Correlating routing configuration changes with forwarding changes

David Lebrun

IIJ Innovation Institute

September 25, 2012

Objectives

- Determine if we can correlate routing conf changes with forwarding changes
- Routing configuration data obtained from a Tier-1 ISP
- Forwarding changes data obtained from own measurements
- We want to measure latency and path changes
- The goal is to detect eBGP events

Objectives

- What measurements ?
  - Pings for latency
  - Traceroutes for path changes
- From where ?
  - Servers
  - Atlas probes
- To what ?
  - ISP's routers or some reachable IP

About Atlas probes

- The probes are distributed world wide within plenty of ASes
- Currently about 1,500 probes are active
- In theory, perfect tool to measure a worldwide ISP
- We need to investigate the capacities and limitations of the Atlas probes

Material overview

- What do we have ?
  - Syslog of routers
  - CVS of configurations (RANCID, runs every two hours)
  - A list of thousands of reachable IPs in the neighborhood of the ISP
  - The list of all active Atlas probes
  - Three dedicated servers with ISP as transit provider (Seattle, Ashburn, Dallas)

Organizing the data

- Prepare reachable IPs
  - Perform traceroutes to each IP to have the exit POP
  - Remove IPs for which there is only one hop in ISP
  - Clusterize IPs with respect to their exit POP
- Prepare probes list
  - Take probes belonging to neighboring ASes
  - For all probes, perform a traceroute to some random reachable IPs
    - So that we can find the entry POP(s)
  - Clusterize the probes w.r.t. their entry POP

Measuring

- Main idea
  - For each probes cluster, select a probe
  - For each selected probe, for each IP cluster, select an IP
  - Ping or traceroute the IP from the probe
  - Repeat periodically

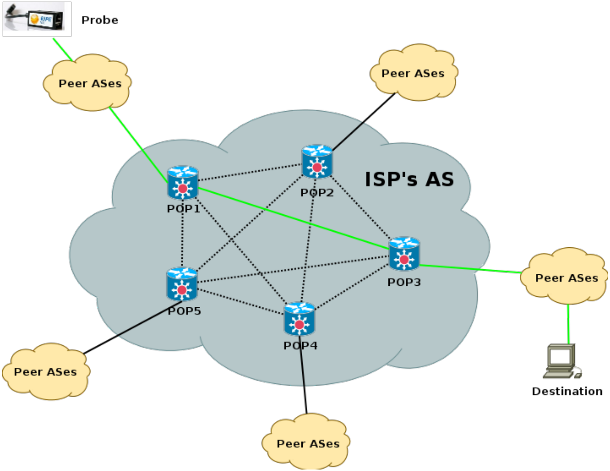
Ping issues

- Reminder: we want to detect eBGP events
- If we ping IPs in the neighborhood (IPs in the previously computed clusters)
  - Will detect unrelated events
  - No guarantee that the ping will go through ISP unless route is enforced by provider (localpref)
  - eBGP events can originate from farther in the as-path
  - The larger the distance from the target IP and the ISP, the more the noise will increase

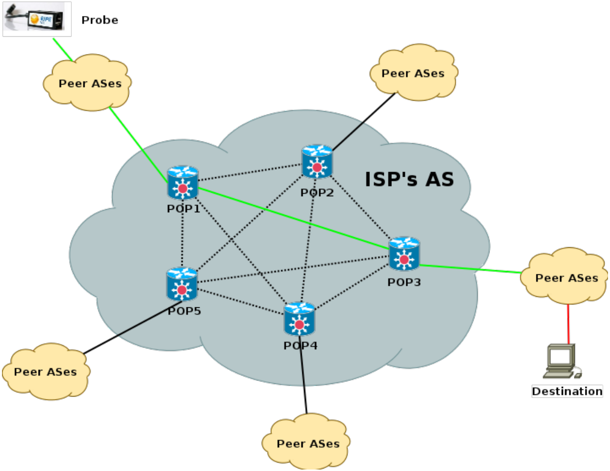
Ping issues

- If we ping ISP's border routers or direct neighbors
  - Difficult to detect eBGP events
  - But we can detect internal changes
  - We can also detect congested links due to eBGP events

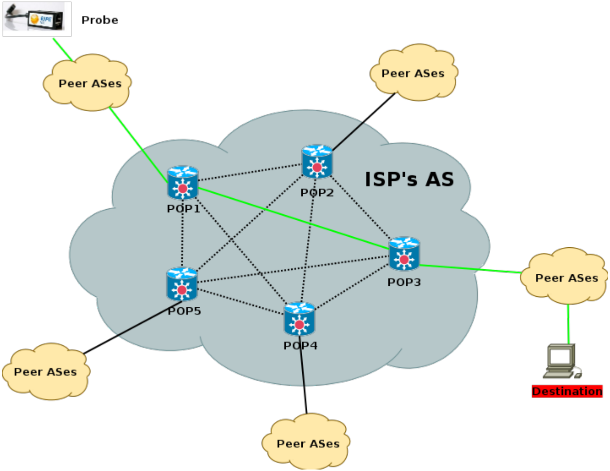
Illustration



Unwanted event



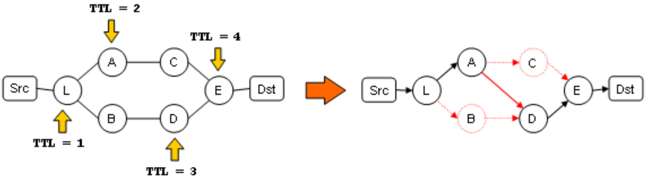
Unwanted event



Traceroute issues

- Big issue with standard traceroute: load balancers
- Standard techniques cannot properly handle equal-cost multipaths
- Leads to detection of non existent links

Traceroute issues



Source: http://www.paris-traceroute.net/about

Solution

- To solve this problem: paris-traceroute
- Use a Multipath Detection Algorithm
- Basic idea
  - Keep constant tuple (srcIP, srcPort, dstIP, dstPort)
  - Mark packets with fields not used to distinguish flows

Paris traceroute

- Now we have a multipath-aware traceroute
- No longer affected by per-flow load balancers

Probes issues

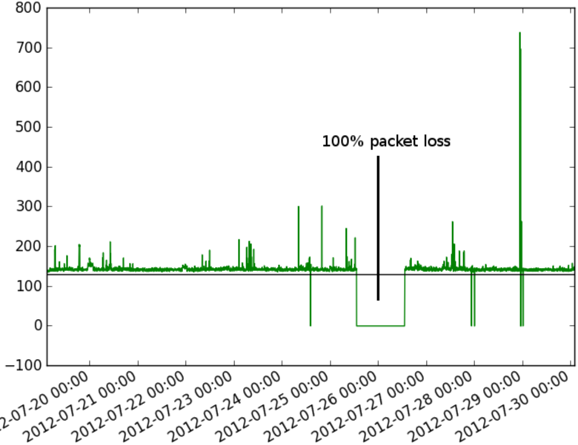
- Most of probes do not use ISP as default provider
  - Some pairs (probe,IP) go through ISP
  - Most of these probes are more than 4 hops away from ISP (and more than 2 ASes away)
- High rate of path changes
  - The probes don't keep constant params for same dest across different traceroutes

Probes issues

- Lot of noise in the ping data
  - High stddev
  - Huge latency peaks
  - Missing data

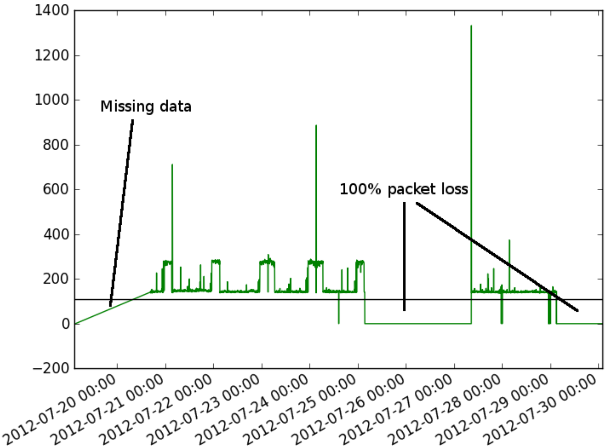
Probes issues

Noisy probe data



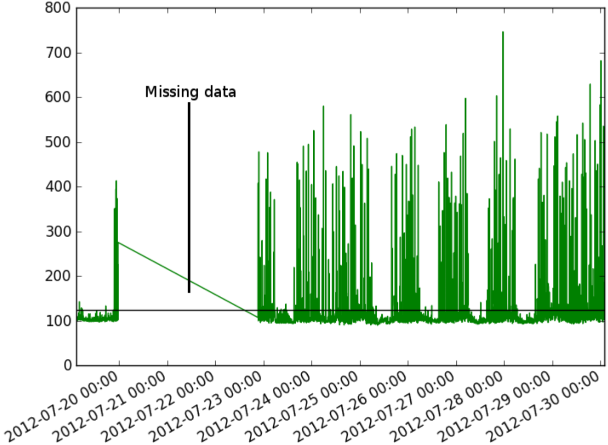
Probes issues

Noisy probe data



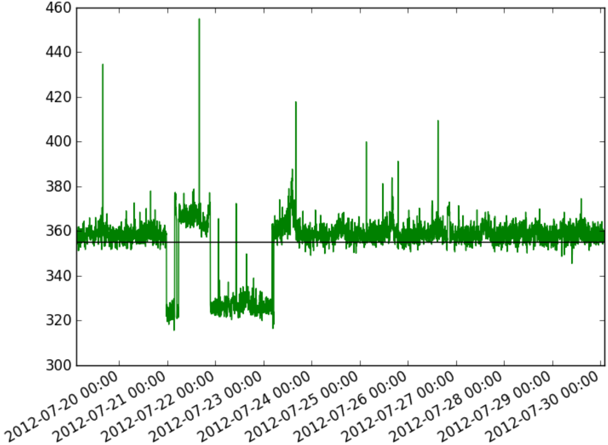
Probes issues

Noisy probe data



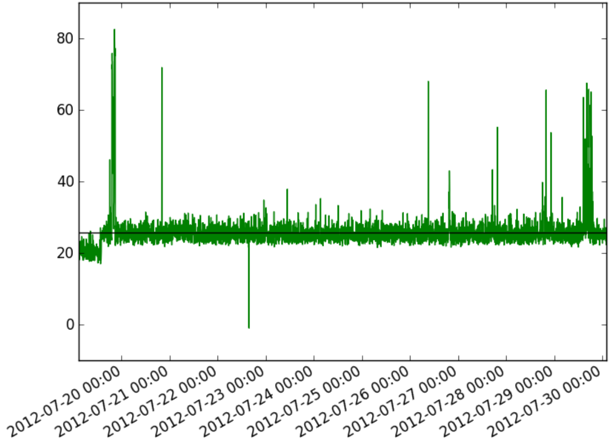
Probes issues

Clean probe data

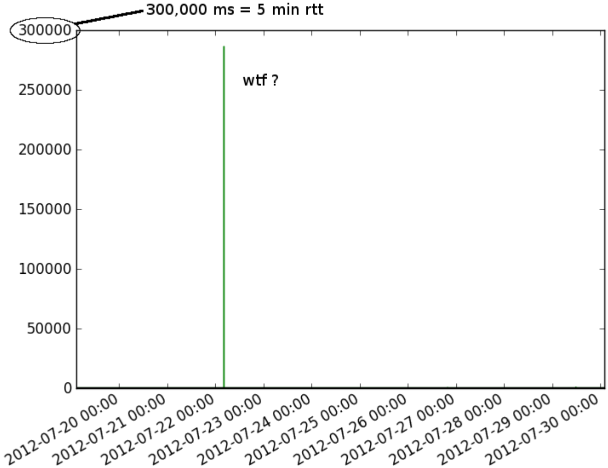


Probes issues

Clean probe data

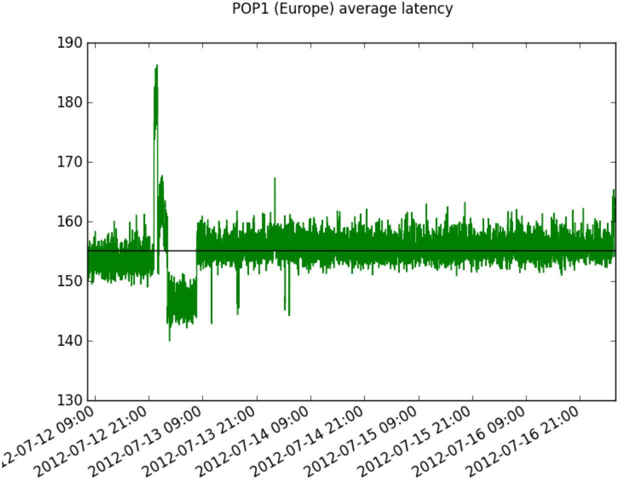


Probes issues



Results

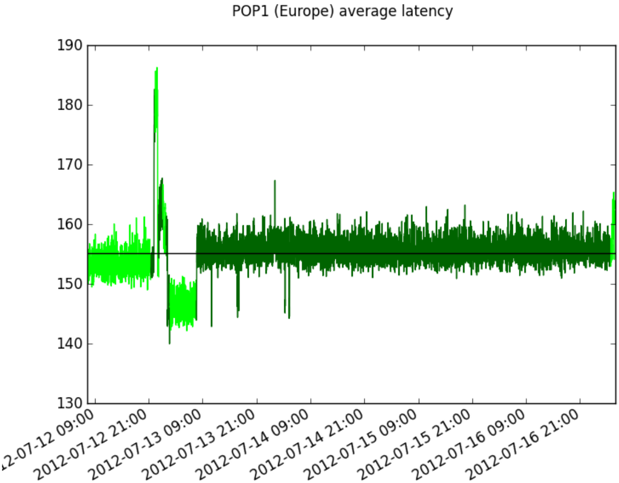
Ping results



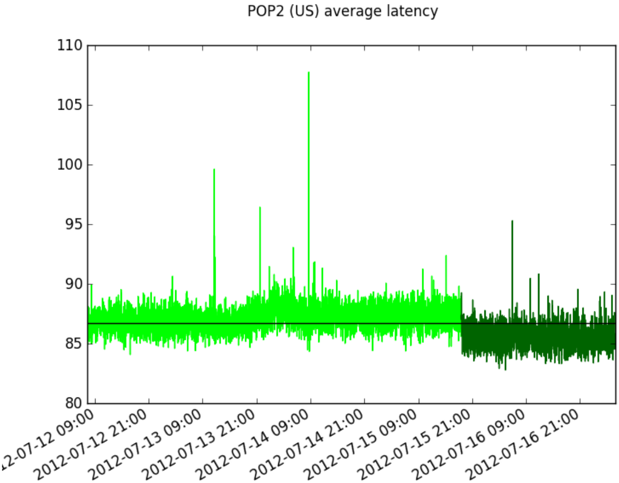
Period detection

- We would like to detect the different periods
- Some heuristic is used to automate this

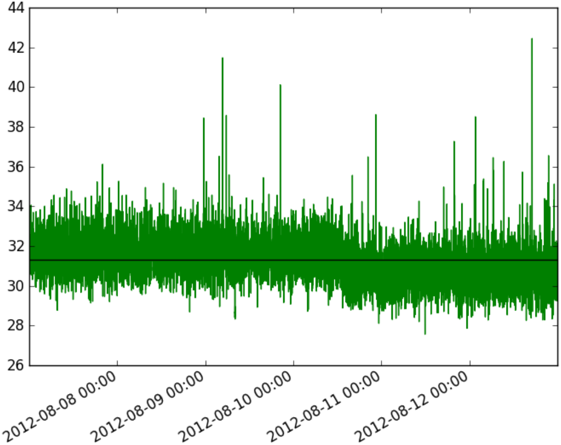
Period detection



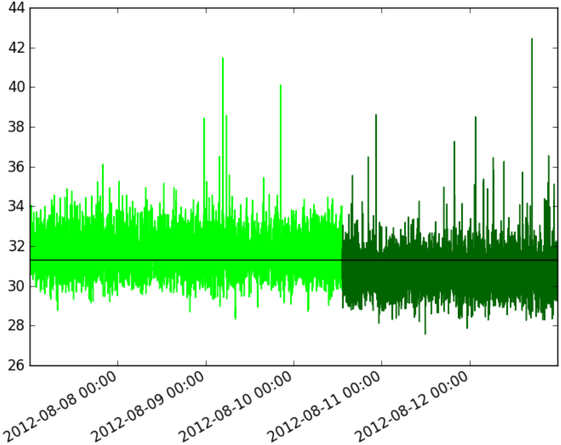
Period detection



Period detection

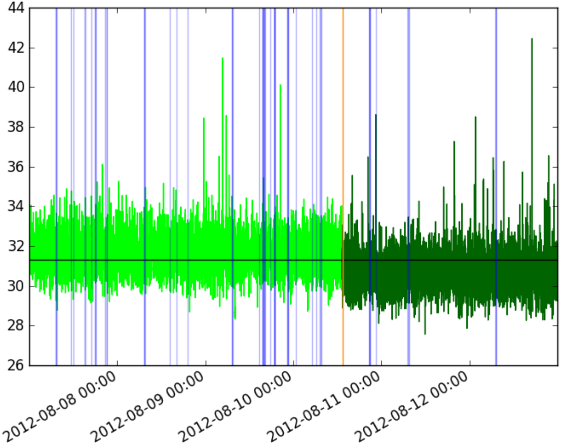


Period detection



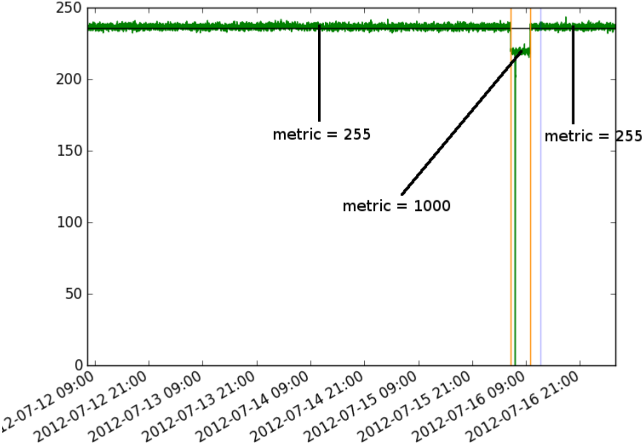
Correlation with commits

Orange line = commit near transition. In this case, addition of a link to an aggregated SONET link



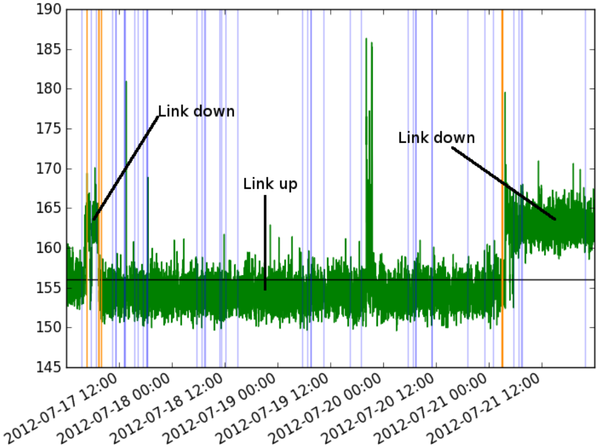
Correlation with commits

IGP metric increase/decrease



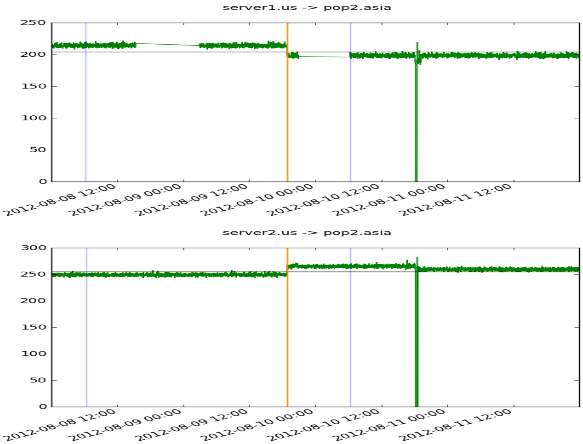
Correlation with commits

Link maintenance



Correlation with commits

MPLS Label-Switched Path configuration change



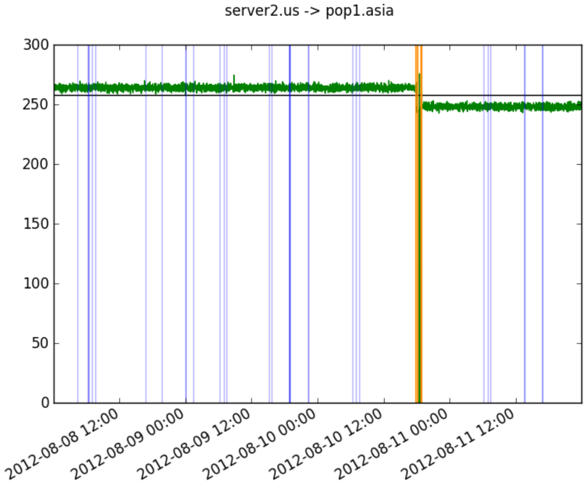
Correlation with commits

What happened here ?

- Change in bandwidth allocation of some LSPs
- Between server1.us and pop2.asia
    - Bw server1.us → pop2.asia: +38%
    - Bw pop2.asia → server1.us: +300%
    - RTT decreased
- Between server2.us and pop2.asia
    - Bw server2.us → pop2.asia: -7%
    - Bw pop2.asia → server2.us: +90%
    - RTT increased

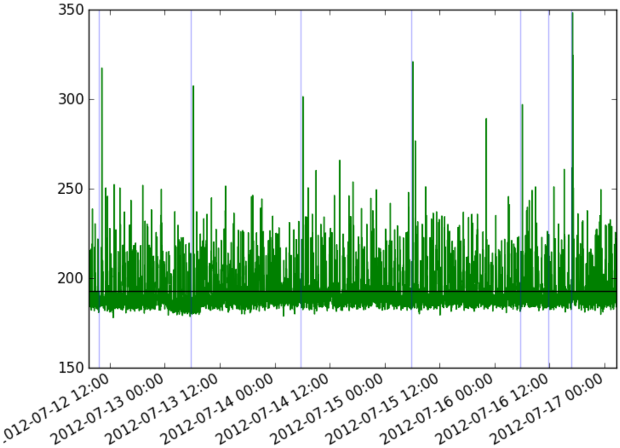
Correlation with commits

Router upgrade



Correlation with commits

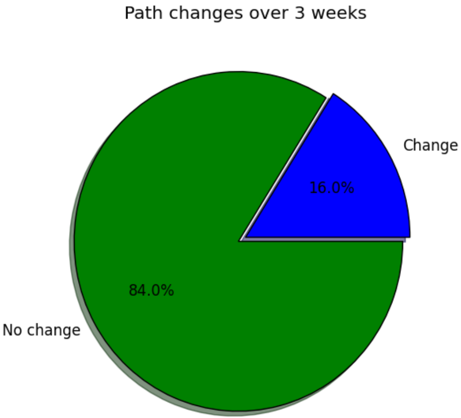
RTT peak: prefix-list change



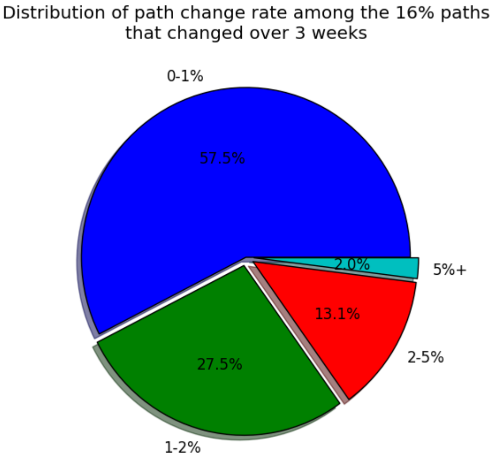
Traceroutes

- Traceroutes performed from servers, to 5,300 targets
- Batch run every hour
- About 16% of the paths changed at least once

Traceroutes



Traceroutes

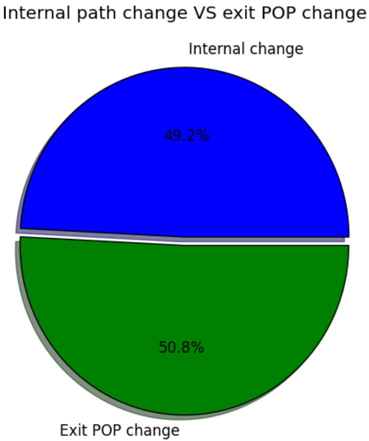


Traceroutes

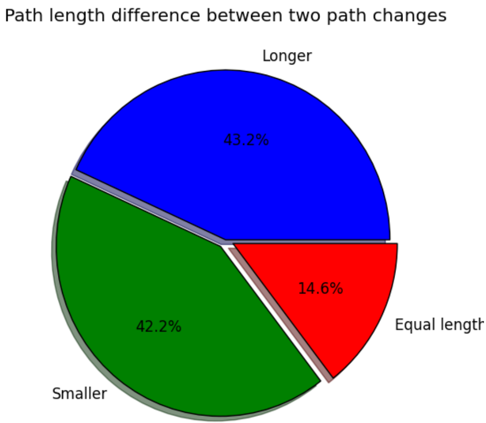
Two classes of traceroutes

- Internal path change
  - IGP change, LBs, link/router failure
- Exit POP change
  - eBGP change inside/outside ISP's AS, link/router failure

Traceroutes

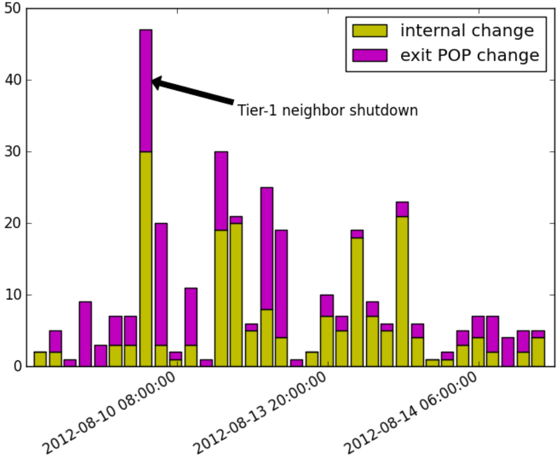


Traceroutes

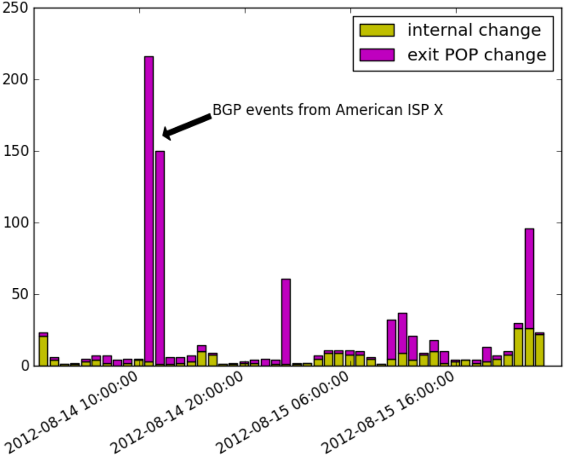


Traceroutes

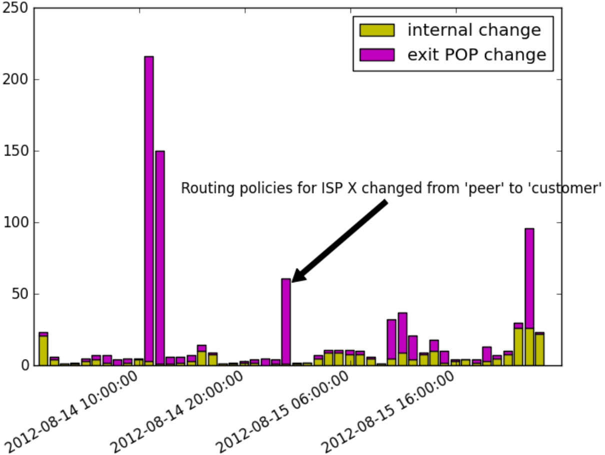
Aggregated data showing the number of path changes in fct of time



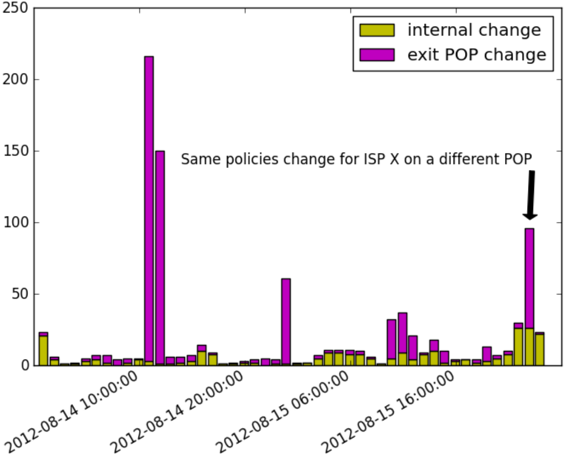
Traceroutes



Traceroutes



Traceroutes



Summary of the measurements

- Most events are IGP related
- Most of IGP events are link maintenances
- Less eBGP events than expected
- Most of eBGP events come from outside the ISP's AS and thus cannot be correlated with a configuration change
- Important events are easily detected even from a few sources
- Atlas probes need more work to be really usable in a project of this scale

Conclusion

- Prefix-list updates can generate RTT peak on the router
- Interface shutdown causes permanent RTT change and internal path changes
- MPLS changes can cause RTT changes
- eBGP events (from the inside or the outside) cause exit POP change (unsurprisingly)

Conclusion

- IGP events easy to detect and correlate
- Significant eBGP events are easy to detect
- More difficult to correlate as it can originate from the outside

Conclusion

What was cool with the Atlas probes ?

- The JSON interface is very useful for scripting
- The ability to manually specify multiple probes for a single UDM
- Geographical dispersion of the probes
- Overall, the system works quite well
- The Atlas team for granting us credits :)

Conclusion

What can be improved ?

- Handling of bulk measurements (probes per UDM and UDMs per probe)
- Fetching results (interface response time)
- Paris-traceroute implementation

Conclusion

Questions ?