

BGPデシジョン統計

～1st実験～

JaNOG / Kanazawa

2011.01.20

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/110121.janog-statistics.pdf>>

モチベーション

- BGPの経路選択プロセスはCPUと時間を消費する「高コスト」プロセス
- 経路選択プロセス順番の下にいくほどより時間がかかっている事を意味する
- 実際にどこが選択ポイントになっているのか我々は知っているのか？

3月 2010, Keyurさんが特殊バージョン を作ってくれた

```
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

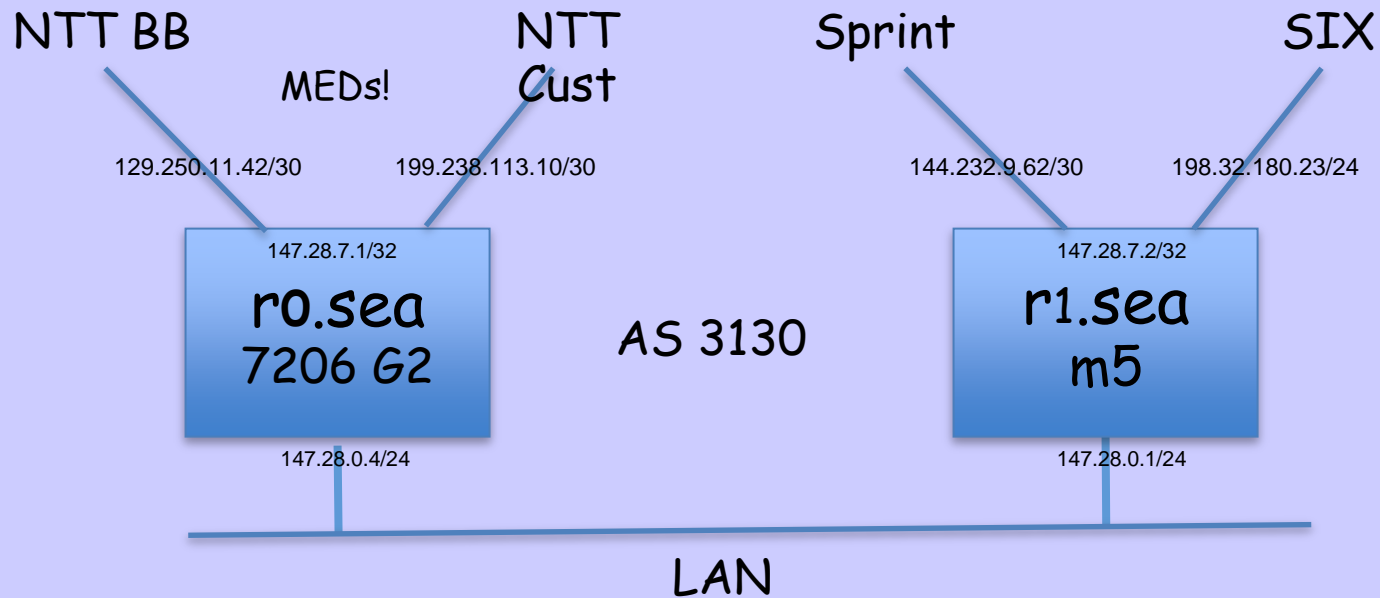
  BGP future scanner version: 459
  BGP scanner version: 0

  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 Cisco内の者に発見されてしまう。以下参照
draft-raszuk-bgp-diagnostic-message-00.txt

Westin 構成

NTTと2本でつながっているためMEDが有効



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

属性変更 - Westin

Metric	0	0.0%
Origin	825,088	1.7%
MED	15,900,280	32.3%
ASpath	9,361,386	19.0%
LocalPref	0	0.0%
Aggregator	202,911	0.4%
Atomic Aggregator	13,771	0.0%
Cluster List	0	0.0%
Originator	0	0.0%
Community	13,535,923	27.5%
Ext Community	2,453	0.0%
SSA	0	0.0%
Connector	0	0.0%
Nexthop	90,131	0.2%
Label	0	0.0%
Dampening	0	0.0%
DMZ BW	0	0.0%
New	5,393,812	11.0%
Duplicate	3,862,781	7.9%

Best Path - Westin

経路変更基準となったもの

新ベストパスになったもの

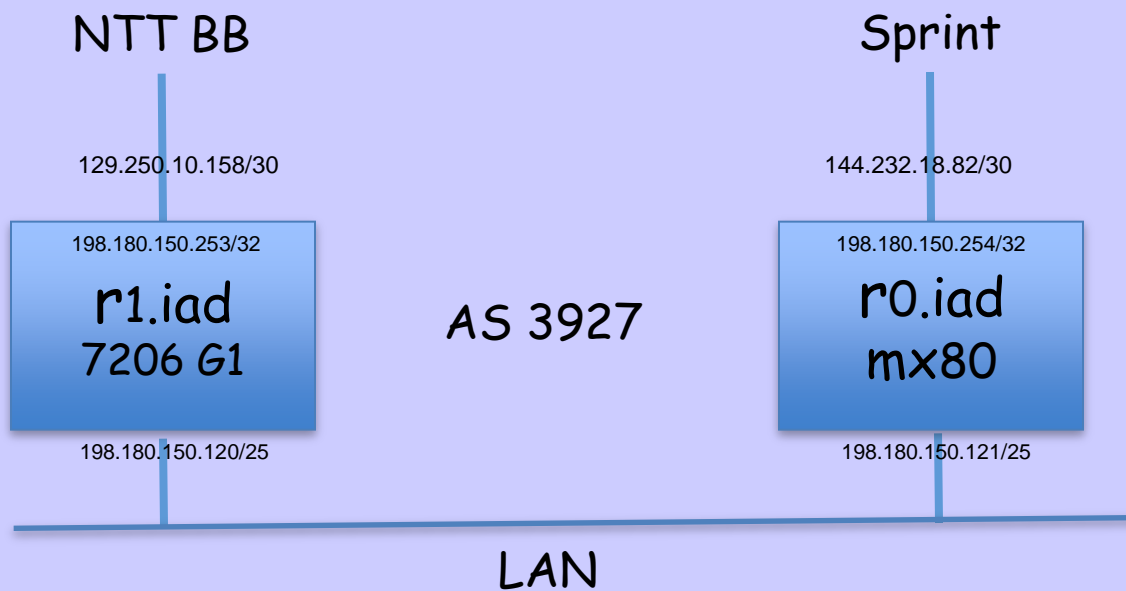
Cost	0	0.00%	0	0
Weight	7,680	0.00%	0	0.0%
Local Pref	0	0.00%	0	0.0%
ASpath	187,915,428	4.62%	2,503,892	31.1%
Origin	2,218,073	0.05%	142,583	1.8%
Med	1,412,154,425	34.72%	5,348,040	66.5%
IGP	0	0.00%	0	0.0%
Router ID	24,827,335	0.61%	43,741	0.5%
Cluster ID	0	0.00%	0	0.0%
Missing	2,440,555,210	60.00%	0	0.0%
Only Path	0	0.00%	0	0.0%
NBR Addr	0	0.00%	0	0.0%

なぜ25億も Missingがあるか？

- NTTと2本の回線で接続していて、
deterministic-med が有効
- Scanner が10分に1回best path計算をし直す
- Prefix毎にカウンタをどこかで計上していてそれがMissingにあがっている

Ashburn 構成

NTTへの接続は1本なのでMEDなし



属性変更 - Ashburn

Metric	0	0.0%
Origin	329,549	2.0%
MED	3,592,634	21.7%
ASpath	4,301,742	26.0%
LocalPref	0	0.0%
Aggregator	460,413	2.8%
Atomic Aggregator	8,323	0.1%
Cluster List	0	0.0%
Originator	0	0.0%
Community	4,726,186	28.5%
Ext Community	0	0.0%
SSA	0	0.0%
Connector	0	0.0%
Nexthop	0	0.0%
Label	0	0.0%
Dampening	0	0.0%
DMZ BW	0	0.0%
New	1,604,910	9.7%
Duplicate	1,551,500	9.4%

Best Path - Ashburn

	経路変更		新しいベストパス	
Cost	0	0.0%	0	0.0%
Weight	0	0.0%	0	0.0%
Local Pref	0	0.0%	0	0.0%
ASpath	312,315,324	97.3%	1,972,413	97.8%
Origin	7,395,908	2.3%	44,022	2.2%
Med	0	0.0%	0	0.0%
IGP	0	0.0%	0	0.0%
Router ID	0	0.0%	0	0.0%
Cluster ID	0	0.0%	0	0.0%
Missing	0	0.0%	0	0.0%
Only Path	1,114,010	0.3%	0	0.0%
NBR Addr	0	0.0%	0	0.0%

暫定結論

- もっと豪華なトポロジーで分析しないといけない
- 7200で動くコードしかない
- 他で動くものもできるかも
- 我々の代わりにコードを動かしてもらえますか？
- 自分たちでコードを動かして分析してみたいと思いますか？