Date: 2014-10-15
Author: Magnus Larsson
Email magnus[at]stril.com
Phone: +46 (0)8 795 82 42

Address:
Stril Networks
Sorterargatan 12
162 50 Vallingby
Sweden

Performance study of RTP header compression on Cisco routers

RTP header compression is a method for reducing traffic bandwidth on links carrying VOIP traffic. However Cisco advises against enabling RTP header compression (cRTP) on link speeds above 1.5Mbps [1]*. This paper summarizes a lab test showing cRTP works fine on high bandwidth links and the maximum packet forwarding performance of the router is reduced to roughly half with cRTP enabled.*

RTP Header Compression Background

RTP is a packet format for delivering real time traffic such as voice and video over internet. It relies on UDP/IP for transport. The combined header size of RTP+UDP+IP packets is 40 bytes. Voice CODECs often have a payload size of around 20 to 60 bytes [2], so packet headers consumes much of the bandwidth in a voice over IP call. RFC2505 specifies a standard for compressing the 40 byte RTP/UDP/IP headers into two bytes (or 4 bytes if UDP checksum is used) [3]. Header compression is especially useful for expensive connections such as satellite links carrying RTP traffic.

This lossless compression scheme relies on the fact that much of the data in the RTP/UDP/IP headers is constant or at least predictable for each RTP session. The compressor sends the full header once for every session. The full header state for each session is stored in a table at both the compressor and decompressor. The compressor replaces the RTP/UDP/IP header with a 2 or 4 byte "compressed" header containing the necessary fields of information so the decompressor can reconstruct the original headers using the stored state information for that session. Compression/decompression is performed per IP hop, such as before and after a PPP link, if needed.

Without header compression

IP (20 bytes)	UPD (8 bytes)	RTP (12 bytes)	Payload (variable size)

With header compression

cRTP Header (2 or 4 bytes)	Payload (variable size)

RTP Header Compression on Cisco routers

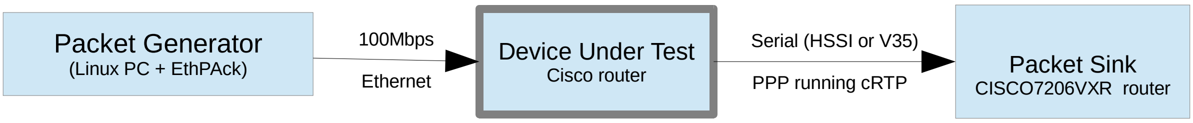
Cisco routers support Compressed RTP (cRTP) over serial lines running High-Level Data Link Control (HDLC), Point Protocol (PPP) or Frame Relay encapsulation. Cisco's configuration guide for cRTP contains the following note [1]:

"Using RTP header compression on any high-speed interfaces--that is, anything over T1 speed--is not recommended. Any bandwidth savings achieved with RTP header compression may be offset by an increase in CPU utilization on the router."

Aha! A cRTP speed limit recommendation. Let's do a test to see what happens if we break it!*

*Disclaimer: These experiments were performed in a controlled lab environment with trained personnel and risk control measures in place. Never attempt to replicate without adult supervision.

Test procedure



The main idea is to establish a PPP connection with cRTP between two Cisco routers and have a packet generator feed a stream of RTP packets over it.
After more than one minute of packet flow the following was recorded:

- Packet rate set in packETH software.
- CPU load of Device under Test (DUT). The routers last 30 second CPU load average were taken after one minute of packet flow. The
- Rate of packets received at the Packet Sink Router (last 30 second average, given by the "show interface" command)

The test is performed twice, one with cRTP enabled and one without, to establish a baseline performance.

If the CPU load of the DUT rises above 90% it is considered to be an unhealthy saturated state with too much traffic and the previous packet rate value is taken as a "healthy" maximum performance. This upper load level of 90% was chosen somewhat arbitrary for this test. It can be seen in the results that the router can still forward additional packets per second beyond the 90% load level, but since the IP Cisco Express Forwarding (CEF) method takes place during the prioritized interrupt handling of each packet arriving at the interface, the CPU load approaching 100% will cause other processes/threads the routers to starve. One of these processes will be the EXEC process which handles console/terminal I/O making it impossible to collect the statistics needed. As a result 90% seemed like a reasonable limit.

Cisco IOS CPU load and interface counter averages
Cisco doesn't use an "normal" arithmetic average but an exponential one to represent the CPU load and packet throughput for an interface. [4]. Every 5 seconds a new sample is taken and weighted with the old value by a "decay time" constant calculated from the load-interval setting. To get a fair "snapshot" reading of these numbers two measures were taken:

1) The load-interval command was set to the minimum, 30 seconds to maximize the decay fade off effect of old values.
2) The packet generator ran for 30 seconds to create a steady state before starting the 30 second test to let old values "fade off" before making the reading.

Packet generator
Hardware: IBM XSeries 345, Dual XEON 2.0GHz, 2.5GB RAM
Interface: 10/100 ethernet interface
Linux OS: Fedora 7
Packet generator software: packETH 1.4.1 packeth.sourceforge.net

Note: PackEth writes the raw packets directly to the Ethernet interface bypassing the routing table and ARP protocol of the OS. Thus the MAC address of the DUT's Ethernet port facing the Packet generator must be specified for the test:

Device Under Test (DUT)
The DUT accepts RTP packets from the Packet generator arriving at its Fast Ethernet interface.
It then forwards the packet over a serial line to the sink router. The serial line hardware is V.35 (2Mbps) and HSSI (52Mbps), depending on platform tested. PPP is used as serial protocol. Cisco forwarding method was "IP CEF". An example configuration for the CISCO7120-4T with HSSI is at the end of the document.

The following Cisco router models were tested:

- CISCO2620 with WIC-1T V35 interface.
- CISCO3640 with NM-1HSSI interface.
- CISCO7120-4T PA-H HSSI interface.

Packet Sink

The router acting as a packet sink receives the RTP packets from the DUT and discards them by routing them to a null interface. The router used in this test was a CISCO7206VXR with a NPE300. This Network Processing Engine is faster than all of the DUT's. Hence the packet sink did not introduce any errors in the measurement.

Test Results

Each test's maximum "healthy" packet forwarding rate has been marked in bold and green.

DUT: Cisco2620 WIC-1T, cRTP turned off.

Packet generator setting [pps out]	DUT CPU load (last one minute)	Sink received pps (last 30 sec average)	Sink bandwidth (measured over 30 sec) [kbps]
1000	13	999	351
2000	27	1999	703
3000 (3003)	41	2999	1407
4000	56	3999	1759
5000	***67***	***4999***	***1826***
6000 (6024)	Unresponsive		

DUT: Cisco2620 WIC-1T, with cRTP compression

Packet generator setting [pps out]	DUT CPU load (last one minute)	Sink received pps (last 30 sec average)	Sink bandwidth (measured over 30 sec) [kbps]
1000	25	1000	184
2000	48	1999	369
3000	***73***	***2999***	***554***
4000	99	3999	738
5000	Unresponsive	4999	923

DUT: 3640 HSSI uncompressed

Packet generator setting [pps out]	DUT CPU load (last one minute)	Sink received pps (last 30 sec average)	Sink bandwidth (measured over 30 sec) [kbps]
5000	37	4999	1759
10000	***67***	***9999***	***7038***
20000	92	19996	10664
30000	Unresponsive	30297	11715

DUT: 3640 HSSI compressed

Packet generator setting [pps out]	DUT CPU load (last one minute)	Sink received pps (last 30 sec average)	Sink bandwidth (measured over 30 sec) [kbps]
5000	**81**	**4999**	**923**
10000	Unresponsive	9936	1842

DUT: 7120-4T HSSI uncompressed

Packet generator setting [pps out]	DUT CPU load (last one minute)	Sink received pps (last 30 sec average)	Sink bandwidth (measured over 30 sec) [kbps]
10000	19	10000	3519
25000	36	24993	8798
50000	***70***	***49960***	***17588***
75000	94	76667	26992
100000	99	99976	35193

DUT: 7120-4T HSSI compressed

Packet generator setting [pps out]	DUT CPU load (last one minute)	Sink received pps (last 30 sec average)	Sink bandwidth (measured over 30 sec) [kbps]
10000	35	9998	1846
25000	74	24967	4612
30000	***88***	***30292***	***5594***
40000	99	39972	7383
45000	Unresponsive	45040	8306

Test summary

Device	*Maximum PPS according to Cisco* [4]	*Max PPS at less than 90% CPU measured in lab (cRTP turned off)*	*Max PPS at less than 90% CPU with cRTP turned on*
CISCO2620 WIC-1T	*25000*	*5000*	*3000*
CISCO3640 HSSI	*50000*	*10000*	*5000*
CISCO7120-4T HSSI	*175000*	*50000*	*30000*

Conclusion

Enabling cRTP on a Cisco router on links faster than 1.5Mbps works fine, but requires a router of about twice the packet forwarding performance compared to handling non cRTP traffic.

References

[1] "Header Compression," Cisco Systems Inc, 5 12 2011. [Online]. Available: http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/qos_hdrcomp/configuration/12-4/qos-hdrcomp-12-4-book/header-compression.pdf.

[2] Cisco Systems Inc, "Voice Over IP - Per Call Bandwidth Consumption," 02 02 2006. [Online]. Available: http://www.cisco.com/c/en/us/support/docs/voice/voice-quality/7934-bwidth-consume.html.

[3] Cisco Systems Inc, "Compressing IP/UDP/RTP Headers for Low-Speed Serial Links," 02 1999. [Online]. Available: http://www.faqs.org/rfcs/rfc2508.html.

[4] Cisco Systems, "Understanding the Definition of bits per second (bits/sec) from the show interfaces Command Output," 01 08 2006. [Online]. Available: www.cisco.com/c/en/us/support/docs/ios-nx-os-software/ios-software-releases-110/12816-3.html.

[5] Cisco Systemns Inc, "Portable Product Sheets – Routing Performance," 03 11 2009. [Online]. Available: http://www.cisco.com/web/partners/downloads/765/tools/quickreference/routerperformance.pdf.

Example of DUT config for CISCO7120-4T with HSSI

```
ip cef

interface FastEthernet0/0
 description packgen receiver
 ip address 1.1.1.2 255.255.255.0
 load-interval 30
 duplex auto
 speed 100
 no keepalive
 no cdp enable
!

interface Hssi3/0
 ip address 2.2.2.1 255.255.255.0
 encapsulation ppp
 ip tcp header-compression iphc-format
 hssi internal-clock
 serial restart-delay 0
 ip rtp header-compression iphc-format

ip route 0.0.0.0 0.0.0.0 Hssi3/0
```