Improving Delay Estimation with Path Stitching

DK Lee, Keon Jang, Changhyun Lee, Sue Moon, Gianluca Iannaccone*

Division of Computer Science, KAIST
Intel Research, Berkeley*

Motivation behind the talk

Microsoft Update

Akamai
The Trusted Choice For Online Business™

XBOX 360

BitTorrent

Distributed
Internet applications

Real-time
query

Performance
measurements

DNS-like
Internet oracle

- *Use* measurements data which are already collected
- *Retrieve* relevant data from existing measurements

Problem statement

"Path stitching"

- *Estimates* Internet forwarding path and latency between two arbitrary points in the Internet
- *Uses* existing two kinds of measurements data
  - *Traceroute outputs* from Ark project
  - *BGP table snapshots* from RouteViews and RIPE RIS
- *Does not make* extra active measurements

Talk outline

- ***Path stitching***
  - End-to-end path and latency prediction method
- Preliminary results of estimation accuracy
- Discussion
  - Sources of error
  - Improving accuracy and AS-to-AS coverage
- Concluding remarks

:X: and X::Y

:X: Intra-domain paths of AS **X**

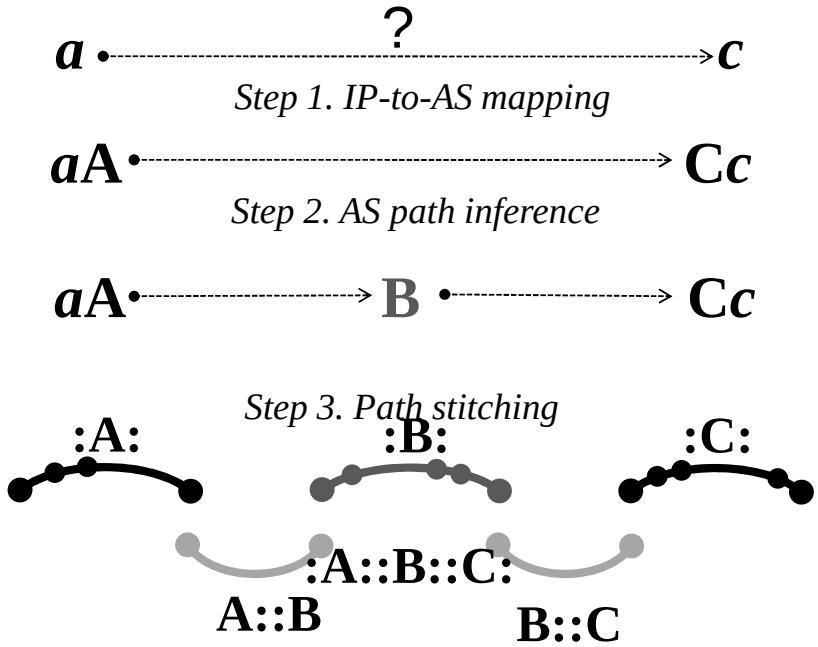
X::Y Inter-domain edges between AS **X** and **Y**

:X: + X::Y + :Y: = :X::Y:

Internet forwarding paths from AS **X** to **Y**

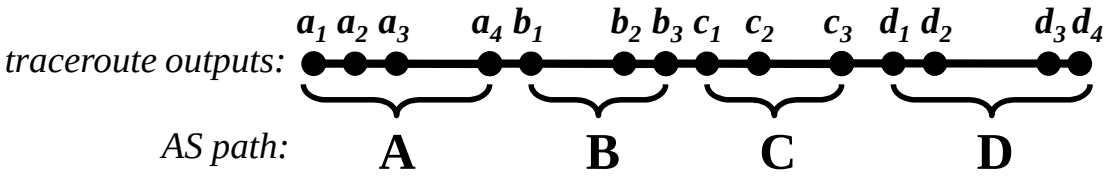
Path stitching - overview

Internet forwarding paths and end-to-end delay between ***a*** and ***c***



Index building

- In order to make a huge number of *traceroute* measurements searchable,

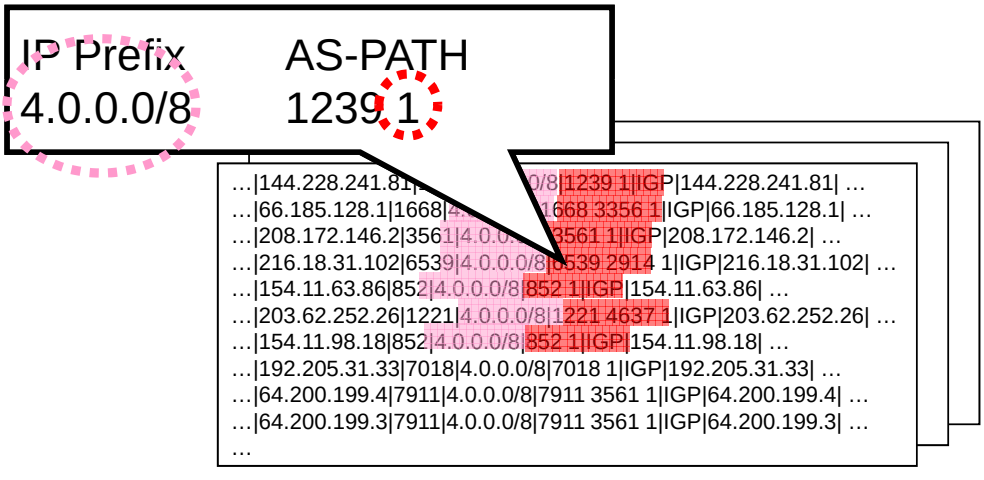


- Build indices for all possible partial paths:
  ABCD, ABC, BCD, AB, BC, CD, A, B, C, D
  Requires space in the order of $O(l^2)$

- Build only $O(l)$ indices:
  A, B, C, D, AB, BC, CD

(1) IP-to-AS mapping

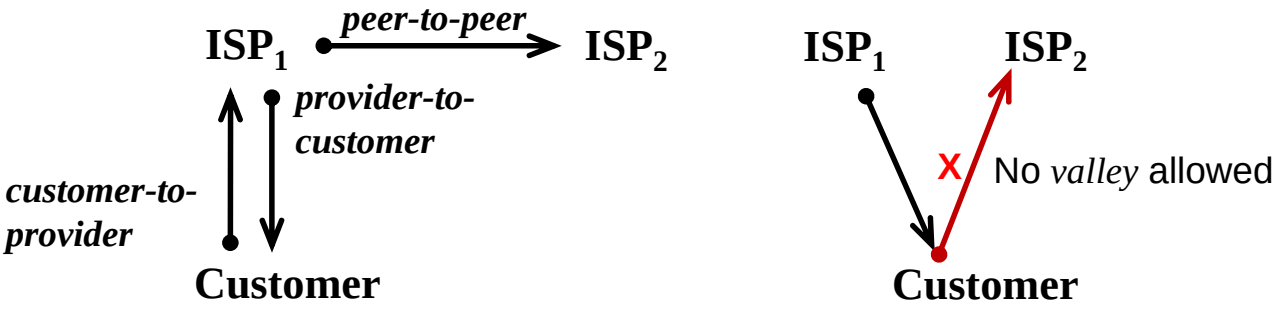
- Step 1. IP-to-AS mapping
  Use BGP Routing table snapshots from:
  University of Oregon, RouteViews' BGP listener
  RIPE RIS' 14 monitoring points (rrc00 ~ rrc07, rrc10 ~ rrc15)



BGP Routing table snapshots.

(2) AS path inference

- Relationship between a pair of ASes



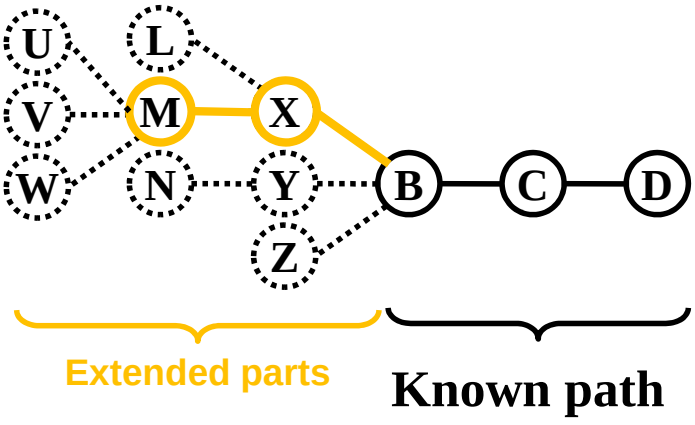
- Valley-free property [L.Gao, TON'00]
  - After traversing a *provider-to-customer* or *peer-to-peer edge*, the AS path can not traverse a *customer-to-provider* or *peer-to-peer* edge.

(2) AS path inference

- Step 2. AS path inference

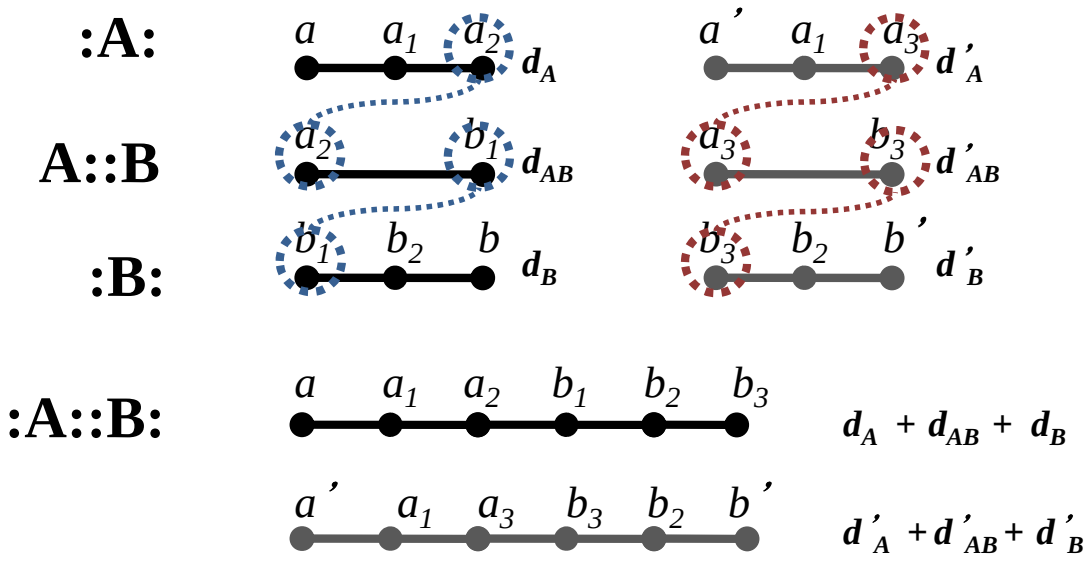
  Qiu and Gao's methodology [GLOBECOM'06]

  - Exploits the AS paths, ***known paths***, appeared in BGP routing tables.
  - Infers AS paths that satisfying ***valley-free property***.



(3) Stitching path segments

- Step 3. Infer path and delay
  along the inferred AS path



Data set

- For end-to-end paths and delay estimation:
  - CAIDA Ark's IPv4 Routed /24 Topology data set
  - Traceroutes from 22 monitors to every /24 routable prefix

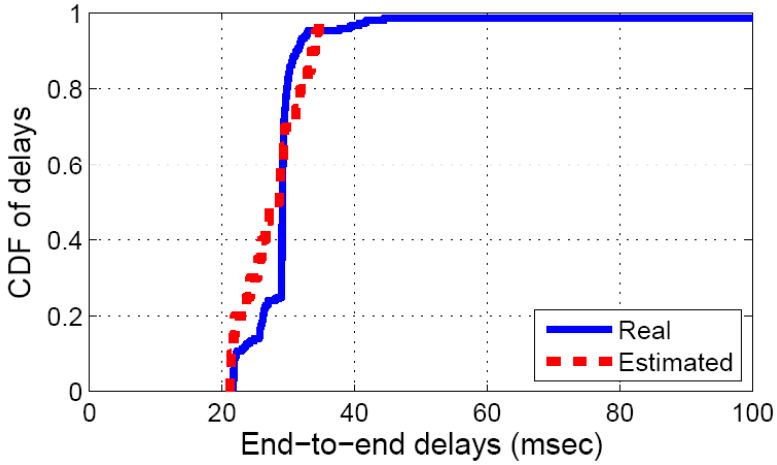
- For accuracy comparison:
  - Perform traceroute 50 times a day between 184 PlanetLab nodes (real live measurements)
  - For every pair of PL-nodes, estimate path and delays using path stitching, and compare the results

Accuracy of estimations – (1)

planetlab2.xeno.cl.***cam.ac.uk*** → lsirextpc02.***epfl.ch***

(cbg-uk) ***(zrh-ch)***

20 stitched paths vs. 150 real measurements



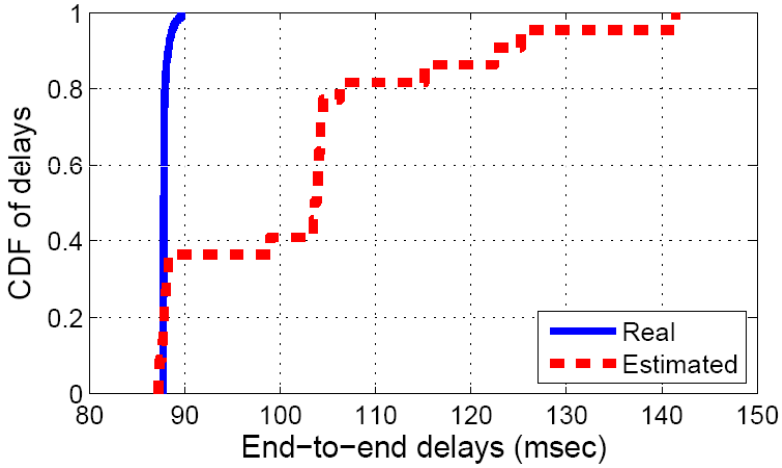
Estimated delays align with real measurements very closely.

Accuracy of estimations – (2)

planetlab1.csail.***mit.edu*** → planet2.scs.***stanford.edu***

(No Ark monitors)

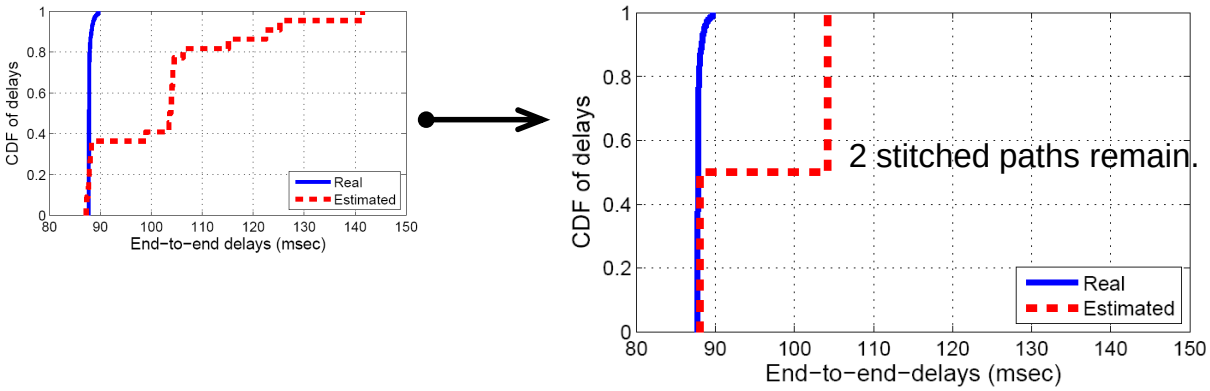
22 stitched paths vs. 150 real measurements



36 % of estimated delays match the real measurements.

Discussion - sources of error

- Dynamic nature of routing in the Internet
- IP-to-AS mapping
- AS path inference
- Not all stitched paths produce good estimates.
  - Tie breaking based on the *destination prefix*



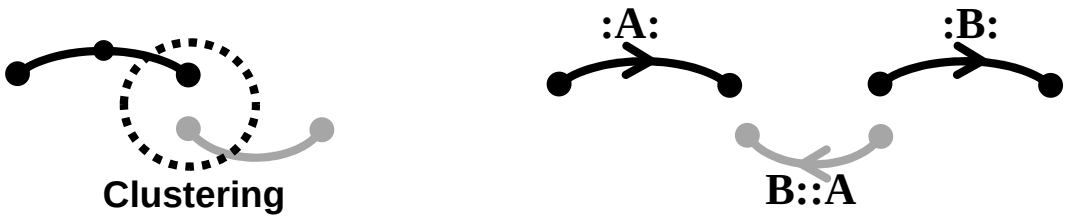
After applying the tie-breaking rule

Discussion - improving coverage

- What if we do not have any stitched path?

  4,091 out of *17,879* PlanetLab pairs are stitched.

  Stitch up paths that do not match at the ending nodes.



  Clustering two ending points:
  - Router-level and PoP-level clustering
  - Clustering based on IP prefix proximity

  Using reverse inter-domain path segments

Concluding remarks

- Path stitching

  End-to-end path and latency prediction method

  using existing measurements data
  without incurring extra active measurements

Thank you!

- Any question?

- For more question:
  dklee@an.kaist.ac.kr