site-perl-install
 - **flow-tools** - This is the collection of programs that includes the collector application I prefer. It can be downloaded from: <http://www.splintered.net/sw/flow-tools/>. You can install it using these commands:
 - tar -zxvf flow-tools-0.66.tar.gz
 - cd flow-tools-0.66
 - ./configure
 - make
 - make install

This will install flow-tools to /usr/local/netflow. Remember this folder. There are many tools included in flow-tools. We will only use flow-capture in this document, but you may find others of use in your situation.

- **Perl Modules** - In addition to Perl5, you will need the modules listed below. To install all of them, except Cflow which is packaged with flow-tools, follow the instructions below.

Notes about Perl: If you are running the CPAN shell for the first time, you will be asked to configure it. To Auto-Configure CPAN (I recommend that you let it auto-configure) enter "no". Once it is configured it will download a database file. This may take several minutes depending on your Internet connection. Also, take care to note that *everything* Perl is case sensitive. CPAN may also alert you to install a newer version of CPAN. The newer versions are valuable, but be warned;

it will take time to install. Lastly, you must be **root** to install the packages. Type the following commands to use CPAN.

- perl -MCPAN -e shell
- install HTML::Table
- install Net::Patricia
- install Boulder::Stream (I had to do a "force install Boulder::Stream" last time I installed this. If you have compile errors, try that.)
- If you are running a current Linux distribution such as CentOS, don't bother trying to install this module in CPAN. Do the following instead:
 - Go to <http://www.cpan.org>
 - Search for ConfigReader
 - Choose ConfigReader-0.5
 - Download the tarball and unpack it (tar -zxvf ConfigReader-0.5.tar.gz)
 - cd ConfigReader-0.5
 - mkdir -p /usr/lib/perl5/site_perl/5.8.0/ConfigReader
 - cp * /usr/lib/perl5/site_perl/5.8.0/ConfigReader*(Update the path to match your version of Perl.)* The above is from CentOS 3.4 with Perl 5.8.0.
- If you are running RedHat 7.2 or earlier, this should work instead of the above steps. *install ConfigReader::DirectiveStyle*

CFlow is included in the current distribution of flow-tools so you do not need to download it separately. Please install it by doing the following:

- cd flow-tools-0.66
- cd contrib
- tar -zxvf Cflow-1.051.tar.gz
- cd Cflow-1.051
- perl Makefile.PL
- make
- make install

*****IMPORTANT***** *(No, really. It's that important that you understand this. This is the #1 FAQ and problem when installing this application.)*

In order for this module to compile with the proper support, you need to run it from a directory under the flow-tools distribution files. The instructions above will build Cflow properly. You **WILL** get errors when you run FlowScan if you do not follow these steps exactly. Also, if you install flow-tools from an RPM package, you'll need to grab the flow-tools source package to get Cflow to work correctly. There is a README file included in the Cflow-1.051 if you'd like more information.

- **Korn Shell** - This is required by CUFlow. pdksh works just as well. On a CentOS box, simply type "yum install pdksh" and CentOS will install this for you. You can also download the source from <http://web.cs.mun.ca/~michael/pdksh/>.
 - **FlowScan** - This is the base report generating application by Dave Plonka. It can be downloaded from: <http://net.doit.wisc.edu/~plonka/FlowScan/>.
 - **CUFlow** - This is the report module and graph generator written by Columbia University for FlowScan. It can be downloaded from: <http://www.columbia.edu/acis/networks/advanced/CUFlow/>.
 - **My Support Files** - This includes the scripts and the updated FlowScan.pm module that are needed to complete the installation using this document. <http://www.linuxgeek.org/netflow/support-files-1.0.tar.gz>
-

Configure Your Routers

***Disclaimer: Make the changes to your routers at your own risk. I recommend that you establish a baseline for processor and memory utilization before making changes to your routers and reexamine the baseline after making the changes. If you have a support contract with Cisco, I suggest opening a case with TAC to have them look at your configuration and determine if these are the best commands for your routers.**

The following commands are the commands that I used to configure my routers running 12.2 and 12.3.

- These are the global configuration mode commands:
 - ip flow-export version 5 peer-as
 - ip flow-export source-interface xxx
 - **xxx** is the source interface. Choose the interface closest to your collector. This simply ensures that there is no confusion as to the source address that will be listed in the flows.
 - ***This command may be "ip flow-export source xxx" in newer versions of IOS.***
 - ip flow-export destination x.x.x.x y
 - **x.x.x.x** is the collector's ip address, **y** is the port you will specify in the flow-capture command line. You may choose any port, just remember what it is and avoid the obvious registered ports like 80. (The flow packets are UDP.)
 - ip flow-cache timeout active 1
 - This syntax is for IOS 12.2 and later. If you are running an 11.x or 12.0/12.1 code, the syntax would be: "ip flow-cache active-timeout 1". This command ensures the timely delivery of flows to the collector.

- In the interface configuration mode of **each major** interface: (major as opposed to sub-interface)
 - ip route-cache flow

I have found that if you do not run NetFlow on every major interface, it does strange things to the flow reports. *Again, consult with Cisco before changing the configuration on a production router.*

A good note from Dave Plonka:

"NetFlow isn't really a switching mode any more; it's just a means of reporting traffic. CEF is used when NetFlow is configured. NetFlow is just configured in this way for historical reasons as it was once proposed and implemented to be a forwarding enhancement."

House Keeping Stuff

You are going to need to add RRDTool to the path variable of your collector. (You're also adding flow-tools.) The following commands will add the paths you need to your current session:

- export PATH=\$PATH:/usr/local/rrdtool/bin
- export PATH=\$PATH:/usr/local/netflow/bin

To add RRDTool and flow-tools to the bootup path on RedHat do the following: (This will ensure that the path remains when you reboot.)

- Open /etc/profile with your favorite editor
- Add the following lines below "pathmunge /usr/local/sbin"
- pathmunge /usr/local/rrdtool/bin
- pathmunge /usr/local/netflow/bin
- Save and Exit

So, the file should look like this when you are done. (As far as I know the "tabs" at the beginning of the pathmunge lines are optional, but it's always good to make it look nice and readable.)

```
# Path manipulation
if [ `id -u` = 0 ]; then
pathmunge /sbin
pathmunge /usr/sbin
pathmunge /usr/local/sbin
pathmunge /usr/local/rrdtool/bin
```

```
pathmunge /usr/local/netflow/bin  
fi
```

Using the instructions above, you will add flow-tools and RRDTool to the path of your current session (so you don't have to reboot) and also add it to the permanent path so its still there the next time you reboot.

If you are running iptables or some other firewall, this is a great time to make sure that you have opened the port you will use to receive data from your router. Remember it is a UDP port, not a TCP port. The default is 2055.

Configure flow-tools

I will not try to explain all the flow-tools programs here. There is plenty of documentation that comes with the package and on their web site. The only program we are concerned with right now is "flow-capture". Everything you need to configure with flow-capture is part of the command line, that's why I like to use it over cflowd. Please don't get the Cflow module and cflowd confused. They are different things. You need Cflow, but we use flow-capture instead of cflowd.

There is one important consideration when planning your installation that we need to discuss before proceeding. If you are not using a dedicated server to collect your flows, you should strongly consider using a separate file system (not /var) for the NetFlow folders. If you are running on a shared server, the flow files could potentially cause issues for other packages by using too much space. Here are a couple of suggestions if you are running on a shared server. 1. Mount /var/netflow on a separate hard drive. Even though this will appear under the same file system in the file tree, it will limit /var/netflow to the amount of space on the new drive. 2. Implement a disk quota to keep /var/netflow from taking over the entire disk. Consult the documentation for your operating system for more information on limiting disk usage. The flow-capture tool also manages the size of the directory that it writes the flow files to. If you follow this document to the "T", you shouldn't have to worry about this, but it is good to keep disk space in mind anyway.

Before we start flow-capture, we need to add a script that will create a symbolic link to the current flow file for FlowScan to process. This script is included in the [Support Files Tarball](#). If you downloaded the support files, copy linkme to /usr/local/netflow/bin. Otherwise, using your favorite editor, create the following file:

```
/usr/local/netflow/bin/linkme
```

Put the following script in this file:

```
#!/usr/bin/perl
```

```

$base = "/var/netflow";

if ($ARGV[0] =~ /\.*/[\^]*(ft-v05[\^\\]*$)/) {
$fileName = $1;
} else {
print "Must specify file\n";
exit 1;
}

unless ( symlink("$base/ft/$fileName","$base/$fileName") ) {
print "Unable to create symbolic link: $base/$fileName\n";
exit 1;
}

```

By using this script, flow-capture can maintain the size of the directory /var/netflow/ft. You will need to make sure this file is executable. You can do with the following command:

- `chmod a+x /usr/local/netflow/bin/linkme`

Now, let's get to the install. First, you need to create a folder to store your flow files. I use /var/netflow for the base directory. This will also be the prefix when you install FlowScan. You will need to create the following directory tree for the raw flow-tools files, rrd files and the toptalkers HTML files.

```

mkdir -p /var/netflow/
mkdir -p /var/netflow/bin
mkdir -p /var/netflow/ft
mkdir -p /var/netflow/rrds
mkdir -p /var/netflow/scoreboard

```

Here is the command to start flow-capture. A sample init script is included with the Support Files Tarball. (Note the case of the options. They are case sensitive)

```

/usr/local/netflow/bin/flow-capture -w /var/netflow/ft 0/0/2055 -S5 -V5 -E1G -n 287 -N 0
-R /usr/local/netflow/bin/linkme

```

Please see the flow-capture man pages for details on what each option is. The main ones that you may want to change are:

- `-w /var/netflow/ft` - This is where flow-capture will store the flow files. You may want to change this for several reasons that we have already discussed.
- `0/0/2055` - This specifies the localip, remoteip, and port in that order. 0 in the local and remote IP spaces represents any IP. You may want to put the router's source interface IP here to make sure that no one can pollute the flows from

somewhere else. NOTE: Using zero for the remote IP is the ONLY way you can have one flow-capture process capture flows from multiple routers.

- E1G - This is the expire size setting. Basically this controls how much raw flow data you want to save. 1G represents 1 Gigabyte. 20M would represent 20 Megabytes. (The other option for saving flow files is to create a directory called "saved" in /var/netflow. The advantage to using the linkme script and this option is that flow-capture will manage the size of your saved data so that it doesn't overrun your hard drive.)
- The other option here that gets lots of questions is -n. What does the 287 mean? It means that flow-capture will rotate the file 287 times in 24 hours. For those that don't want to do the math, this is every five minutes starting at 00:00 and ending with 23:55. A good note to make here is that CUFlow doesn't currently play nice with flow files that do not represent five minutes of traffic. So do not change this setting unless you know exactly what you are doing.

Once flow-capture is running there are several things you can do to verify that it is receiving packets from the router. First, use tcpdump to see the incoming packets. If you do not have tcpdump (it is not installed by default), you can again use yum to install it:

- yum install tcpdump

The command to run tcpdump and see only UDP port 2055 is:

```
tcpdump -n udp port 2055
```

The output will look something like this:

```
03:30:13.928242 192.168.10.1.57218 > 192.168.5.3.2055: udp 264
03:30:16.392830 192.168.5.1.51892 > 192.168.5.7.2055: udp 456
03:30:25.920687 192.168.10.1.57218 > 192.168.5.3.2055: udp 312
03:30:28.393009 192.168.5.1.51892 > 192.168.5.7.2055: udp 456
```

Then use netstat -lnp to see if flow-capture is listening to port 2055. If you don't see either, check /var/log/messages for errors.

The last thing to verify is, look in /var/netflow/ft and see if there has been a tmp file created. If there is, you are good to go.

To verify that the "linkme" script is working correctly, let flow-capture run for a few minutes then look in /var/netflow/ for a new symbolic link to be created. (The file name will be ft-v05.somedate.sometime.) If it is there, everything is working. These links point to the real files. Once FlowScan has processed the real file, it will delete the link. Remember that flow-capture will manage the size of the /var/netflow/ft directory, so you shouldn't have to worry about using up all of your disk space.

Here is the init script to start flow-capture. (This is also included in the Support Files Tarball.)

```
#!/bin/sh
# description: Start Flow-Capture
# chkconfig: 2345 95 00

case "$1" in
'start')

/usr/local/netflow/bin/flow-capture -w /var/netflow/ft 0/0/2055 -S5 -V5
-E1G -n 287 -N 0 -R /usr/local/netflow/bin/linkme
touch /var/lock/subsys/startflows
;;
'stop')

killall -9 /usr/local/netflow/bin/flow-capture
rm -f /var/lock/subsys/startflows
;;

*)

echo "Usage: $0 { start | stop }"
;;

esac
exit 0
```

Go ahead and start this and let it run while we finish ("service flow-capture start" or "/etc/init.d/flow-capture start" will do the trick). This way when you get done with the install, you will have live data to process and verify that things are working!

Install FlowScan

Before proceeding with installing FlowScan, be sure that you have downloaded and installed the needed Perl Modules listed previously. If you have, here are the commands to install FlowScan:

As Dave Plonka notes, "A good way to avoid doing something dumb here is to not run FlowScan's configure nor make as root."

```
./configure --prefix=/var/netflow
make
```

```
make -n install
make install
cd cf
cp flowscan.cf /var/netflow/bin
```

The last two steps copy the FlowScan config file to /var/netflow/bin. Also, make sure that you do both "make -n install" and "make install".

You will need to download a patched FlowScan.pm and copy it to /var/netflow/bin as well. It is included with the support files that you downloaded earlier, or you can download the file from either of the following locations, then just copy it to /var/netflow/bin and replace the one that is there.

- <http://net.doit.wisc.edu/~plonka/list/flowscan/archive/att-0848/01-FlowScan.pm>
- <http://www.linuxgeek.org/netflow/FlowScan.pm>

Also from Dave Plonka's notes, "By the way, in the above commands, all is OK if make says ``Nothing to be done for `target'". As long as make completes without an error, all is OK." I normally see this message when I run "make". The above commands should have installed FlowScan to /var/netflow/bin/. This is where the application and the configuration files reside. You should also verify when you copy the new FlowScan.pm module that it is executable. The easiest way to make sure is do an "ll" while in /var/netflow/bin. You should see something like this: (Normally the file name will also be green if you have a color terminal setup.)

```
total 180
-rwxr-xr-x 1 root root 3318 Jul 31 02:51
add_ds.pl
-rwxr-xr-x 1 root root 2520 Jul 31 02:51 add_tsrx
-rwxr-xr-x 1 root root 70096 Jul 31 02:51
CampusIO.pm
-rw-r--r-- 1 root root 1692 Jul 31 02:56
CUFlow.cf
-rw-r--r-- 1 root root 43533 Jul 31 02:53
CUFlow.pm
-rwxr-xr-x 1 root root 834 Jul 31 02:51
event2vrule
-rwxr-xr-x 1 root root 5098 Jul 31 02:51 flowscan
-r--r--r-- 1 root root 630 Jul 31 02:54
flowscan.cf
-rwxr-xr-x 1 root root 8695 Jul 31 02:52
FlowScan.pm
-rwxr-xr-x 1 root root 2407 Jul 31 02:51
ip2hostname
-rwxr-xr-x 1 root root 1442 Jul 31 02:51 locker
-rwxr-xr-x 1 root root 9130 Jul 31 02:51
SubNetIO.pm
```

If you do not see the "x" attribute, use "chmod a+x FlowScan.pm" to correct the attributes.

Install CUFlow

Unpack the file you downloaded previously (tar -zxvf CUFlow-1.4.tar.gz) and then move the files CUFlow.pm and CUFlow.cf to /var/netflow/bin/. Now, go to /var/netflow/bin and edit flowscan.cf. You need to comment out any existing "ReportClasses" lines and add the following line:

```
ReportClasses CUFlow
```

You will also need to change file name listed in the "FlowFileGlob" variable. It is best to specify the full path to the flow file links (remember the linkme script?) Use: "FlowFileGlob /var/netflow/ft-v05.*". (Don't forget the * on the end, otherwise FlowScan will complain loudly.)

NOTE: You do not need to modify any of the other FlowScan files, such as CampusIO.cf. We are not using those reports in this setup. If you want to use the built in FlowScan reports, please read the FlowScan documentation at: <http://net.doit.wisc.edu/~plonka/FlowScan/INSTALL.html> for instructions on installing and configuring them.

Configure CUFlow

The most detailed documentation on CUFlow can be found at: <http://www.columbia.edu/acis/networks/advanced/CUFlow/CUFlow.html>. I highly suggest reading this document in its entirety. It goes into far greater detail than I will. There are several things that you will need to change in the CUFlow.cf file before running any reports.

First, you need to configure your "Subnet" statements. These are used to determine what is local and what isn't. These can be as general as you want, but you'll want to make sure all of your local subnets are configured here. Use a separate line for each block. The syntax is "Subnet x.x.xx/y label". x.x.xx/y represents your network block in CIDR format. Example:

```
Subnet 172.16.0.0/16
```

Using all zeros (0.0.0.0/0) on a subnet will not produce useful graphs since CUFlow will think that EVERYTHING is local, so don't use it.

Second, you need to list any network groups that you want to get separate usage reports for. These are OPTIONAL settings. These groups only record the amount of traffic, not the detailed protocol and service break downs, but are useful non-the-less. A good example of how they can be useful can be [seen here](#). This graph shows how much bandwidth each of the dialup locations that backhaul thru the local router are using.

The syntax here is similar to the subnet syntax. "Network x.x.xx/y label". You can specify as many network blocks as needed, separated by commas. Examples:

```
Network 172.16.1.0/24 routers
Network 172.16.2.0/24,172.16.3.0/24 data_center
```

Next, you must change the OutputDir variable. I use /var/netflow/rrds/.

```
OutputDir /var/netflow/rrds
```

NOTE - Remember this directory. You will use it again in CUGrapher.pl.

To add the Top Talker reports, AKA Scoreboard, you'll want the scoreboard and aggregatescore lines to look something like this:

```
Scoreboard 25 /var/netflow/scoreboard /var/netflow/scoreboard/toptalkers.html
AggregateScore 25 /var/netflow/rrds/agg.dat /var/netflow/scoreboard/overall.html
```

25 is the number of toptalkers that will be listed. This can be changed to anything you want. Then, I add a symbolic link to /var/netflow/scoreboard/ in my apache document root. (ln -s /var/netflow/scoreboard toptalkers). Then you can access the reports by going to <http://yourserver/toptalkers/>. If you get a "forbidden" or "not found" error after adding this symbolic link, make sure that your Apache configuration allows it. (Options FollowSymLinks) You may also want to add a Directory section in your apache configuration file for /var/netflow/scoreboard. The default options will work just fine. This just gives you extra control.

You can add new services and protocols as you wish. For more information on these features, please consult the CUFlow documentation at <http://www.columbia.edu/acis/networks/advanced/CUFlow/CUFlow.html> or the file comes with the package.

Some extra notes about the config file.

1. Be sure to capitalize the variable names. If you do not, CUFlow will not recognize them. (i.e. Subnet, not subnet)

2. You cannot have any spaces in the labels. This will cause FlowScan to error out.
3. The CUFlow.cf file that is included with the package has many settings that are specific to the author's network. Be sure to comment out or delete these lines. The latest version of the config file has been changed to not include information specific to their network, but make sure you comment out what's there either way.
4. Make sure that you do not use the same directory as the RRDs, or a subdirectory of RRDs, for the scoreboard reports. It causes strange issues with CUGrapher.
5. If you remove or comment out any of the services, protocols, TOS or AS settings from the CUFlow config after you've started FlowScan, you will need to delete the corresponding RRD file from the /var/netflow/rrds folder. (I.e. If you delete eDonkey from the services section, delete service_edonkey_* in /var/netflow/rrds.)

Starting FlowScan

Now that the configuration file is complete, you should be able to start FlowScan. I use the following script to start it automatically at startup. (This file is included in the Support Files Tarball.)

```
#!/bin/sh
# description: Start Flowscan
# chkconfig: 2345 99 00

case "$1" in
'start')
/var/netflow/bin/flowscan >>/var/log/flowscan 2>&1 </dev/null &
>/dev/null
touch /var/lock/subsys/flowscan.1
;;
'stop')
killall -9 flowscan
rm -f /var/lock/subsys/flowscan.1
;;
*)
echo "Usage: $0 { start | stop }"
;;
esac
exit 0
```

This script will send all messages generated by FlowScan to /var/log/flowscan. If you use this script to start FlowScan, you can monitor whether or not it is working by tailing the log. (tail -f /var/log/flowscan)

If everything is working correctly, you should see logs like this:

```
sleep 30...
sleep 30...
2002/06/24 03:05:09 working on file flows.20020624_03:00:00...
2002/06/24 03:05:53 FlowScan-1.020 CUFlow: Cflow::find took 44 wallclock secs
(43.57 usr + 0.11 sys = 43.68 CPU) for 8729930 flow file bytes, flow hit ratio:
156224/158726
2002/06/24 03:05:56 FlowScan-1.020 CUFlow: report took 3 wallclock secs ( 0.00 usr
0.03 sys + 1.47 cusr 1.30 csys = 2.80 CPU)
sleep 30...
sleep 30...
sleep 30...
sleep 30...
sleep 30...
sleep 30...
sleep 30...
sleep 30...
sleep 30...
sleep 30...
2002/06/24 03:10:28 working on file flows.20020624_03:05:00...
2002/06/24 03:11:10 FlowScan-1.020 CUFlow: Cflow::find took 42 wallclock secs
(42.72 usr + 0.08 sys = 42.80 CPU) for 8568285 flow file bytes, flow hit ratio:
153537/155787
2002/06/24 03:11:13 FlowScan-1.020 CUFlow: report took 3 wallclock secs ( 0.00 usr
0.03 sys + 1.64 cusr 1.26 csys = 2.93 CPU)
sleep 30...
sleep 30...
sleep 30...
sleep 30...
```

If you see nothing or the file cannot be found, try starting FlowScan without the script. (From /var/netflow type in "bin/flowscan") Most error messages that FlowScan generates are self-explanatory.

There is one last thing to talk about since we have a log file here. I suggest you setup a logrotate process to rotate the /var/log/flowscan log file. Here is what you'll need to do to rotate it daily and keep 8 weeks of logs (current log plus the 7 past logs). Using your favorite editor, add a file to the directory /etc/logrotate.d called flowscan (vi /etc/logrotate.d/flowscan) and enter the following: (This file is also included in the Support Files Tarball.)

```
/var/log/flowscan {
rotate 7
weekly
}
```

Logrotate should already be setup to run each of the configuration files in this directory daily. For more information on Logrotate, type "man logrotate" at the command line.

CUGrapher.pl

Now, copy the CUGrapher.pl to your cgi-bin directory. You will need to change at least two things here before running it. First is the \$rddir variable. If you followed this document, that is /var/netflow/rrds. The other is \$organization. This variable is shown at the top of each generated graph.

If you installed Apache from the RPM, the cgi-bin directory is /var/www/cgi-bin/. If you installed Apache from the tarball, the default is /usr/local/apache/cgi-bin/. If your server is already up and running, you can see the grapher by opening <http://yourserver/cgi-bin/CUGrapher.pl>.

You can use the URLs created by CUGrapher.pl in IMG SRC tags if you want to build any summary HTML pages. You can also link to them in a web page. The next version of CUGrapher will have the option to not have a legend. This will make the graphs fit on summary pages more easily.

Example Sites

To see an example web site and reports generated by CUFlow and FlowScan, please visit Dave Plonka's sample graphs at: <http://wwwstats.net.wisc.edu>. Due to policy changes with my company, I can no longer allow access to our FlowScan sites. If anyone has a site they would like to submit, please let me know.

Appreciation

I have spent a lot of hours of my free time working on compiling, fine tuning and testing this document. I am happy to share it with the world. If it has helped save you time, please consider purchasing a DVD or other item from my [Amazon.com wish list](#) or make a contribution via PayPal. (E-mail me for PayPal information.) I will post a list of people that have shown their appreciation for this document.

Mailing Lists

Here are the mailing lists and instructions for subscribing to them.

FlowScan

There are two mailing lists having to do with FlowScan:

- **flowscan** -- a general mailing list for FlowScan users.
- **flowscan-announce** -- a low-volume, restricted post mailing list to keep FlowScan users informed of news regarding FlowScan.

The lists' respective archives are available at:

<http://net.doit.wisc.edu/~plonka/list/flowscan> and
<http://net.doit.wisc.edu/~plonka/list/flowscan-announce>

Announcements will be "cross-posted" to both lists, so there's no need to join both. These lists are hosted by the Division of Information Technology's Network Engineering Technology group at the University of Wisconsin - Madison. To subscribe to either of them, send email to: majordomo@net.doit.wisc.edu containing either: **subscribe flowscan** OR **subscribe flowscan-announce**.

You should receive an automatic response that will request that you verify your request to become a member of the list, to which you must reply with the authentication information there-in. Then, in response to your reply, you should receive a welcome message.

CUFlow

There is a mailing list having to do with CUFlow:

- **cufLOW-users** -- a general mailing list for CUFlow users.

The list's archives are available at: <https://www1.columbia.edu/sec/bboard/mj/cufLOW-users/>

This list is hosted by the Academic Information Systems department at Columbia University. To subscribe to the list, send email to: majordomo@columbia.edu containing: **subscribe cufLOW-users**.

You should receive an automatic response that will request that you verify your request to become a member of the list, to which you must reply with the authentication information there-in. Then, in response to your reply, you should receive a welcome message.

flow-tools

Please visit the following web site to subscribe and view the archives of the flow-tools maillist. <http://www.pairlist.net/mailman/listinfo/flow-tools>.

RRDTool

I've never had a reason to subscribe to RRDTool's mailing lists, but here is the page with their information. <http://people.ee.ethz.ch/~oetiker/webtools/rrdtool/maillinglists.html>

What else can FlowScan do? What if I need more?

There are many ways to take advantage of FlowScan. The flow-tools package includes many command line utilities to gather information from the flow files. There are also other FlowScan modules such as JKFlow.

flow-tools

Here are some of the other applications that I use in flow-tools. In every case that I've used these tools, you have to use flow-cat first, then pipe that into the next program, such as flow-nfilter, then flow-print or flow-stat. (i.e. *flow-cat ft-v05.2005-03-01.222500-0700 /flow-nfilter -F simple-ip-filter /flow-stat -f9 -s1*) This is specific to my setup at work, so I won't explain it all, but this will give me a report of all of the traffic going to one host on my network sorted by the second column data (flows). The best advice I can give you is play with these tools.

- flow-stat - gives you basic ASCII reports based on flow data. Read the man page here: <http://www.splintered.net/sw/flow-tools/docs/flow-stat.html>. Here is an example command line to see the current overall summary:
flow-cat ft-v05.2005-03-01.222500-0700 /flow-stat -f0
There are many more reports that you can get from flow-stat. Again, play around with it.
- flow-nfilter - allows you to filter the flow-data before sending it to a report. For example this could be useful for generating a report to support that a user has a virus. flow-nfilter has too many options to get into here, so read the man page. <http://www.splintered.net/sw/flow-tools/docs/flow-nfilter.html>
- flow-print - allows you to print out the flow data in a human readable (well kind of) format. It can be used with flow-nfilter to filter data based on anything that nfilter can specify. Note that the format options allow for a significant amount of detail. <http://www.splintered.net/sw/flow-tools/docs/flow-print.html>

There are other tools, but those are the ones I use most.

JKFlow

JKFlow is another FlowScan module that you would use in place of CUFlow. It is capable of far greater detail than CUFlow, but that comes at the cost of being very complex to configure. It is designed for enterprise class deployment and offers many advantages over CUFlow including an SMP version to take advantage of multiple processors. I am currently spending a lot of time banging my head against the wall trying to get JKFlow configured for our Primary NOC. I'll let you know how it goes in the next version. The link to the JKFlow site is in the appendix.

Appendix

Links

Here are is a list of links that you will find useful if you need more information:

1. <http://www.centos.org/> - This is the CentOS homepage. This is a replacement for RedHat.
2. <http://www.cisco.com/go/fn> - This is the Cisco Feature Navigator
3. <http://httpd.apache.org/> - This is the home page for Apache. This is the web server that I recommend.
4. <http://www.rrdtool.org/> - This is the RRDTool home page.
5. <http://www.splintered.net/sw/flow-tools/> - This is the flow-tools home page.
6. <http://net.doit.wisc.edu/~plonka/FlowScan/> - This is the FlowScan home page.
7. <http://www.columbia.edu/acis/networks/advanced/CUFlow/> - This is the CUFlow home page.
8. <http://net.doit.wisc.edu/~plonka/list/flowscan/> - This is the FlowScan mailing list home page.
9. <http://wwwstats.net.wisc.edu> - Examples of FlowScan and CUFlow.
10. <https://www1.columbia.edu/sec/bboard/mj/cufLOW-users/> - CUFlow mailing list archive. The mailing list is cufLOW-users@columbia.edu.
11. <http://users.telenet.be/jurgen.kobierczynski/jkflow/JKFlow.html> - Home of JKFlow, another reporting module for FlowScan. This is an enterprise class module and is extremely configurable. However, that extra detail comes at the price of being very complicated to configure.

Other Hardware Configs:

- Cisco 6500 Series *suggested* configuration:

```
mls netflow
mls aging normal 60
```

```
mls aging long 64
mls flow ip interface-full
mls nde sender version 5
mls nde interface
```

```
ip flow-export source xxx (replace with your interface)
ip flow-export version 5 peer-as
ip flow-export destination x.x.x.x y (replace "x.x.x.x y" with your collector
IP address and port as discussed previously)
```

"ip route-cache flow" should then be included in the interface configurations, including the VLAN interfaces as needed.

- Enterasys 8600 and ER-16 *suggested* configuration:

```
netflow set interval 5
netflow set ports all-ports
netflow set collector 192.168.1.1
netflow enable
```

Remember, these are not tested. Refer to the appropriate documentation for more information.

Change Log

Version 1.4a -- April 21, 2005

- Minor corrections submitted by Francois Caen. Thanks!

Version 1.4 -- March 1, 2005

- Updated HowTo for CentOS as a replacement for RedHat.
- Added suggested configurations for some hardware platforms.
- Added more information about the flow-tools applications that I use regularly.
- Added references to JKFlow.
- Included updates suggestions from Eric Laird.
- Removed references to my employer.
- Other minor changes to grammar, etc.

Version 1.3 -- October 4, 2003

- Changed document to use the native flow-tools file format. This change includes a new export script and should increase the performance of the application.

- Updated version numbers.
- Added specific instructions for adding to the RedHat path.
- Corrected error in the CUFlow configuration section. Thanks to Thomas Wiebe for pointing out my error.
- Added a killall to the flowscan startup script. Thanks to Velimir Kalik for the suggestion.
- Added support tarball to download list.

Version 1.2 -- January 16, 2003

- Minor spelling/grammar corrections.

Version 1.1 -- December 1, 2002

- Numerous updates made to the document to correct spelling/grammar and update to RedHat 8.0.
- Added instructions to add RRDTOol to the path on a RedHat 7.x/8.x system.
- Added instructions to get the ConfigReader::DirectiveStyle module installed correctly on RedHat 7.2+ with Perl 5.6.1+.
- Added pdksh to the list of needed applications. (Needed by CUFlow.)
- Added --prefix option to RRDTOol install to make updating RRDTOol less painful.

Version 1.0 -- July 24, 2002

This is a new document. No changes have been made yet.

Version 1.4a-Final

© copyright 2005 by Robert S. Galloway <rgalloway>

All Rights Reserved. The author believes that appropriate credit has been given. If anyone has been missed, please alert [me](#).

This document may be reproduced and distributed in its entirety (including this authorship, copyright, and permission notice), provided that no charge is made for the document itself.