Traceroute em Detalhes

Autores:
Gustavo Ramos
Artur Araujo

Grupo de Trabalho de Engenharia e Operação de Redes - 36ª Reunião
5 de Dezembro 2013

PINPOINT

IPivOt Telecom

Agenda

- Introdução e objetivos
- Fundamentos
- Programas
- Caminhos Assimétricos
- Um caso real …
- Resumo: Boas Práticas

Introdução e Objetivos

- **Objetivo**: "Desmistificar" a utilização de traceroute na análise de problemas complexos de roteamento Internet.

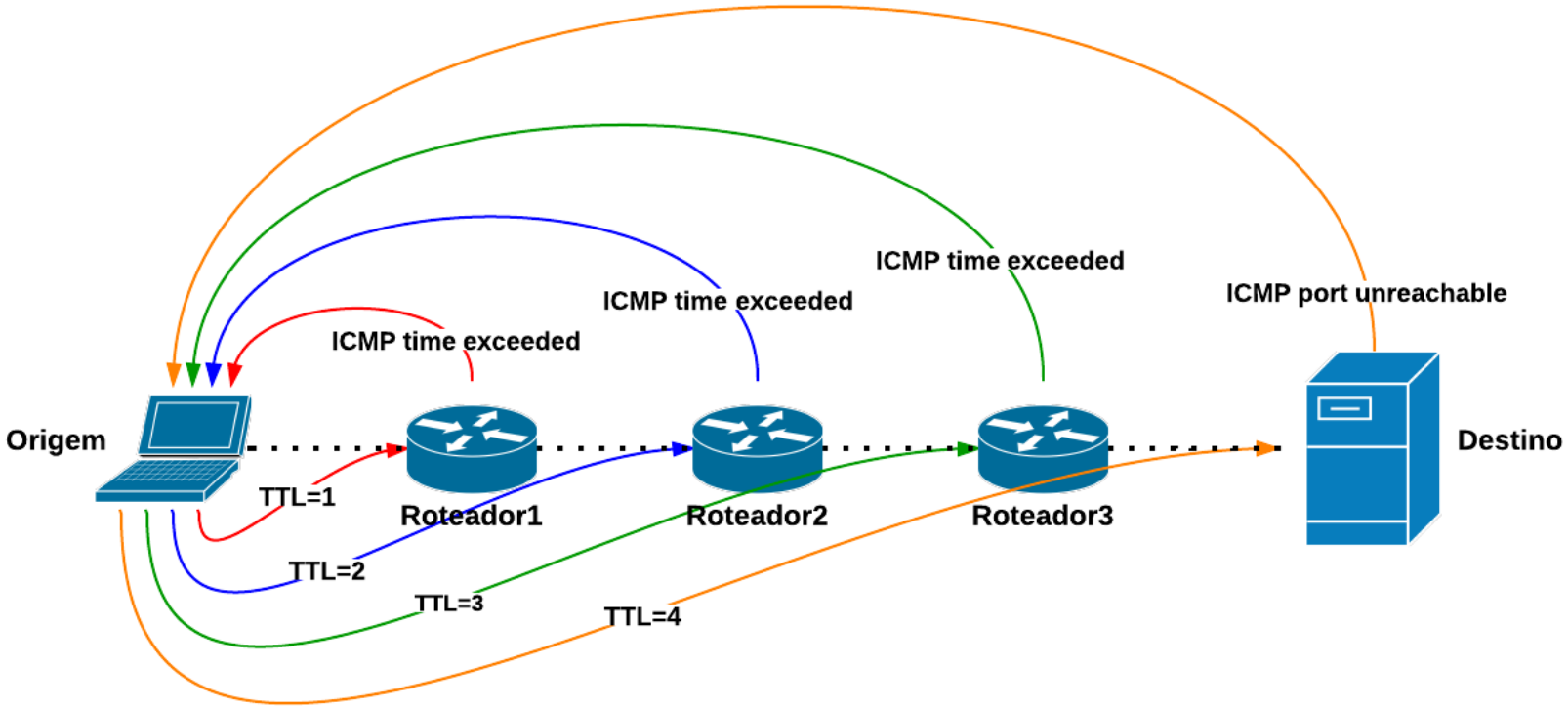
- O que é o traceroute?
  - Programa de computador desenvolvido por Van Jacobson (e outros) para permitir a visualização do caminho percorrido de um ponto a outro em uma rede IP.

Fundamentos

Como mapear o caminho?

- Origem envia pacotes com TTL crescente; ou seja, o primeiro pacote será enviado com TTL=1, o segundo com TTL=2, etc.
- Cada roteador intermediário que receber um pacote com TTL=1, irá decrementar o TTL, descartar o pacote e responder para a origem um pacote "ICMP Time Exceeded in-transit".
- Normalmente, como o traceroute envia pacotes UDP, o destino final irá responder com um pacote "ICMP port unreachable".

Fundamentos



Traceroute utilizando ICMP

```
host:~ $ traceroute -I -q1 201.6.42.6
traceroute to 201.6.42.6 (201.6.42.6), 64 hops max, 72 byte packets
 1  b3d18c01.virtua.com.br (179.209.140.1)  38.188 ms
 2  spojabrtd01.virtua.com.br (201.6.0.95)  14.738 ms
 3  spotvtrtd10.virtua.com.br (201.6.0.10)  12.682 ms
 4  c9062a06.virtua.com.br (201.6.42.6)  12.918 ms
```

```
17:15:03.113394 IP (tos 0x0, ttl 1, id 38556, offset 0, flags [none], proto ICMP (1), length 72)
   192.168.0.15 > 201.6.42.6: ICMP echo request, id 38555, seq 1, length 52
17:15:03.151121 IP (tos 0x0, ttl 255, id 18393, offset 0, flags [none], proto ICMP (1), length 56)
   179.209.140.1 > 192.168.0.15: ICMP time exceeded in-transit, length 36
         IP (tos 0x0, id 38556, offset 0, flags [none], proto ICMP (1), length 72)
   192.168.0.15 > 201.6.42.6: ICMP echo request, id 38555, seq 1, length 52

17:15:03.151994 IP (tos 0x0, ttl 2, id 38557, offset 0, flags [none], proto ICMP (1), length 72)
   192.168.0.15 > 201.6.42.6: ICMP echo request, id 38555, seq 2, length 52
17:15:03.166531 IP (tos 0xc0, ttl 254, id 792, offset 0, flags [none], proto ICMP (1), length 56)
   201.6.0.95 > 192.168.0.15: ICMP time exceeded in-transit, length 36
         IP (tos 0x0, ttl 1, id 38557, offset 0, flags [none], proto ICMP (1), length 72)
   192.168.0.15 > 201.6.42.6: ICMP echo request, id 38555, seq 2, length 52

17:15:03.167126 IP (tos 0x0, ttl 3, id 38558, offset 0, flags [none], proto ICMP (1), length 72)
   192.168.0.15 > 201.6.42.6: ICMP echo request, id 38555, seq 3, length 52
17:15:03.179629 IP (tos 0xc0, ttl 253, id 21818, offset 0, flags [none], proto ICMP (1), length 56)
   201.6.0.10 > 192.168.0.15: ICMP time exceeded in-transit, length 36
         IP (tos 0x0, ttl 1, id 38558, offset 0, flags [none], proto ICMP (1), length 72)
   192.168.0.15 > 201.6.42.6: ICMP echo request, id 38555, seq 3, length 52

17:15:03.180353 IP (tos 0x0, ttl 4, id 38559, offset 0, flags [none], proto ICMP (1), length 72)
   192.168.0.15 > 201.6.42.6: ICMP echo request, id 38555, seq 4, length 52
17:15:03.193101 IP (tos 0x0, ttl 252, id 38559, offset 0, flags [none], proto ICMP (1), length 72)
   201.6.42.6 > 192.168.0.15: ICMP echo reply, id 38555, seq 4, length 52
```

Traceroute utilizando UDP

```
# traceroute -f1 -m4 -q1 69.31.135.129
traceroute to 69.31.135.129 (69.31.135.129), 4 hops max, 60 byte packets
 1  router1-atl.linode.com (64.22.106.73)  0.495 ms
 2  64.22.106.9 (64.22.106.9)  0.395 ms
 3  xe-2-0-5-101.ar1.atl1.us.nlayer.net (69.31.135.41)  1.686 ms
 4  ae0-50g.cr1.atl1.us.nlayer.net (69.31.135.129)  4.853 ms
```

Porta UDP de destino entre 33434 e 33534.

```
18:27:34.027116 IP (tos 0x0, ttl 1, id 6283, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.60530 > 69.31.135.129.33434: UDP, length 32
18:27:34.027169 IP (tos 0x0, ttl 2, id 6284, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.58536 > 69.31.135.129.33435: UDP, length 32
18:27:34.027204 IP (tos 0x0, ttl 3, id 6285, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.37972 > 69.31.135.129.33436: UDP, length 32
18:27:34.027243 IP (tos 0x0, ttl 4, id 6286, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.47830 > 69.31.135.129.33437: UDP, length 32

18:27:34.027550 IP (tos 0xc0, ttl 254, id 32739, offset 0, flags [none], proto ICMP (1), length 56)
   64.22.106.9 > 50.116.36.165: ICMP time exceeded in-transit, length 36
        IP (tos 0x0, ttl 1, id 6284, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.58536 > 69.31.135.129.33435: UDP, length 32
18:27:34.027572 IP (tos 0x0, ttl 255, id 6283, offset 0, flags [none], proto ICMP (1), length 56)
   64.22.106.73 > 50.116.36.165: ICMP time exceeded in-transit, length 36
        IP (tos 0x0, ttl 1, id 6283, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.60530 > 69.31.135.129.33434: UDP, length 32
18:27:34.028876 IP (tos 0x0, ttl 253, id 51768, offset 0, flags [DF], proto ICMP (1), length 56)
   69.31.135.41 > 50.116.36.165: ICMP time exceeded in-transit, length 36
        IP (tos 0x0, ttl 1, id 6285, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.37972 > 69.31.135.129.33436: UDP, length 32
18:27:34.032085 IP (tos 0x0, ttl 252, id 60803, offset 0, flags [DF], proto ICMP (1), length 56)
   69.31.135.129 > 50.116.36.165: ICMP 69.31.135.129 udp port 33437 unreachable, length 36
        IP (tos 0x0, ttl 1, id 6286, offset 0, flags [none], proto UDP (17), length 60)
   50.116.36.165.47830 > 69.31.135.129.33437: UDP, length 32
```

Fundamentos: “Lendo o resultado”

```
# traceroute -q1 allspice.lcs.mit.edu
traceroute to allspice.lcs.mit.edu (18.26.0.122), 30 hops max, 40 byte packets
 1  *
 2  *
 3  187-92-134-181.customer.tdatabrasil.net.br (187.92.134.181)  8.584 ms
 4  187-100-7-5.dsl.telesp.net.br (187.100.7.5)  8.859 ms
 5  187-100-24-93.dsl.telesp.net.br (187.100.24.93)  11.271 ms
 6  187-100-44-114.dsl.telesp.net.br (187.100.44.114)  8.752 ms
 7  200-100-98-181.dial-up.telesp.net.br (200.100.98.181)  9.863 ms
 8  et-6-0-0-101-GRTRIOTW2.red.telefonica-wholesale.net (84.16.9.65)  47.408 ms
 9  176.52.255.70 (176.52.255.70)  124.742 ms
10  176.52.251.53 (176.52.251.53)  151.014 ms
11  te0-4-0-11.ccr21.dfw03.atlas.cogentco.com (154.54.11.97)  182.149 ms
12  be2032.ccr22.dfw01.atlas.cogentco.com (154.54.6.53)  185.834 ms
13  be2141.mpd22.mci01.atlas.cogentco.com (154.54.5.157)  194.112 ms
14  be2157.ccr22.ord01.atlas.cogentco.com (154.54.6.118)  169.293 ms
15  be2140.ccr22.bos01.atlas.cogentco.com (154.54.43.186)  178.004 ms
16  *
17  *
18  backbone-rtr-1-dmz-rtr-1.mit.edu (18.168.5.1)  192.324 ms
19  *
20  mitnet.trantor.csail.mit.edu (18.4.7.65)  191.166 ms
21  trantor.helicon.csail.mit.edu (128.30.0.246)  195.237 ms
22  mercury.lcs.mit.edu (18.26.0.122)  200.232 ms !X
```

Códigos comuns que podem aparecer no resultado do traceroute:

!H **Received a reply telling that the destination host is unreachable.**

!N **Received a reply telling that the destination network is unreachable.**

!P **Received a reply telling that the desired protocol is unavailable.**

!S **Received a reply telling that source routing failed.**

!X **Communication administratively prohibited.**

PINPOINT

IPivot Telecom

Fundamentos: “Lendo o resultado”

```
# traceroute -I 200.221.2.45
traceroute to 200.221.2.45 (200.221.2.45), 30 hops max, 60 byte packets
 1  router1-atl.linode.com (64.22.106.73)  0.493 ms  0.625 ms  0.754 ms
 2  64.22.106.9 (64.22.106.9)  0.464 ms  0.535 ms  0.603 ms
 3  atl-core-g-g1-6.gnax.net (63.247.69.178)  1.308 ms  1.385 ms  1.441 ms
 4  xe-2-0-5-103.ar1.atl1.us.nlayer.net (69.31.135.53)  1.752 ms  1.806 ms *
 5  ae-8.r04.atlnga05.us.bb.gin.ntt.net (204.2.241.93)  1.492 ms  1.505 ms  1.502 ms
 6  ae-8.r20.asbnva02.us.bb.gin.ntt.net (129.250.5.214)  20.514 ms  15.043 ms  17.839 ms
 7  ae-1.r04.asbnva02.us.bb.gin.ntt.net (129.250.3.17)  14.576 ms  14.972 ms  16.569 ms
 8  xe-3.telefonica-data.asbnva02.us.bb.gin.ntt.net (129.250.9.102)  14.265 ms  14.165 ms
    xe-1.telefonica-data.asbnva02.us.bb.gin.ntt.net (129.250.8.250)  25.529 ms
 9  176.52.250.17 (176.52.250.17)  40.100 ms  39.541 ms
    Te0-2-0-6-grtmiabr6.red.telefonica-wholesale.net (84.16.12.46)  45.289 ms
10  Xe0-1-11-0-grtsanem3.red.telefonica-wholesale.net (84.16.15.45)  154.025 ms  156.370 ms  154.055 ms
11  TEBRASIL-Et1-0-0-101-grtsanem3.red.telefonica-wholesale.net (84.16.10.154)  163.036 ms  163.023 ms  163.015 ms
12  187-100-53-10.dsl.telesp.net.br (187.100.53.10)  177.482 ms
    187-100-53-86.dsl.telesp.net.br (187.100.53.86)  154.543 ms
    187-100-53-70.dsl.telesp.net.br (187.100.53.70)  165.372 ms
13  187-100-53-178.dsl.telesp.net.br (187.100.53.178)  154.724 ms
    187-100-53-186.dsl.telesp.net.br (187.100.53.186)  158.706 ms  163.846 ms
14  187-11-166-2.dsl.telesp.net.br (187.11.166.2)  135.937 ms  135.323 ms  138.180 ms
15  200.221.136.158 (200.221.136.158)  138.911 ms  138.406 ms  136.085 ms
16  home.uol.com.br (200.221.2.45)  132.010 ms  132.618 ms  132.130 ms
```

PINPOINT

IPivot Telecom

Fundamentos: IPv4 vs. IPv6

 # traceroute -q1 -I www.facebook.com

```
traceroute to www.facebook.com (31.13.65.17), 30 hops max, 60 byte packets
 1  router1-atl.linode.com (64.22.106.73)  0.558 ms
 2  64.22.106.9 (64.22.106.9)  0.354 ms
 3  atl-core-g-g1-6.gnax.net (63.247.69.178)  0.399 ms
 4  xe-8-2-3.edge4.Atlanta2.Level3.net (4.35.6.113)  0.623 ms
 5  vlan52.ebr2.Atlanta2.Level3.net (4.69.150.126)  0.684 ms
 6  4.69.159.57 (4.69.159.57)  0.700 ms
 7  FACEBOOK-IN.edge5.Atlanta2.Level3.net (4.28.26.46)  1.071 ms
 8  ae1.bb01.atl1.tfbnw.net (74.119.78.214)  1.195 ms
 9  ae2.pr02.atl1.tfbnw.net (74.119.78.217)  1.154 ms
10  po126.msw01.02.atl1.tfbnw.net (31.13.27.123)  1.204 ms
11  edge-star-shv-02-atl1.facebook.com (31.13.65.17)  1.078 ms
```

 # traceroute6 -q1 -I www.facebook.com

```
traceroute to www.facebook.com (2a03:2880:f011:101:face:b00c:0:1), 30 hops max, 80 byte packets
 1  2600:3c02::8678:acff:fe5a:1941 (2600:3c02::8678:acff:fe5a:1941)  0.641 ms
 2  2604:b900:1::1 (2604:b900:1::1)  0.498 ms
 3  2604:b900:0:2::1 (2604:b900:0:2::1)  0.506 ms
 4  xe-8-2-3.edge4.Atlanta2.Level3.net (2001:1900:2100::24b1)  0.500 ms
 5  vl-51.edge5.Atlanta2.Level3.net (2001:1900:1c:1::f)  0.498 ms
 6  FACEBOOK-IN.edge5.Atlanta2.Level3.net (2001:1900:2100::fae)  0.563 ms
 7  ae1.bb01.atl1.tfbnw.net (2620:0:1cff:dead:beef::40a)  0.840 ms
 8  ae2.pr02.atl1.tfbnw.net (2620:0:1cff:dead:beef::407)  0.837 ms
 9  po126.msw01.02.atl1.tfbnw.net (2620:0:1cff:dead:bef0::85)  1.053 ms
10  edge-star6-shv-02-atl1.facebook.com (2a03:2880:f011:101:face:b00c:0:1)  0.577 ms
```

Notas:

- Opção “-I” (letra “i” maiúscula) para forçar o **envio** de pacotes ICMP.
- Opção “-q1” para enviar somente um pacote para cada TTL.

PINPOINT

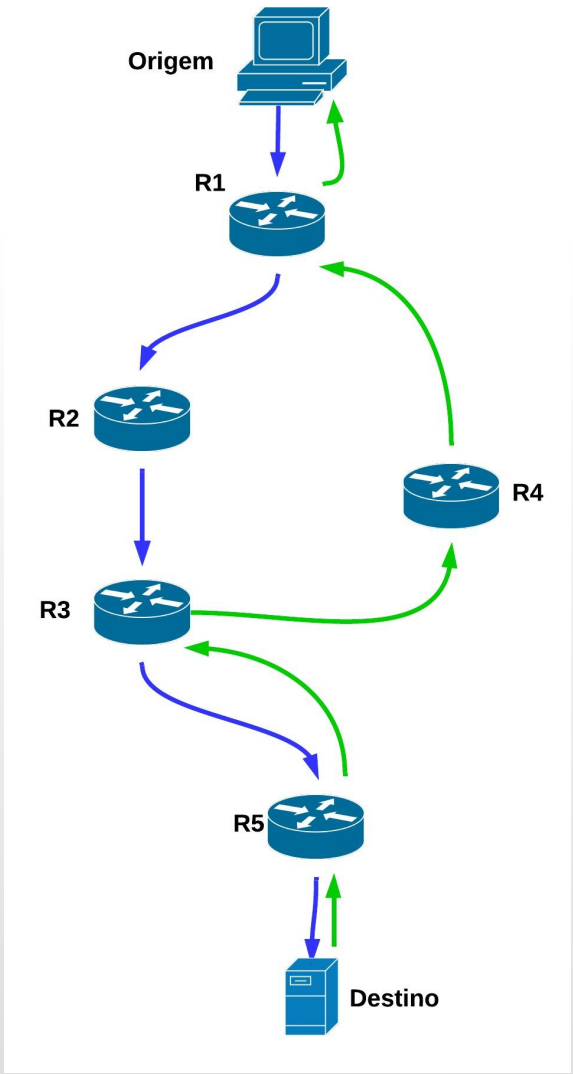
IPivot Telecom

Programas

- traceroute e traceroute6 (Linux, BSD, etc)
- tracert (Windows)
- MTR (Linux, etc) ou WinMTR (Windows)
- PingPlotter (Windows)
- Path Analyzer Pro (Mac OS X)
- Etc.

- Importante: lembre-se que programas diferentes implementam o traceroute de forma diferente. Por exemplo, por padrão o “tracert” utiliza ICMP e o “traceroute” utiliza UDP.
- Protocolo padrão utilizado no traceroute:
  - Windows: ICMP
  - Linux: UDP
  - Cisco: UDP
  - Juniper: UDP
  - Mikrotik: ICMP

Caminhos Assimétricos

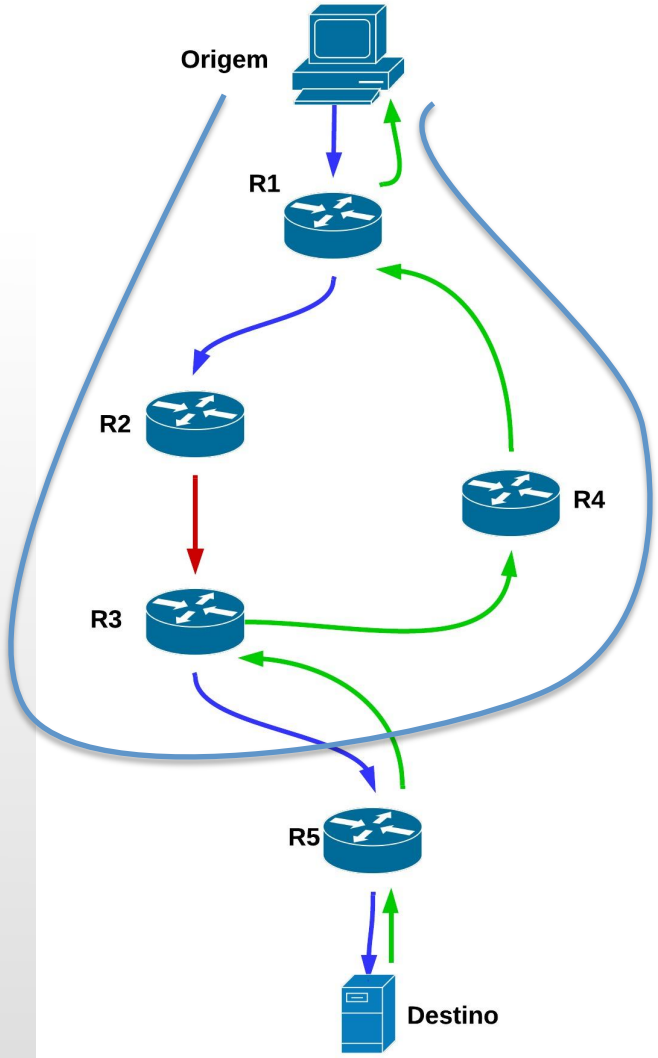


1 R1 (10.10.1.1) 1.178 ms 1.420 ms 0.898 ms
2 R2 (10.10.2.1) 3.511 ms 3.710 ms 2.871 ms
3 **R3 (10.10.3.1) 200.005 ms 211.471 ms 214.643 ms**
4 R5 (10.10.4.1) 210.785 ms 220.816 ms 211.575 ms
5 DST (10.10.5.1) 220.185 ms 230.816 ms 221.005 ms

PINPOINT

IPivot Telecom

Caminhos Assimétricos

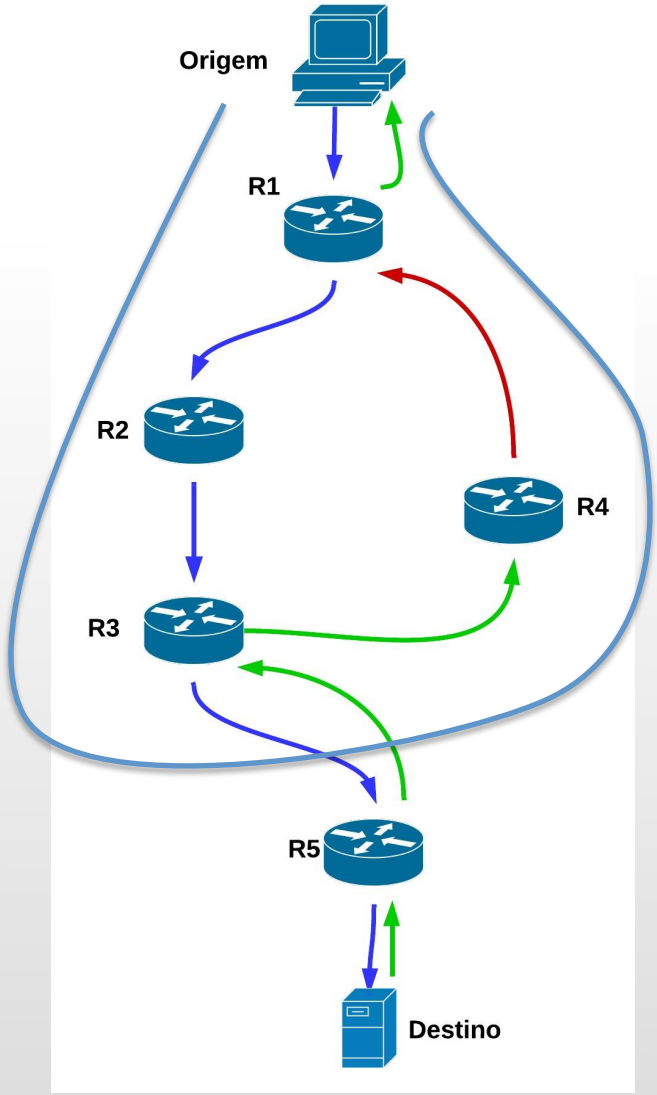


1 R1 (10.10.1.1) 1.178 ms 1.420 ms 0.898 ms
2 R2 (10.10.2.1) 3.511 ms 3.710 ms 2.871 ms
3 **R3 (10.10.3.1) 200.005 ms 211.471 ms 214.643 ms**
4 R5 (10.10.4.1) 210.785 ms 220.816 ms 211.575 ms
5 DST (10.10.5.1) 220.185 ms 230.816 ms 221.005 ms

PINPOINT

IPivot Telecom

Caminhos Assimétricos



1 R1 (10.10.1.1) 1.178 ms 1.420 ms 0.898 ms
2 R2 (10.10.2.1) 3.511 ms 3.710 ms 2.871 ms
3 **R3 (10.10.3.1) 200.005 ms 211.471 ms 214.643 ms**
4 R5 (10.10.4.1) 210.785 ms 220.816 ms 211.575 ms
5 DST (10.10.5.1) 220.185 ms 230.816 ms 221.005 ms

Traceroute nas duas direções!

PINPOINT

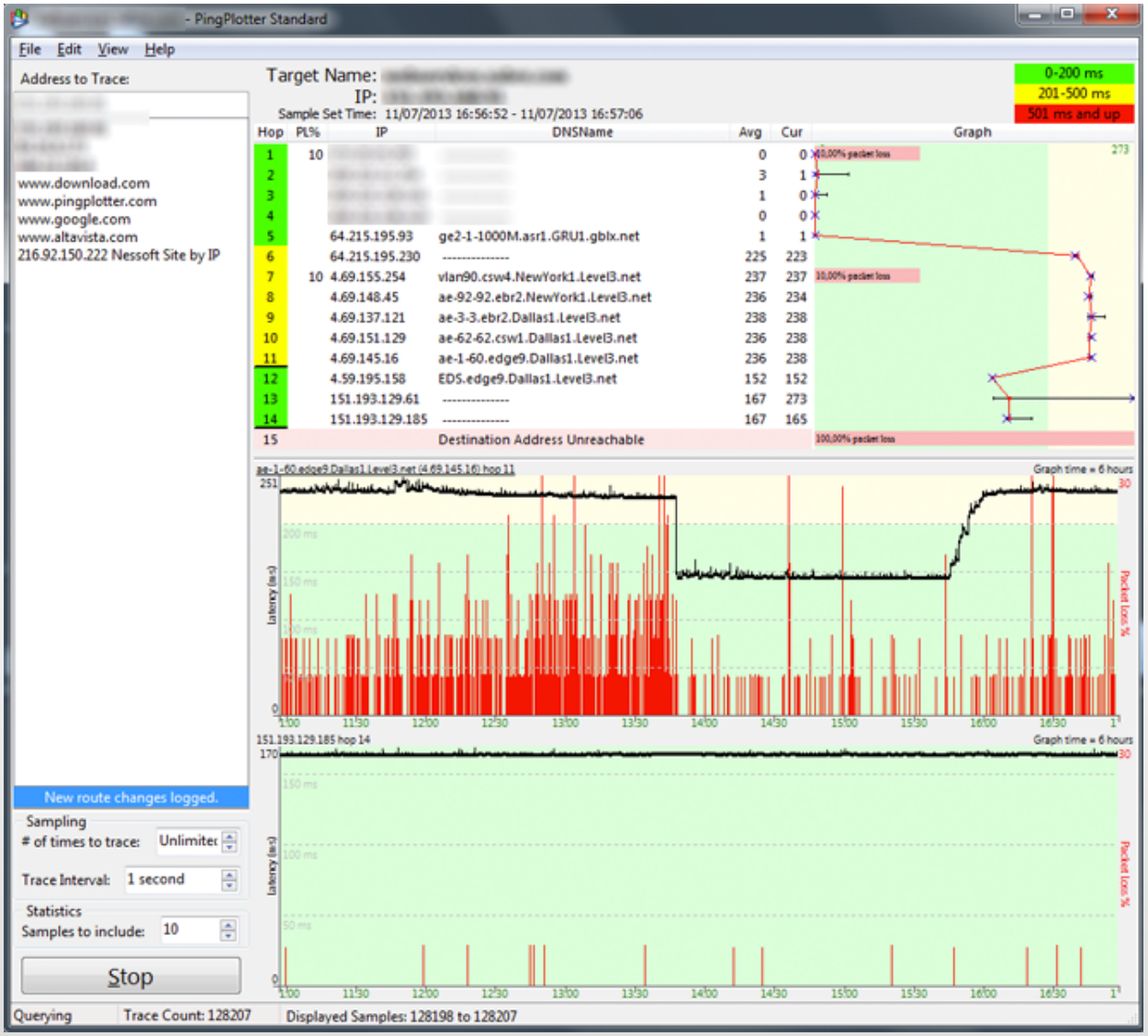
IPivot Telecom

Identificando caminhos diferentes

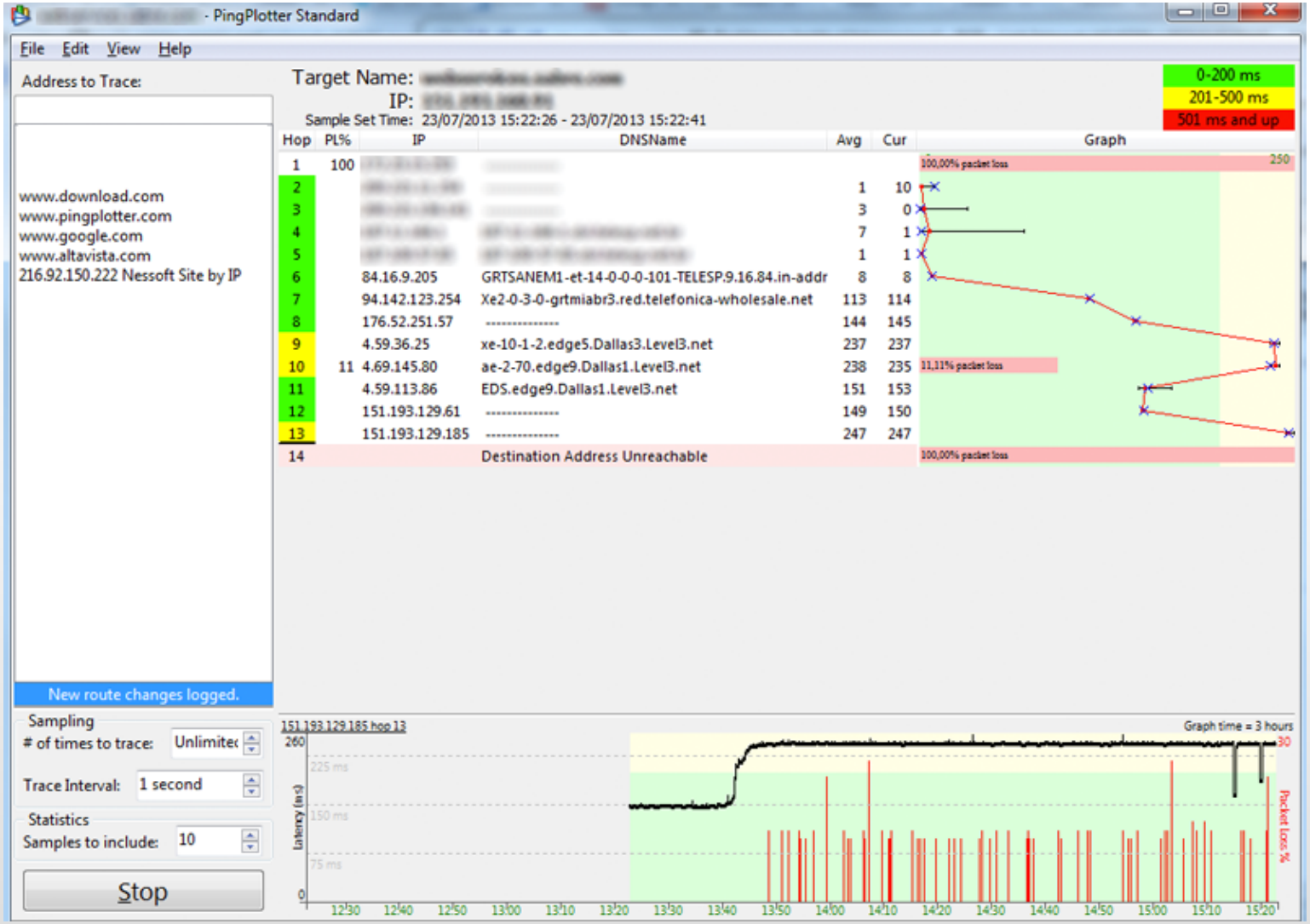
```
# ping 200.147.67.142
PING 200.147.67.142 (200.147.67.142) 56(84) bytes of data.
64 bytes from 200.147.67.142: icmp_seq=1 ttl=247 time=5.72 ms
64 bytes from 200.147.67.142: icmp_seq=2 ttl=247 time=9.45 ms
64 bytes from 200.147.67.142: icmp_seq=3 ttl=246 time=6.77 ms
64 bytes from 200.147.67.142: icmp_seq=4 ttl=246 time=6.04 ms
64 bytes from 200.147.67.142: icmp_seq=5 ttl=246 time=4.93 ms
64 bytes from 200.147.67.142: icmp_seq=6 ttl=246 time=7.52 ms
64 bytes from 200.147.67.142: icmp_seq=7 ttl=247 time=9.36 ms
64 bytes from 200.147.67.142: icmp_seq=8 ttl=247 time=11.3 ms
64 bytes from 200.147.67.142: icmp_seq=9 ttl=246 time=3.69 ms
64 bytes from 200.147.67.142: icmp_seq=10 ttl=247 time=3.31 ms
64 bytes from 200.147.67.142: icmp_seq=11 ttl=246 time=3.24 ms
64 bytes from 200.147.67.142: icmp_seq=12 ttl=246 time=4.84 ms
64 bytes from 200.147.67.142: icmp_seq=13 ttl=246 time=3.52 ms
64 bytes from 200.147.67.142: icmp_seq=14 ttl=246 time=3.21 ms
64 bytes from 200.147.67.142: icmp_seq=15 ttl=246 time=4.56 ms

--- 200.147.67.142 ping statistics ---
15 packets transmitted, 15 received, 0% packet loss, time 14005ms
rtt min/avg/max/mdev = 3.219/5.836/11.336/2.490 ms
```

Um caso real …


- PingPlotter Standard
File Edit View Help
Address to Trace:
www.download.com
www.pingplotter.com
www.google.com
www.altavista.com
216.92.150.222 Nessoft Site by IP
Target Name:
IP:
Sample Set Time: 11/07/2013 16:56:52 - 11/07/2013 16:57:06
0-200 ms
201-500 ms
501 ms and up
Hop PL% IP DNSName Avg Cur Graph
1 10 0 0
2 3 1
3 1 0
4 0 0
5 64.215.195.93 ge2-1-1000M.asr1.GRU1.gblx.net 1 1
6 64.215.195.230 225 223
7 10 4.69.155.254 vlan90.csw4.NewYork1.Level3.net 237 237
8 4.69.148.45 ae-92-92.ebr2.NewYork1.Level3.net 236 234
9 4.69.137.121 ae-3-3.ebr2.Dallas1.Level3.net 238 238
10 4.69.151.129 ae-62-62.csw1.Dallas1.Level3.net 236 238
11 4.69.145.16 ae-1-60.edge9.Dallas1.Level3.net 236 238
12 4.59.195.158 EDS.edge9.Dallas1.Level3.net 152 152
13 151.193.129.61 167 273
14 151.193.129.185 167 165
15 Destination Address Unreachable
New route changes logged.
Sampling
# of times to trace: Unlimited
Trace Interval: 1 second
Statistics
Samples to include: 10
Stop
Graph time = 6 hours
151.193.129.185 hop 14
Graph time = 6 hours
Querying
Trace Count: 128207
Displayed Samples: 128198 to 128207

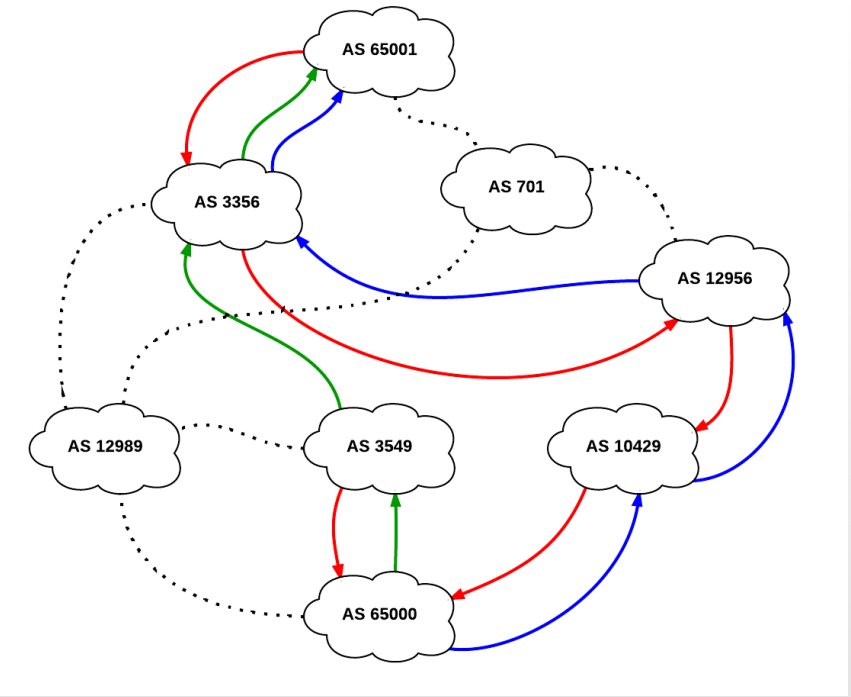
Um caso real …



PINPOINT

IPivot Telecom

Um caso real ...



- Ponto de vista: **AS65000**
- Traceroute:
  - Origem: AS65000
  - Destino: AS65001
- Consideramos que o tráfego de saída pode ser **controlado** e o tráfego de entrada pode ser apenas **influenciado**.
- Os dois caminhos de ida (**verde** de **azul**) não deixam claro qual é o problema.
- Entender que o retorno acontece somente pelo caminho **vermelho**, possibilita identificar o problema entre dois AS's.
- Formas de identificar o caminho no sentido contrário:
  - Perguntar ao administrador da rede remota?
    - INOC-DBA
    - E-mail
  - Looking-glass?
    - http://www.traceroute.org
  - Ferramentas de visualização do BGP?
    - http://bgp.he.net
    - https://stat.ripe.net/widget/bgplay

PINPOINT

IPivot Telecom

Boas Práticas

- Ao analisar a saída de um traceroute, lembre-se:
  - Na Internet o roteamento normalmente é assimétrico.
  - As possíveis perdas de pacote podem ser na verdade restrições de firewall ou ACL.
  - Saber qual ferramenta e quais parâmentros de protocolo utilizar nos testes com traceroute.
  - Importante escolher corretamente o endereço de origem quando o traceroute for disparado de um roteador.

- E atenção:
  - Para equipamentos intermediários que não decrementam TTL.
  - Redes que utilizam MPLS (label switching) para "roteamento" dos pacotes IP.

Perguntas?

Obrigado!

Contato:

gustavo@pinpoint.com.br

artur@pinpoint.com.br

PINPOINT

IPivot Telecom