Enabling NetFlow on Virtual Switches

ESX Server 3.5

NetFlow is a general networking tool with multiple uses, including network monitoring and profiling, billing, intrusion detection and prevention, networking forensics, and SOX compliance. NetFlow sends aggregated networking flow data to a third-party collector (an appliance or server). The collector and analyzer report on various information such as the current top flows consuming the most bandwidth in a particular virtual switch, which IP addresses are behaving irregularly, and the number of bytes a particular virtual machine has sent and received in the past 24 hours.

NetFlow is a mature technology, developed by Cisco, that is widely supported by third-party collectors. NetFlow enables visibility into virtual machine traffic in a virtualized server farm.

ESX Sever 3.5 NetFlow Experimental Support

NetFlow support in ESX Server 3.5 is experimental and supports only a limited set of the standard NetFlow features commonly found on physical switches today. Although the activation of NetFlow should not create stability issues, overall performance of the ESX Server host may be affected.

ESX Server 3.5 supports only a subset of the NetFlow Version 5 specification. Details about the limitations of the ESX Server implementation are described in "Limitations of NetFlow in ESX Server 3.5" on page 4.

How to Activate NetFlow Support in ESX Server 3.5

Enter the commands needed to activate NetFlow using the service console on the ESX Server host. You can enter the commands over an SSH connection or directly at the ESX Server host.

Take the following steps to activate NetFlow:

1 Prepare ESX Server for NetFlow configuration.

   Make sure the VMkernel TCP/IP stack is properly configured and that a VMkernel virtual interface (vmknic) exists on the network where your collector is located.

   The ESX Server implementation of NetFlow uses the ESX Server TCP/IP stack to send NetFlow packets on the network.

2 Load the NetFlow module.

   Enter the following command:

   ```
   vmkload_mod netflow
   ```

   You see the following response from the system:

   ```
   Using /usr/lib/vmware/vmkmod/netflow
   Module load of netflow succeeded.
   ```

3 Confirm that the NetFlow module is loaded.

Enter the following command:

```
vmkload_mod –l | grep netflow
```

You see a response from the system similar to the following:

```
netflow 0xc2e000 0x6000 0x2bf53a0 0x1000 16 Yes
```

4 Configure NetFlow

Use the application called `net–netflow` for the remaining configuration steps. The application is located in `/usr/lib/vmware/bin/` and takes multiple parameters, as shown in the following example:

```
net–netflow –e <vswitchname> <collectorhostname:port>
```

The minimum parameters required include:

- Names of the virtual switches where NetFlow will be activated
- Host IP address and port number of the NetFlow collector/analyzer

NOTE Multiple virtual switch names can be provided at one time, separated by commas.

NOTE If the port number is not provided, port 2055 is the default.

Configuration Examples

Use the `vmkload_app` wrapper to execute the `net–netflow` command. The `vmkload_app` wrapper is also in `/usr/lib/vmware/bin/`. The following examples show how to enter the commands for specific configurations.

Example 1

To enable NetFlow on vSwitch0 and send NetFlow data to 10.6.125.84 on port 4242, enter the following command on one line with no line break:

```
/usr/lib/vmware/bin/vmkload_app –i vmktcp /usr/lib/vmware/bin/net–netflow –e vSwitch0
10.6.125.84:4242
```

Example 2

To enable NetFlow on the virtual switches named vSwitch0, prod1, and test2 and send NetFlow data to 10.6.125.84 on port 4242, enter the following command on one line with no line break:

```
/usr/lib/vmware/bin/vmkload_app –i vmktcp /usr/lib/vmware/bin/net–netflow –e vSwitch0,prod1,test2
10.6.125.84:4242
```

NOTE `vmkload_mod` and `vmkload_app` are unsupported commands in ESX Server. Do not try to use them for other purposes.

Detailed Configuration Notes

If a virtual switch name does not exist, or registration with one of the virtual switches fails, `net–netflow` prints an error and exits immediately. Because NetFlow Version 5 uses UDP as the transfer protocol, no point-to-point connection is initiated with the collector and `net–netflow` cannot immediately detect if it is not able to contact the collector. In such a case, `net–netflow` continues to run but periodically and asynchronously prints messages about the loss of its UDP packets.

NOTE NetFlow analysis begins as soon as `net–netflow` is started. NetFlow quickly starts to send packets through the VMkernel stack.

To deactivate NetFlow, halt `net-netflow` by sending it a signal. To do so, press Ctrl-C if the instance is in the foreground of the current command prompt or enter a `kill` command for its corresponding PID. When you kill `net-netflow`, all the virtual switches registered for NetFlow are automatically unregistered and exporting of NetFlow export packets promptly stops.

NOTE This implies that in order to change the set of virtual switches on which NetFlow is enabled, you must first kill the current `net-netflow` process instance, then initiate a new instance specifying a new set of virtual switches. This completely resets the exporting process and does not cause issues with any of the collectors VMware tested.

You may find it impractical to monopolize the service console for `net-netflow`. To run `net-netflow` as a daemon, add the `-S` parameter to `vmkload_app` as shown in the following example, which you enter on one line with no line break:

```
/usr/lib/vmware/bin/vmkload_app -S -i vmktcp /usr/lib/vmware/bin/net-netflow -e vSwitch0
nf-collector1.mycompany.com:4242
```

This command launches `net-netflow` in the background, and you can use or terminate the command console session without affecting the `net-netflow` process.

To halt NetFlow when `net-netflow` is running as a daemon, you must enter a `kill` command for its corresponding PID.

Additional Configuration Options

The `net-netflow` application has several option parameters.

- `-e <vSwitch1[,vSwitch2[,...]]>`
  The list of virtual switches on which NetFlow should be activated, as shown in "Configuration Examples" on page 2.
- `-h`
  Prints usage— a quick way to retrieve usage information.
- `-v`
  Verbose. Causes `net-netflow` to output a brief summary of the flows sent on the network. This output may be useful for debugging purposes. This option may have slight performance impact if used on a busy network.
- `-s <v>`
  The `net-netflow` program periodically polls the VMkernel for flows to export. The default polling period is 1 second (1000ms). If you see a "Maximum purgatory size reached" message in the VMkernel log, you have reached the maximum flow queue limit. In this situation, some flows may be dropped. Lowering the polling period may solve this problem. For example, `-s 100` changes the polling period to 100ms.
- `-p`
  Use an alternate method of recording the port IDs in flow records.

  In ESX Server, it is possible to run multiple virtual switches on the same host. The NetFlow specification and current collectors do not deal well with the fact that one source IP address does not necessarily correlate to a single virtual switch.

  NetFlow on ESX Server embeds the virtual switch ID into the engineType and engineID fields of the header of each NetFlow export packet. (Flows from different virtual switches are always sent in separate packets.) Most collectors ignore these fields. In this case, interfaces on different virtual switches that have the same local ID are merged or aggregated into a single interface by the collector, mixing their flows together. If your collector shows this behavior, you can use this option to change the way source port IDs and destination port IDs are encoded.

When you use the `-p` option, the port IDs written in each flow record (the srcPort and dstPort fields of each flow record) are modified using the virtual switch's ID. The virtual switch's ID is stored in the seven most significant bytes (MSBs) of these fields, along with the original port ID located in the nine least significant bytes (LSBs). This is summarized in the following table (p = port ID; v = virtual switch ID):

Bit	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
ID	v	v	v	v	v	v	v	p	p	p	p	p	p	p	p	p

For example, if the source port of a flow is located on interface 13 of virtual switch 2, the srcPort is filled as shown in the following table:

Bit	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
ID	0	0	0	0	0	1	1	0	0	0	0	0	1	1	0	1

The interface appears as interface 1549 in your collector.

This option provides a way to generate unique port IDs per ESX Server host while still offering the possibility of retrieving information identifying the specific virtual machine referred to in each ID.

Limitations of NetFlow in ESX Server 3.5

NetFlow export is supported on ESX Sever 3.5 only as an experimental feature. Additionally, ESX Sever 3.5 exports flows in the NetFlow Version 5 format, with the following information missing (the corresponding fields have a value of zero):

- IP address of next hop router ("nexthop" in the specification)
- Autonomous system number of the source, either origin or peer ("src_as" in the specification)
- Autonomous system number of the destination, either origin or peer ("dst_as" in the specification)
- Source address prefix mask bits ("src_mask" in the specification)
- Destination address prefix mask bits ("dst_mask" in the specification)

The lack of any of these fields should not cause problems with any of the major collectors.

The idle timeout and active timeout are respectively statically set to 15 seconds and 5 minutes. Currently these values are not changeable.

There is no way to dynamically change the set of virtual switches where NetFlow is enabled. To make a change, you must kill the previous `net-netflow` process and launch a new one.

There is no sampling mode available on this version of ESX Server.

This implementation should not be used to do strict traffic accounting. Although the implementation is generally quite accurate, memory pressure conditions and network congestion may result in dropped flows, without any way to retrieve them.

In order to enable and configure NetFlow, you must use the service console directly. With ESX Sever 3.5, there is no support for configuring or managing NetFlow using VirtualCenter or the remote command line interface. This means NetFlow is not supported in ESX Server 3i environments.

Cautions on Using NetFlow in ESX Server 3.5

- Do not launch more than one instance of `net-netflow`.

  Do not launch two `net-netflow` instances at the same time. Although in theory this should be safe, the configuration has not been tested. It may lead to unexpected behavior.

- Do not delete a virtual switch that has NetFlow activated.

  Deleting a virtual switch with NetFlow activated may cause ESX Server to become unstable.

VMware, Inc. 3401 Hillview Ave., Palo Alto, CA 94304 www.vmware.com

Revision 20071108