pS Traceroute LS

Ivan Garnizov — ivan.garnizov@fau.de

Antoine Delvaux — antoine.delvaux@man.poznan.pl

June 12th, 2020

Agenda

- Problem statement and Idea
- Metrics to build on
- Possible solution
- Data sources
- DB model proposal
- High level implementation plan
- Risks and Discussion!

Main idea: Topologically relevant LS

- Collect traceroute data from multiple sources
    - Partners
    - Projects
    - Community
    - Even external sources.. like RIPE Atlas in the future
- Provide a service, which presents the most relevant pS nodes along a requested path or even a single IP address
- Dynamically updated
    - Always up to date while we collect traceroute data
   - short lived records + benefits

Key elements

- (nearly|all)? pS nodes are running traceroute measurements
  - For PMP nodes
    - 5 different traces
    - about 5 IP addresses per trace
    - 25 IP addresses we have pS topology information for
  - More than 2.000 MP around the world
  - More than 50.000 IP addresses for which we can collect pS topology information
- results collected at pS MA
  - central MA
  - distributed MA (pS Toolkit)
- The collected information is current!
  - Search engines vs web directories of the 1990's

Metrics we can build on

- TTL distance
- RTT distance
- ASN, Organisation
- Timestamp
- Test URI (opt)

- ...and derivatives from above: range queries

How it could look like
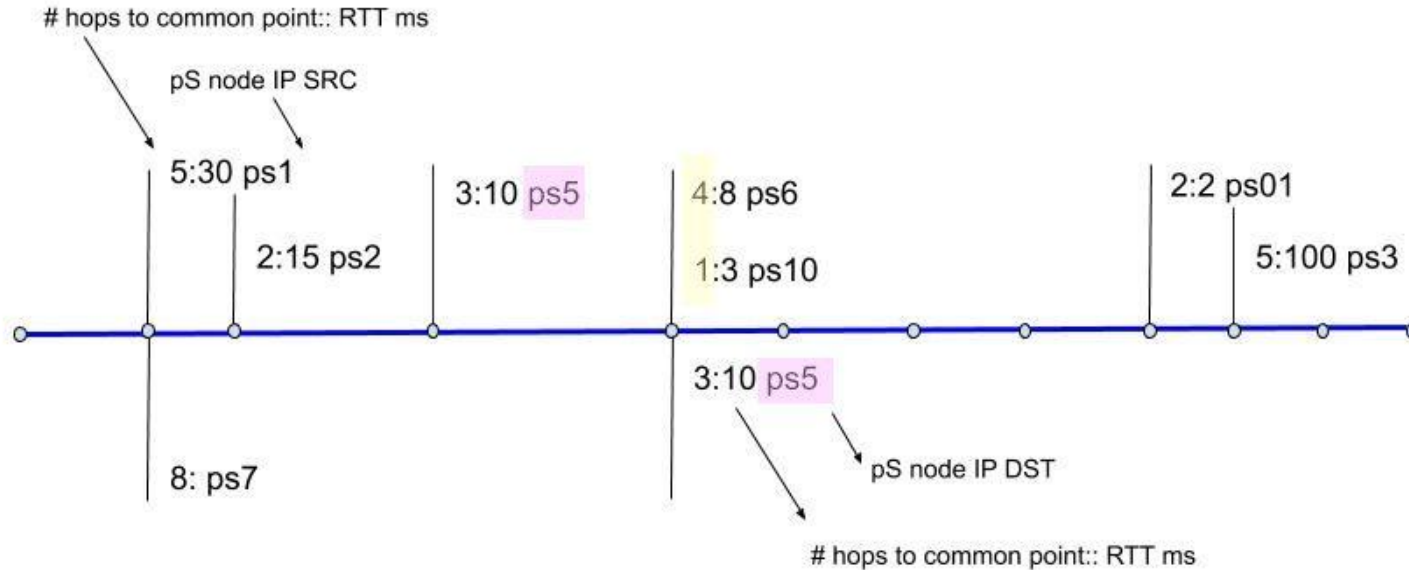
Data sources options

- Direct collection of results from pS nodes
  - As simple as adding a pScheduler archiver config
- Querying of central pS MA
  - When people advertise those MA and grant us access
- Triggering and querying of RIPE Atlas
  - perhaps even based on a cooperation
  - still IMO only based on registered pS nodes

Path matching

- Match sections between queried and collected path
- Requires collection and processing of full traceroute paths (simple on data collection)
- Can provide a limited insight for the connectivity
  - IF pS performance measurements are collected
- Does not fit well with external data sources: RIPE Atlas
- IMO can be covered 70-80% by an improved data preservation model

DB model proposal

- preserve pairs of pS node IP and router IP
- calculate for each pS node the TTL and RTT data
  - at best preserve the direction with positives and negatives
- preserve the rest of the info ASN, domain
- At best preserve URI of measurement
- LINK
- Review the graphic again

High Level Implementation Plan

- Have a deployment of the ELK stack
  - follow the examples from Andy and others
- configure 2 test pS instances with traceroute tests towards:
  - each other
  - external pS nodes
  - external non pS servers
- prepare a conversion
- prepare a query based on full traceroute data
- work on the results output format for API
- work on GUI - loong term

Engagement

- Use the opportunity to promote perfSONAR framework by inviting community to participate
  - Projects
  - Community
- Credits for the use?... when we grow

Summary and discussion

- technology: ELK
- Data collection process: parsed and/or injected
- Risks: "restricted" nodes in DB/results - mitigated, other?
- Extension with RIPE Atlas
- pS Traceroute LS +/in pS LS
- GUI vs API only
- DB model
- A name for this service?
- more