NBAR support for HTTP

Last Updated: January 2011

Introduction

Network-Based Application Recognition (NBAR) is a Cisco IOS® Software feature that performs stateful, deep-packet inspection on a dataflow to identify the packet type and the protocol the flow belongs to. In conjunction with Cisco IOS Software quality of service (QOS), NBAR can apply application-level policies like guaranteeing/limiting bandwidth, policing, marking, and dropping to mention a few.

NBAR can distinguish more than 900 different protocols using protocol signatures inside the packet content. It can also inspect custom protocols by using a custom Protocol Description Language Module (PDLM) that has the protocol signatures.

This document discusses insights of NBAR support for HTTP.

Q. How does the classification work in practice?

A. Once enabled on any interface, every incoming packet will be subjected to deep-packet inspection on its IP header or on the payload content depending on the match policies defined under the respective class map.

The signature and attributes to be matched are included in the Packet Description Language Module (PDLM) uploaded to the router. NBAR can detect HTTP transactions on any port number and on any packet. It can also perform subport classification, which is a classification per URL, MIME type, or other parameters going beyond plain HTTP.

For example, NBAR can match based on User-Agent (such as “Mozilla/4.0”), Referrer, From, information about the software used by the server, Content-Encoding (such as “gzip”), and so on.

Q. What is the maximum number of concurrent URLs or hosts that can be matched at the same time?

A. This depends on the platform. The following are the performance numbers:

- Integrated Services Routers (ISR): 24
- ASR 1000 running RLS7: 20
- PISA-32 cards: 55

  (This is because they use a larger Ternary

  Content Addressable Memory (TCAM))

These numbers are per router, per protocol. Practically this is a big limitation.

Q. Does it work with IPv6?

A. Currently the IPv6 classification for HTTP is port-based, so only port 80 traffic will be classified. Subclassification is not supported.

Q. Does it support full payload inspection anywhere in the packet or only for the first X bytes?

A. This is platform dependant. The NBAR engine does not restrict inspecting anything. There might be some limitations if hardware acceleration is used.

For custom protocols, only the first 255 bytes of the payload can be inspected.

© 2011 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Q. How does NBAR handle fragmented traffic?

A. NBAR does the signature and regular expression matching on a per packet basis. Therefore, if fragmentation happens in the middle of the regular expression or signature, the matching may not happen correctly.

However, if the whole signature is within one packet (or one fragment), then the protocol would be discovered. If the first fragmented packet is classified as protocol X, then all of the remaining fragments will be classified as X by default.

Q. What happens to marking when NBAR finds a protocol in a dataflow?

A. If there is a match on any packet in a dataflow, all the remaining packets are classified based on the first match, and the configured policy is applied.

The remainder of the packets that belong to the same dataflow will be inspected, but the classification may not change. However additional attributes may be added if needed over the lifetime of the transaction.

For instance, if a dataflow has been wrongly classified as HTTP in the beginning and has later been found to be something else (like, say, RTP), the dataflow will still be marked as HTTP.

Q. With pipelining, multiple objects can be requested at the same time. What is to be done if there's a match?

A. NBAR does not support pipelining, and thus the classification based on URLs might not work at all for such connections. Hostname may still work, but that is still unsupported and therefore not guaranteed. Some headers will not work either.

Q. If the name of the server is replaced by its IP address, will NBAR be able to mark it?

A. No. If the server's hostname is replaced by its IP address, the matching will not work. Add hostnames and IP addresses to the NBAR matching filters, or if you need to match a specific host, use an IP-based access list instead.

Q. How does NBAR detect HTTPS traffic?

A. Since the traffic itself is encrypted, the recognition is based on the port number.

Q. Is it right to assume that nothing can be done for HTTPS?

A. Yes, NBAR will identify this traffic as HTTPS, and there's no other information that we can gather out of the encrypted payload.

Q. What happens when an HTTP proxy is used?

A. The proxied traffic will be classified as HTTP (assuming this is not an encryption proxy over SSL).

Q. What is the NBAR subclassification for HTTP?

A. NBAR can also analyze the HTTP flow deeper to extract more information called subclassification. Examples are user agent, host, Referer, Server, Location, Content-Encoding, Content-Type, URL, MIME Type", and so on.

For instance, this will match traffic going to a particular URL:

```
class-map match-all <classname>
    match protocol http c-header-field "http://www.cisco.com/routers"
```

Note: The number of subclassifications per protocol per router is limited to 20.

Q. Can URL and host contain regular expressions?

A. Yes, absolutely. Regular expressions can be used. See Table 1 for a list of regular expressions.

© 2011 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Table 1. Regular Expressions

Regular Expression	Description	Example
*	Matches zero or more sequences of the character preceding the asterisk. Also acts as a wildcard for matching any number of characters.	*.cisco.com matches www.cisco.com, and tools.cisco.com.
|	Acts as an “and” between multiple expressions.	*.com|*.org|*.net will match anything that contains .com, .org or .net.
^	The caret matches the beginning of the string.	^www will match www.cisco.com but not test.www.cisco.com

For a complete overview of regular expressions and how they are used in Cisco IOS Software, please check out the document at http://www.cisco.com/en/US/docs/ios/12_2/termserv/configuration/guide/tcfaapre_ps1835_TSD_Products_Configuration_Guide_Chapter.html.

Q. How can I match for a host or URL?

A. Every line in a class map is either true (if there is a match) or false (if there’s none). The class map itself can be of type “match-any” (if any of the criteria are true, then the packet matches the class) or “match-all” (all the criteria must be true for the packet to belong to this class).

A typical HTTP request is formatted like http://hostname/url . NBAR can match based on the hostname using the "host" keyword or based on the URL using the "url" keyword.

Example to match based on Host and URL

This example will match the page "cisco.com/go/nbar", as well as www.cisco.com/go/nbar and other variants.

```
class-map match-all <classname>
 match protocol http host "*cisco.com"
 match protocol http url "/go/nbar"
```

This will, for instance, match packets like this (in bold the matched content):

```
CF[4  ] TCP 10.10.10.2(54454) -> 72.163.4.161(80  ) ACK
        45 00 05 14 7d 24 40 00 3f 06 58 70 0a 0a 0a 02   E....$@.?.Xp....
        48 a3 04 a1 d4 b6 00 50 cd 93 ee f3 d5 2b 18 f5   H......P.....+..
        50 10 ff ff 08 81 00 00 47 45 54 20 2f 67 6f 2f   P.......GET /go/
        6e 62 61 72 20 48 54 54 50 2f 31 2e 31 0d 0a 48   nbar HTTP/1.1..H
        6f 73 74 3a 20 77 77 77 2e 63 69 73 63 6f 2e 63   ost: www.cisco.c
        6f 6d 0d 0a 55 73 65 72 2d 41 67 65 6e 74 3a 20   om..User-Agent:
        4d 6f 7a 69 6c 6c 61 2f 35 2e 30 20 28 4d 61 63   Mozilla/5.0 (Mac
        69 6e 74 6f 73 68 3b 20 49 6e 74 65 6c 20 4d 61   intosh; Intel Ma
        63 20 4f 53 20 58 20 31 30 2e 36 3b 20 72 76 3a   c OS X 10.6; rv:
        32 2e 30 62 36 29 20 47 65 63 6b 6f 2f 32 30 31   2.0b6) Gecko/201
        30 30 31 30 31 20 46 69 72 65 66 6f 78 2f 34 2e   00101 Firefox/4.
        30 62 36 0d 0a 41 63 63 65 70 74 3a 20 74 65 78   0b6..Accept: tex
        74 2f 68 74 6d 6c 2c 61 70 70 6c 69 63 61 74 69   t/html,applicati
```

Important: The classification needs to be configured with a separate host and URL as in the example below; you cannot match on http://cisco.com/go/nbar directly.

© 2011 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Example to match based on host or URL

This example is similar to the previous one except that it matches on either the host name or the URL. Note the "match-any" statement:

```
class-map match-any <classname>
 match protocol http host "www.cisco.com"
 match protocol http url "*.exe"
```

Example to block Facebook traffic

This example will block HTTP traffic to the social network site Facebook.

```
class-map match-all facebook
 match protocol http host "*facebook.com"
!
policy-map nofacebook
 class facebook
   drop
!
interface FastEthernet4
service-policy output nofacebook
```

This example will effectively block packets like this one (in bold is the matched content):

```
FF[3  ] TCP 10.10.10.2(54178) -> 69.63.189.34(80  ) ACK PSH FIN
        45 00 03 70 53 30 40 00 3f 06 ce ea 0a 0a 0a 02   E..pS0@.?.......
        45 3f bd 22 d3 a2 00 50 af 43 4c af e0 79 ac 84   E?."...P.CL..y..
        80 19 ff ff e6 c4 00 00 01 01 08 0a 18 48 46 0b   .............HF.
        e2 b0 ad 59 47 45 54 20 2f 20 48 54 54 50 2f 31   ...YGET / HTTP/1
        2e 31 0d 0a 48 6f 73 74 3a 20 77 77 77 2e 66 61   .1..Host: www.fa
        63 65 62 6f 6f 6b 2e 63 6f 6d 0d 0a 55 73 65 72   cebook.com..User
        2d 41 67 65 6e 74 3a 20 4d 6f 7a 69 6c 6c 61 2f   -Agent: Mozilla/
        35 2e 30 20 28 4d 61 63 69 6e 74 6f 73 68 3b 20   5.0 (Macintosh;
        49 6e 74 65 6c 20 4d 61 63 20 4f 53 20 58 20 31   Intel Mac OS X 1
```

Q. What's happening to HTTP when an HTTP proxy is used?

A. Typically when an HTTP proxy is used, the URL contains both the hostname and the URL, as opposed to only the URL in direct mode.

Packet capture without proxy configured, direct connection on port 80:

```
CF[4  ] TCP 10.10.10.2(54454) -> 72.163.4.161(80  ) ACK
        45 00 05 14 7d 24 40 00 3f 06 58 70 0a 0a 0a 02   E....$@.?.Xp....
        48 a3 04 a1 d4 b6 00 50 cd 93 ee f3 d5 2b 18 f5   H......P.....+..
        50 10 ff ff 08 81 00 00 47 45 54 20 2f 67 6f 2f   P.......GET /go/
        6e 62 61 72 20 48 54 54 50 2f 31 2e 31 0d 0a 48   nbar HTTP/1.1..H
```

© 2011 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

```
        6f 73 74 3a 20 77 77 77 2e 63 69 73 63 6f 2e 63   ost: www.cisco.c
        6f 6d 0d 0a 55 73 65 72 2d 41 67 65 6e 74 3a 20   om..User-Agent:
        4d 6f 7a 69 6c 6c 61 2f 35 2e 30 20 28 4d 61 63   Mozilla/5.0 (Mac
        69 6e 74 6f 73 68 3b 20 49 6e 74 65 6c 20 4d 61   intosh; Intel Ma
        63 20 4f 53 20 58 20 31 30 2e 36 3b 20 72 76 3a   c OS X 10.6; rv:
```

Packet capture with proxy configured; see that the GET statement is different as well as the destination IP (proxy) and port (Squid Server):

```
CF[4  ] TCP 10.10.10.2(54631) -> 192.168.2.3(3128) ACK
        45 00 05 14 2d c8 40 00 3f 06 32 65 0a 0a 0a 02   E...-.@.?.2e....
        c0 a8 02 03 d5 67 0c 38 30 3b 01 cf 5c a0 41 2f   .....g.80;..\.A/
        80 10 ff ff f8 af 00 00 01 01 08 0a 18 48 bd b6   .............H..
        16 f8 74 4b 47 45 54 20 68 74 74 70 3a 2f 2f 77   ..tKGET http://w
        77 77 2e 63 69 73 63 6f 2e 63 6f 6d 2f 67 6f 2f   ww.cisco.com/go/
        6e 62 61 72 20 48 54 54 50 2f 31 2e 31 0d 0a 48   nbar HTTP/1.1..H
        6f 73 74 3a 20 77 77 77 2e 63 69 73 63 6f 2e 63   ost: www.cisco.c
        6f 6d 0d 0a 55 73 65 72 2d 41 67 65 6e 74 3a 20   om..User-Agent:
        4d 6f 7a 69 6c 6c 61 2f 35 2e 30 20 28 4d 61 63   Mozilla/5.0 (Mac
        69 6e 74 6f 73 68 3b 20 49 6e 74 65 6c 20 4d 61   intosh; Intel Ma
        63 20 4f 53 20 58 20 31 30 2e 36 3b 20 72 76 3a   c OS X 10.6; rv:
```

Q. What's happening when a SOCKS proxy is used?

A. Generally speaking the traffic will be classified as SOCKS and not as the protocol transported by SOCKS.

Exceptions are made for some messenger protocols over SOCKS that are classified as messenger protocols and not SOCKS, for example, Yahoo, MSN, AOL, and Bittorent over SOCKS.

Q. Is the local traffic (from/to the router) also inspected, or only the traffic going through?

A. All traffic going through the router and also local traffic from and to the router can be inspected by NBAR. Local traffic destined to the router always hits NBAR on the ingress data plane. However local traffic originating from the router may or may not hit NBAR depending on the data-plane architecture for different platforms and thus will go directly for egress queuing.

Q. How should application PDLs be used, and would an HTTP match be required as well along with an application match?

A. No. The application PDL will do everything. For instance NBAR can be configured to match for YouTube video. This will let the user browse the youtube.com website but not watch the video even if—for instance—the video is not streamed from youtube.com but from googlevideos.com.

This example will match all browsing on the youtube.com website:

```
class-map match-all youtube
  match protocol http hostname *youtube.com
```

This example will not match browsing youtube.com, but will match any streaming video off YouTube:

```
class-map match-all youtube
  match protocol youtube
```

© 2011 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Therefore, the YouTube PDL is enough and does not need to be mixed with traditional HTTP inspection unless traffic to the YouTube website needs to be included as well. The same reasoning applies to other PDLs also.

Q. Are there any known bugs or possible issues?

A. No.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Cisco and the Cisco Logo are trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and other countries. A listing of Cisco's trademarks can be found at www.cisco.com/go/trademarks. Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1005R)

Printed in USA c67-646888-00 02/11

© 2011 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.