Cache Hit Analysis

EOF / Istanbul

2006.04.25

Olaf Maennel <olaf@maennel.net>
University of Adelaide

Motivation

Caching only makes sense if we see updates in regular intervals again and again…

So, how often do we see updates again…

1. on a per prefix / per peer basis?
2. across different feeds (including iBGP), but still on a per prefix basis?
3. across prefixes
   (what a cache finally has to deal with)

Analyzed data

Data-sets: *RRC 00-15 + GEANT*

Time: **(start)** Fri Feb 3 16:24:27 2006

(start of study) Sun Feb 5 16:24:27 2006

(end of study) Sat Mar 4 05:29:09 2006

(remove edge-effects: ~ 2 days)

Prefixes: 220,708 prefixes

(most de-aggregated view;
including 1,229 IPv6;
including 16k: /25-/32; 11k: /30-/32;…)

Emulating workload of a router

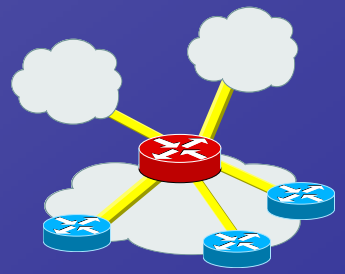
i-BGP workload *("GEANT")*:

GEANT collector records i-BGP feeds from:

- 23 PoPs

GEANT has 2 upstreams: (Telia + Level 3)

- 5 PoPs give full table (121k-151k)
- remaining announce only
  6-600 prefixes per session



additional e-BGP workload *("RIPE")*:

- "full"-feeds (>100k) : up to 60 feeds (+0, +5, +10)
- "peering"-feeds (>2k) : up to 34 feeds (+5,+15,+34)
- "customer"-feeds (<2k) : up to 275 feeds (+100)

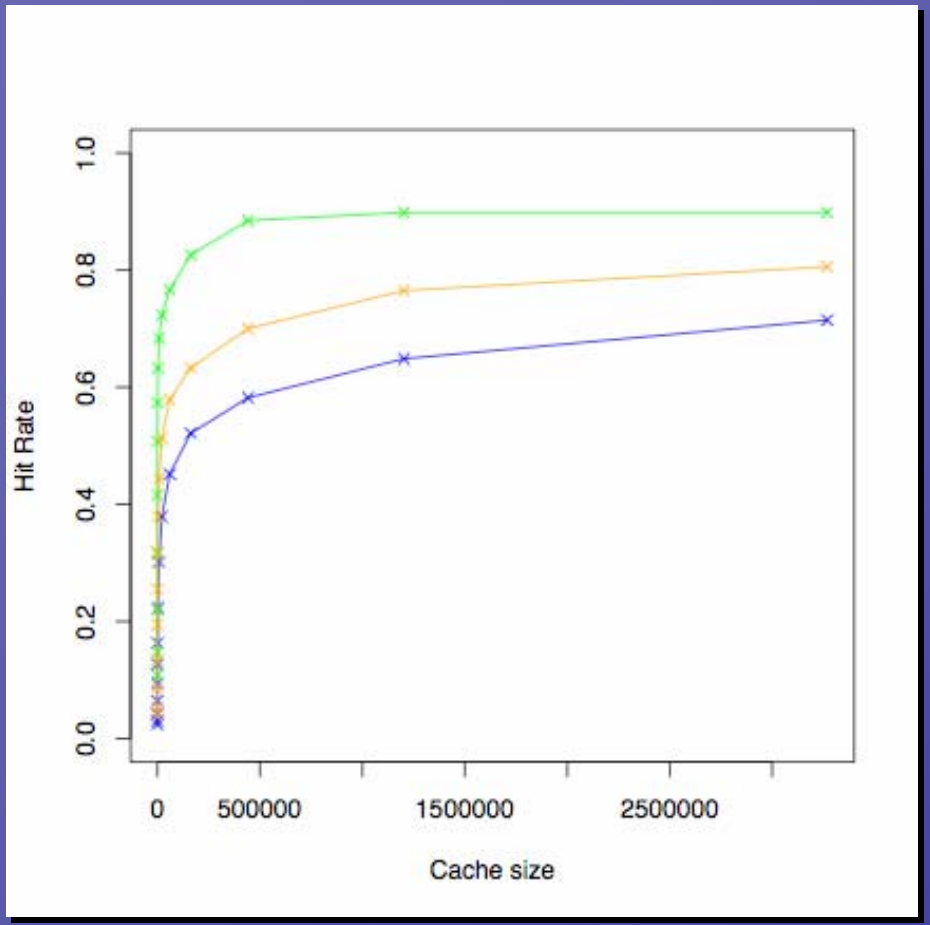
Motivation

Caching only makes sense if we see updates in regular intervals again and again…

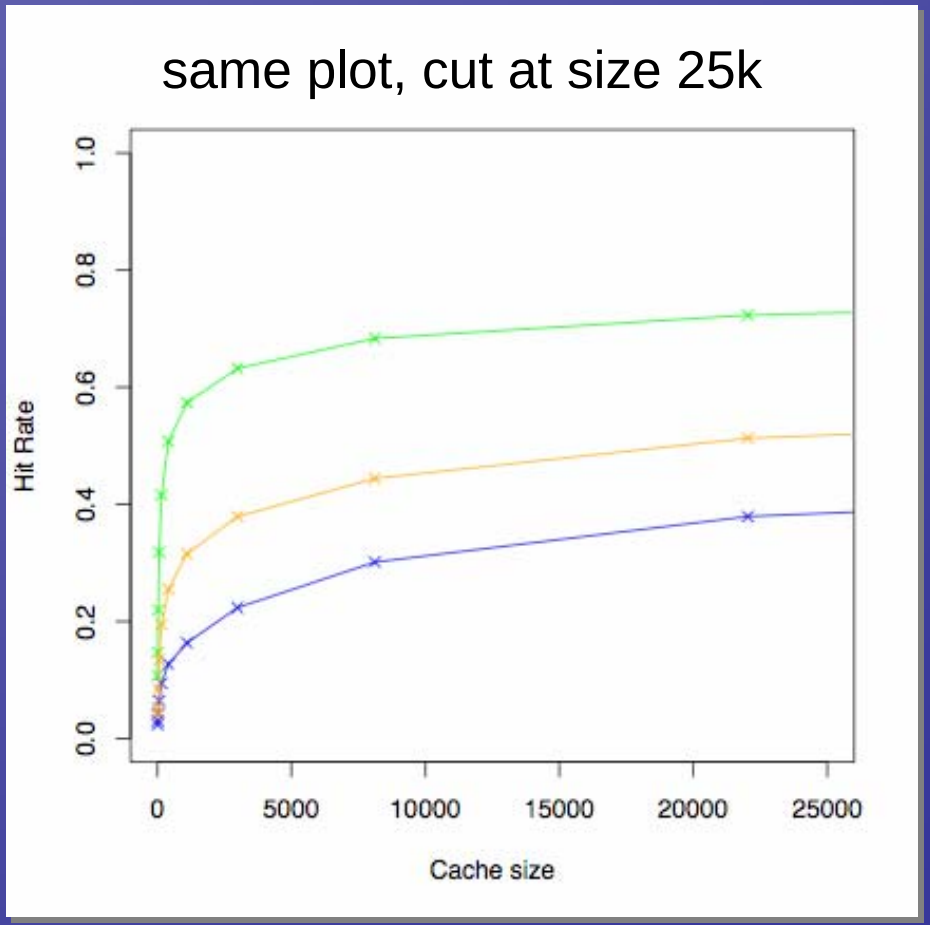
So, how often do we see updates again…

1. on a per prefix / per peer basis?
2. across different feeds (including iBGP), but still on a per prefix basis?
3. across prefixes
   (what a cache finally has to deal with)

Cache Size vs. Hit Rate

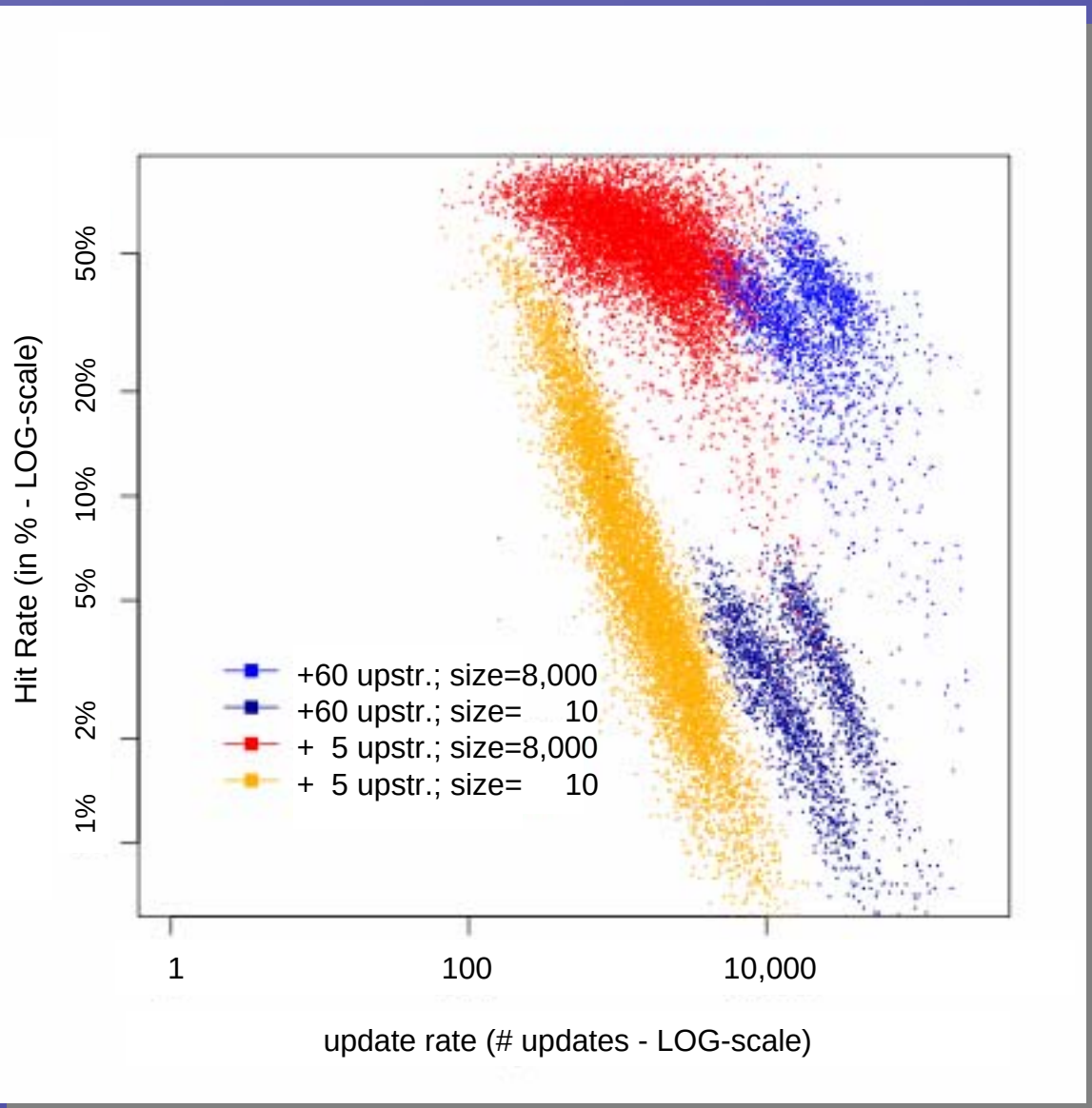


same plot, cut at size 25k



green: iBGP data only (+0 added feeds)
orange: iBGP data +5 upstreams added
blue: iBGP data + all data (~60 upstreams added)

Hit Rate vs. Update Rate



Correlation between hit rate and update rate shown in 5 min bins.

Amazing variation in hit rates over time!

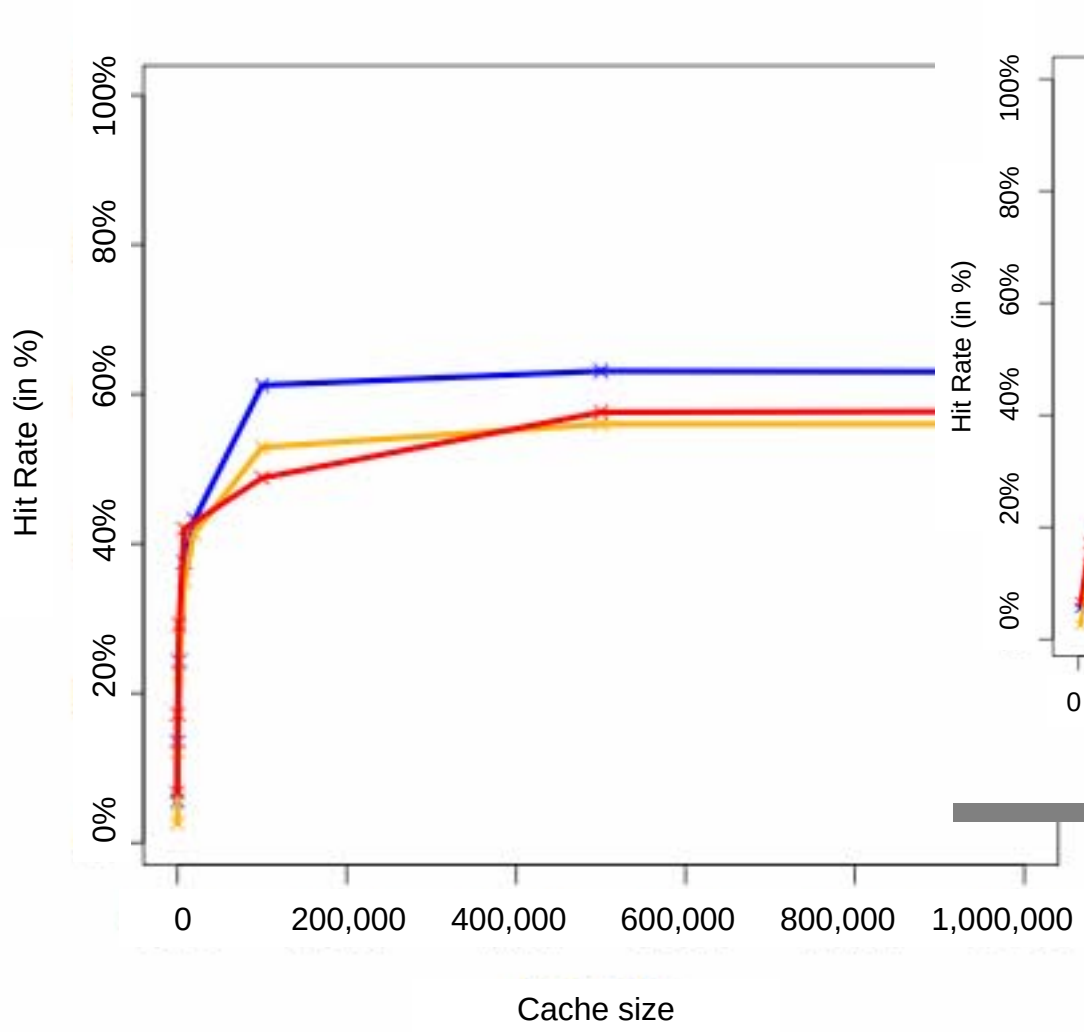
Hit Rate depends on the number of updates on peers.

e-BGP only study

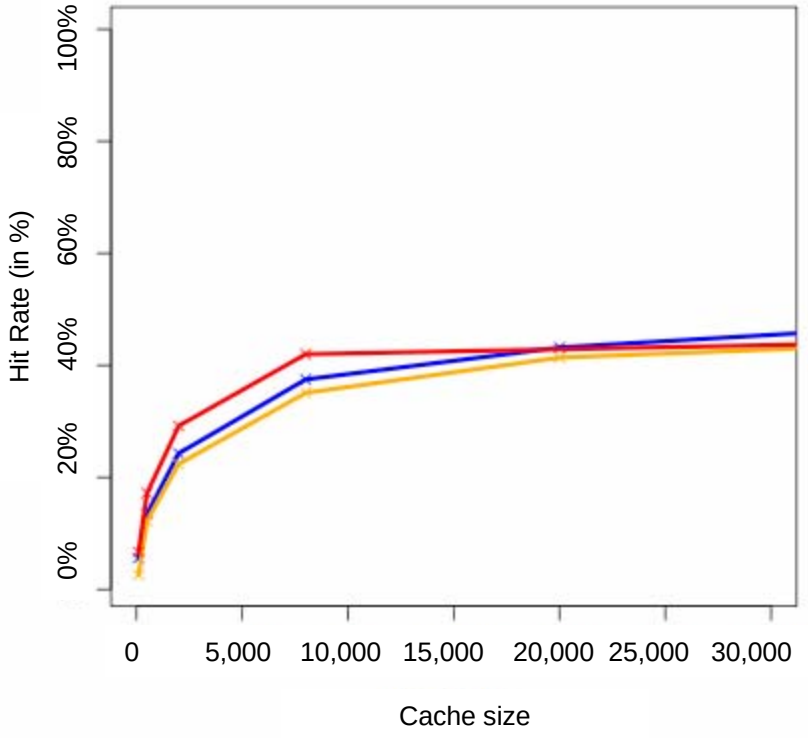
- High correlation between feeds in iBGP show good hit-rates. Yet, iBGP updates do not necessarily have to be validated (again).

- How does cache perform in e-BGP only scenarios?

Cache Size vs. Hit Rate



same plot, cut at size 30k



blue: 1 upstream
orange: 2 upstreams
red: 7 upstreams

NO iBGP data!

Summary

- Cache size and hit rate seem to follow a logarithmical distribution!

Hit rate strongly depends on:

- Type of neighbors (e.g., number of prefixes, thus upstreams have worst hit-rates)
- Number of neighbors
- Number of updates received at certain times.