

BGP Decision Statistics a First Experiment

RIPE / Roma

2010.11.17

Randy Bush <randy@psg.com>
Cristel Pelsser <cristel@iij.ad.jp>
Keyur Patel <keyupate@cisco.com>
Olaf Maennel <olaf@maennel.net>

<<http://archive.psg.com/101117.ripe-statistics.pdf>>

Motivation

- BGP Decision Process presumed to be costly in CPU and time
- The further down the tie-break the more has been spent
- Do we know where the decisions are actually being made?

March 2010, Keyur Agreed to a Hack

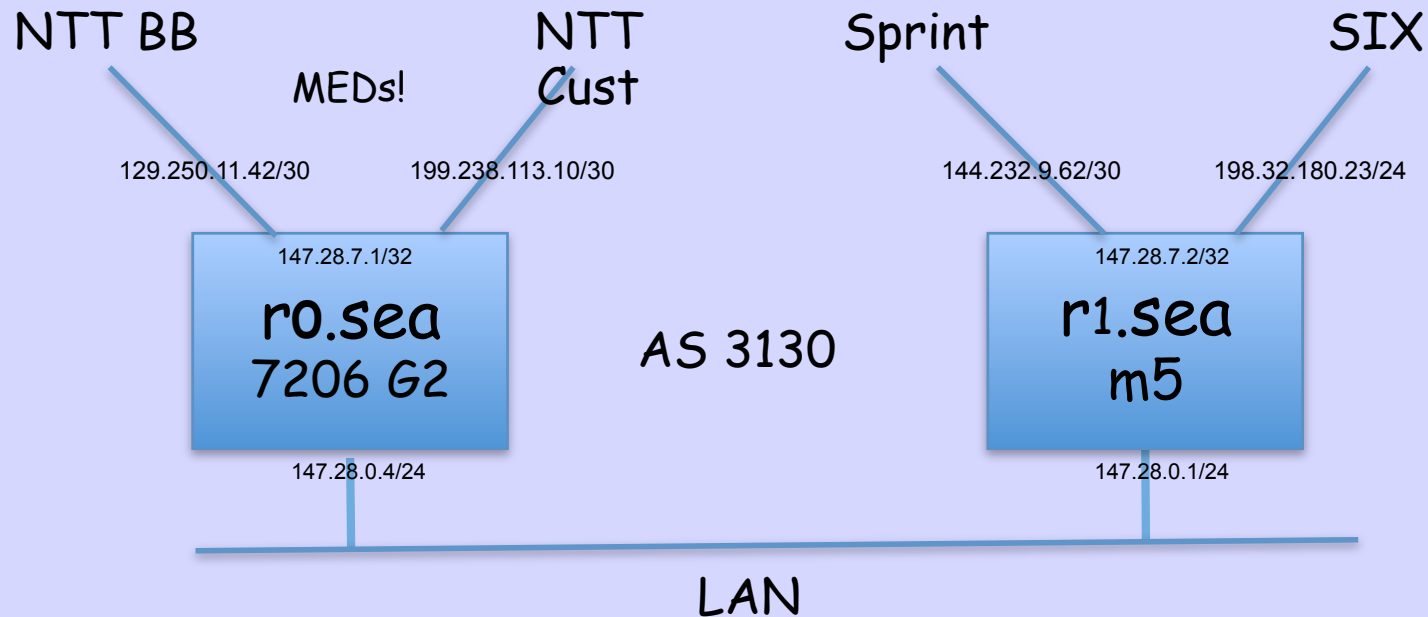
```
r0.sea#sh ip bg internal
Time left for bestpath timer: 12 secs
Address-family IPv4 Unicast, Mode : RW
  Table Versions : Current 625197, RIB 625197
  Start time : 00:00:12.804    Time elapsed 07:43:56.300
  First Peer up in : 00:00:20    Exited Read-Only in : 00:01:31
  Done with Install in : 00:01:38    Last Update-done in : never
  3331 updates expanded
  Attribute list queue size: 9357
  Slow-peer detection is disabled
  BGP Nexthop scan:-
    penalty: 0, Time since last run: never, Next due in: none
    Max runtime : 0 ms Latest runtime : 0 ms Scan count: 0
  BGP General Scan:-
    Max runtime : 1724 ms Latest runtime : 352 ms Scan count: 458

  BGP future scanner version: 459
  BGP scanner version: 0

  BGP Attribute counters:
    BGP Attr Metric changed count 0
    BGP Attr Origin changed count 47848
    BGP Attr MED changed count 472530
    BGP Attr Aspath changed count 172996
```

2010.10.18 the Cisco internal copycat complicators discovered it, see
draft-raszuk-bgp-diagnostic-message-00.txt

Westin Set-Up



Neighbor	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
129.250.11.41	2914	915011	23516	3581566	0	0	1w0d	325999
147.28.7.2	3130	485431	1133945	3581587	0	0	1w0d	254777
199.238.113.9	2914	1410879	11870	3581566	0	0	1w0d	326000

Attribute Changes - Westin

Metric	0	0.00%
Origin	47,848	3.06%
MED	472,530	30.25%
Aspath	172,996	11.08%
LocalPref	0	0.00%
Aggregator	3,701	0.24%
Atomic Aggregator	417	0.03%
Cluster List	0	0.00%
Originator	0	0.00%
Community	313,099	20.05%
Ext Community	11	0.00%
SSA	0	0.00%
Connector	0	0.00%
Nexthop	828	0.05%
Label	0	0.00%
Dampening	0	0.00%
Dmz Bw	0	0.00%
New	434,072	27.79%
Duplicate	116,439	7.45%

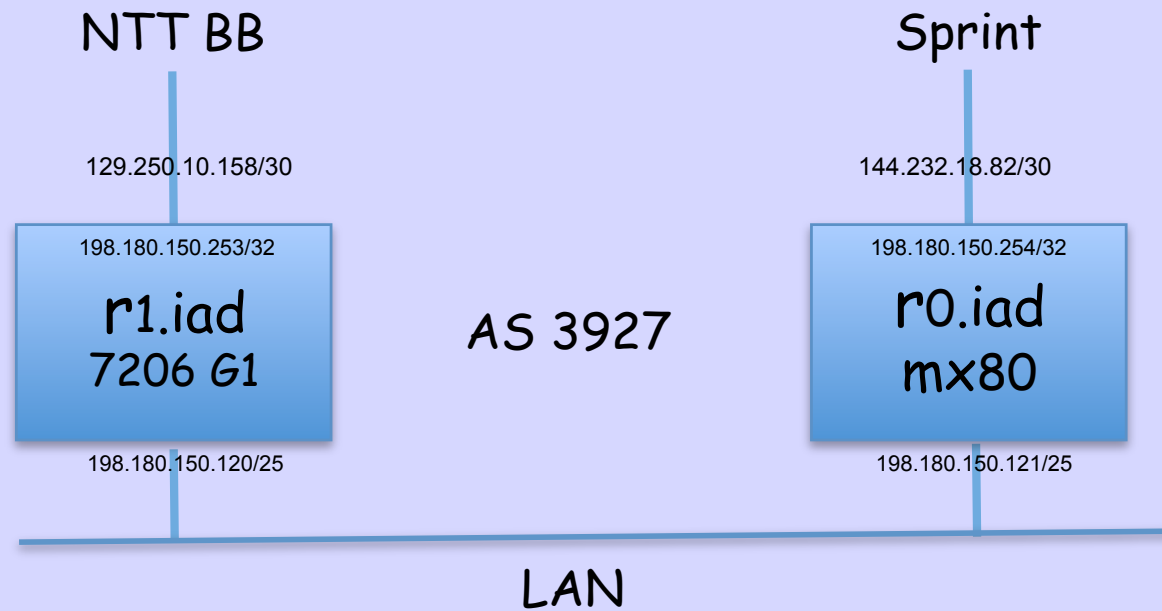
Best Path - Westin

	Decision Made		New Best Path	
Cost	0	0.00%	0	0.00%
Weight	4,458	0.00%	0	0.00%
LocalPref	0	0.00%	0	0.00%
AsPath	105,746,367	4.26%	1,497,266	31.31%
Origin	1,341,944	0.05%	99,129	2.07%
Med	882,802,290	35.55%	3,156,239	66.01%
IGP	0	0.00%	0	0.00%
RouterID	14,999,026	0.60%	28,876	0.60%
ClusterID	0	0.00%	0	0.00%
Missing	###	59.54%	0	0.00%
OnlyPath	0	0.00%	0	0.00%
NbrAddr	0	0.00%	0	0.00%

Why the Big Missing Value

- Because of the two links to NTT, deterministic-med is enabled
- Scanner runs bestpath computation for the entire table every ten minutes
- And the counter gets bumped at least once for every net that is evaluated

Ashburn Set-Up



Attribute Changes - Ashburn

Metric	0	0.00%
Origin	1,466	0.46%
MED	26,165	8.22%
Aspath	47,634	14.97%
LocalPref	0	0.00%
Aggregator	1,667	0.52%
Atomic Aggregator	148	0.05%
Cluster List	0	0.00%
Originator	0	0.00%
Community	50,885	15.99%
Ext Community	0	0.00%
SSA	0	0.00%
Connector	0	0.00%
Nexthop	0	0.00%
Label	0	0.00%
Dampening	0	0.00%
Dmz Bw	0	0.00%
New	187,424	58.89%
Duplicate	2,860	0.90%

Best Path - Ashburn

	Decision		New Best Path	
Cost	0	0.00%	0	0.00%
Weight	0	0.00%	0	0.00%
Local Pref	0	0.00%	0	0.00%
Aspath	3,310,422	91.47%	71,910	98.10%
Origin	76,688	2.12%	1,392	1.90%
Med	0	0.00%	0	0.00%
IGP	0	0.00%	0	0.00%
Router ID	0	0.00%	0	0.00%
Cluster ID	0	0.00%	0	0.00%
Missing	0	0.00%	0	0.00%
only path	232,006	6.41%	0	0.00%
nbr addr	0	0.00%	0	0.00%

Tentative Conclusions

- Need to run in a richer topology
- But current code is only 7200
- Other images may be coming
- Would you run for us?
- Would you run for yourself?!