SIBILLA: A Step Towards a Planet-Scale Measurements Retrieval Infrastructure

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

FISC Student Workshop
Aug 28, 2008

Division of Computer Science, KAIST
Intel Research, Berkeley*

Motivation behind the talk

- Distributed applications are popular in today's Internet.
  - Peer-to-peer file sharing, CDNs, multi-player online games
- They can benefit from information about the Internet path and proximity between their nodes.
  1. Nearest neighbor discovery
  2. Leader node selection
  3. Distribution tree construction
- Our goal is a DNS-like system that provides network-internal performance information

Key idea behind Path Stitching

- Internet separates ***inter-*** and ***intra-domain*** routing
  - » Path stitching ***splits paths*** into path segments, and ***stitches path segments*** together using BGP routing information to ***predict a new path***

- Many measurement data are available already, and we use them and ***do no*** additional measurement

Talk outline

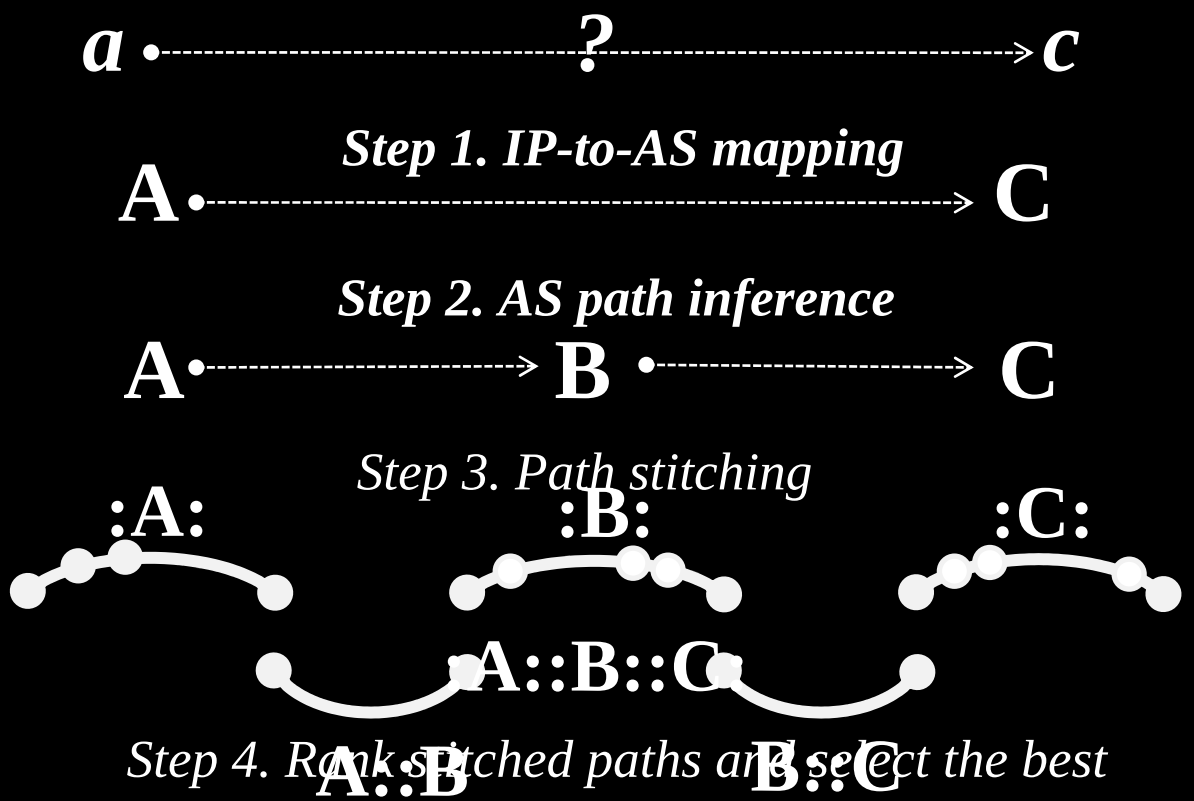
- ***Path Stitching*** algorithm
- When Path Stitching produces no stitched path
  - Approximation heuristics
- When Path Stitching produces multiple paths
  - Preference rules
- Implementation and deployment issues
- Concluding remarks

Data set

- *CAIDA Ark's* ***traceroutes***
  - One round of *traceroute* outputs from 18 sources to every /24 prefix
  - 14 millions of *traceroute* outputs
- BGP routing tables
  - University of Oregon, ***RouteViews***' BGP listener
  - *RIPE RIS*' 14 monitoring points (rrc00 ~ rrc07, rrc10 ~ rrc15)
- Notations
  - **: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**

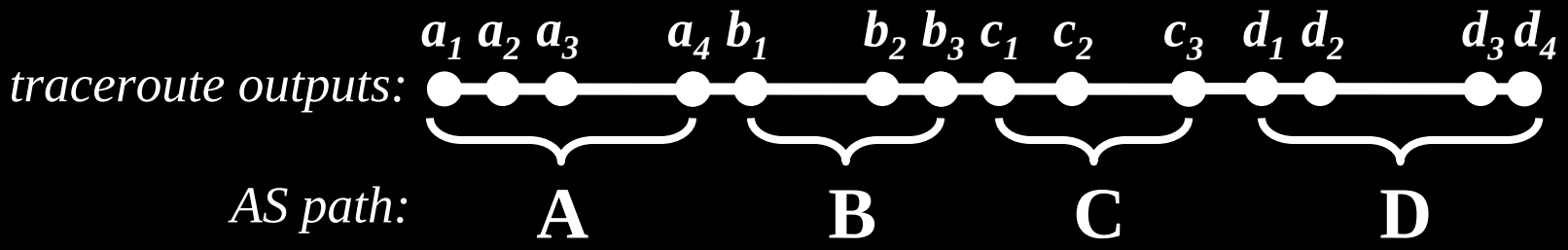
Overview of Path Stitching

- What are Internet forwarding paths and end-to-end delay between two arbitrary Internet host ***a*** and ***c?***



Index building

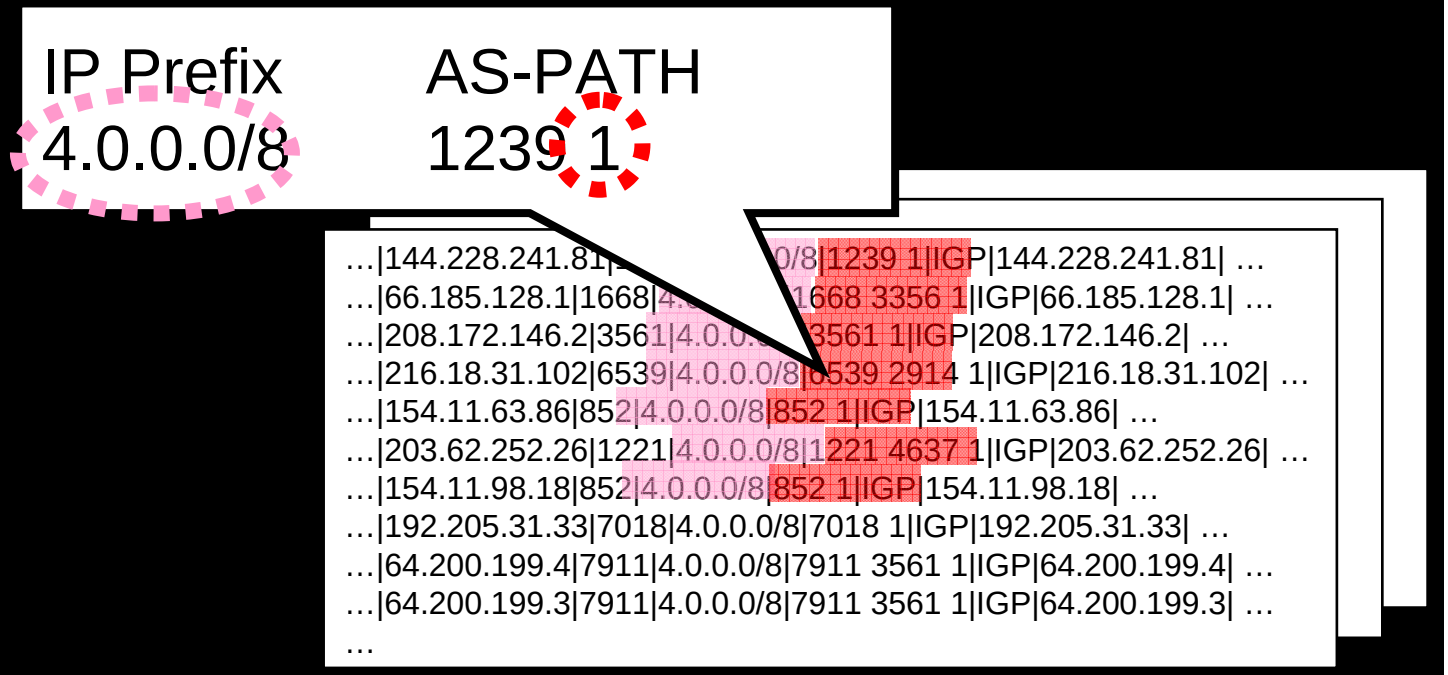
- In order to make a huge number of *traceroute* measurements *searchable*,



- Choices
  - Build indices for all possible partial paths
    - **ABCD, ABC, BCD, AB, BC, CD, CD, A, B, C, D**
    - Requires O(l^2) space
  - Build indices for intra AS and inter AS segments
    - **A, B, C, D, AB, BC, CD**
    - Requires O(l) space

Step 1. IP to AS mapping

- Use BGP routing table snapshots:
  - An IP address is mapped to the *longest matching IP prefix* in a table,
  - Take the *last hop in the AS-PATH* as the origin AS



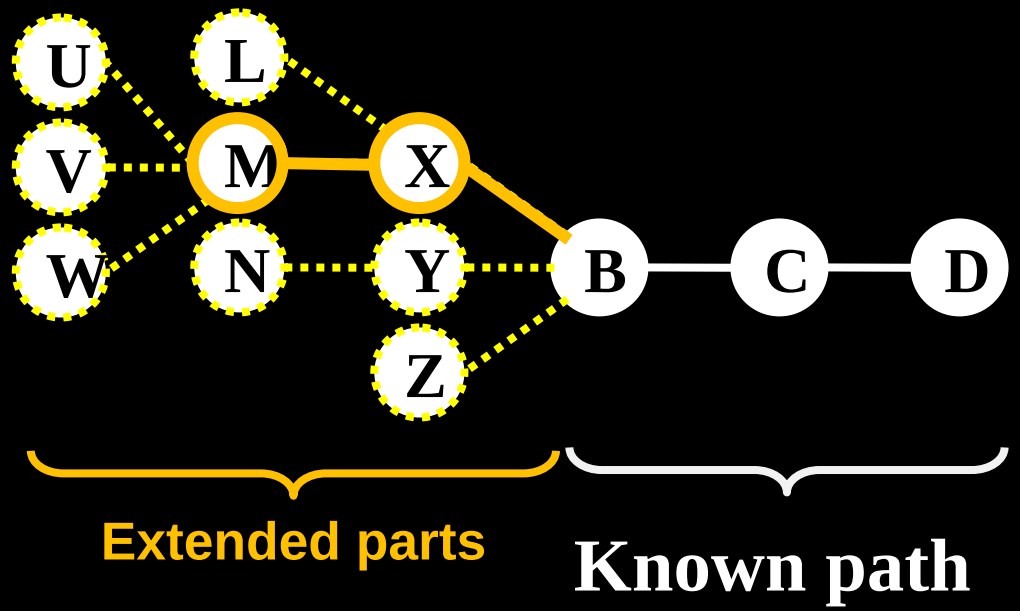
BGP Routing table snapshots.

Errors in IP to AS mapping

- Single origin AS mismatch
  - Mao et al reported that inaccurate mapping result in
    - Missing AS hop, extra AS hop, substitute AS hop, two hop AS loops
  - 8.9% AS paths contain two-hop AS loops
  - If we use the same IP-to-AS mapping for a query, the outcome would be consistent although mismatched.
- Multiple origin AS (MOAS)
  - 2,651,387 out of 14M traceroutes have MOAS conflicts
  - 22.61% of MOAS are caused by Internet exchange prefixes
  - Infer AS paths from all MOASes

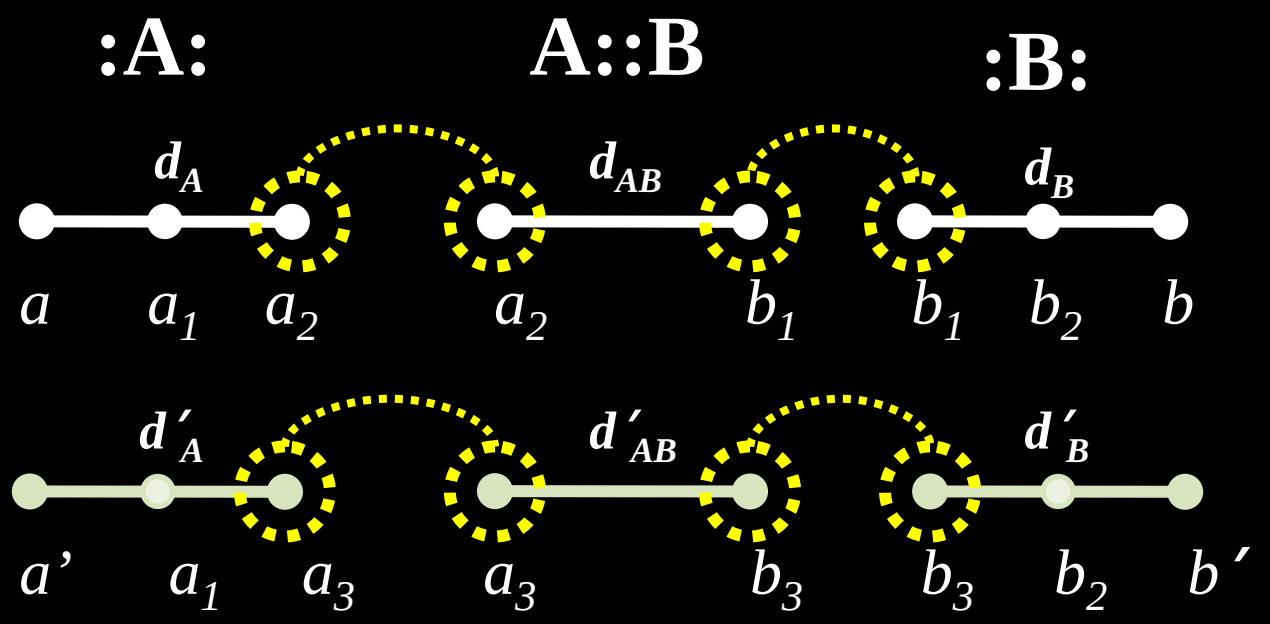
Step 2. AS path inference

- Qiu and Gao's methodology [GLOBECOM'06]
  - Exploits the AS paths, *known paths*, appeared in BGP routing tables.
  - Infer AS paths that satisfying *valley-free property* [L.Gao, TON'00]

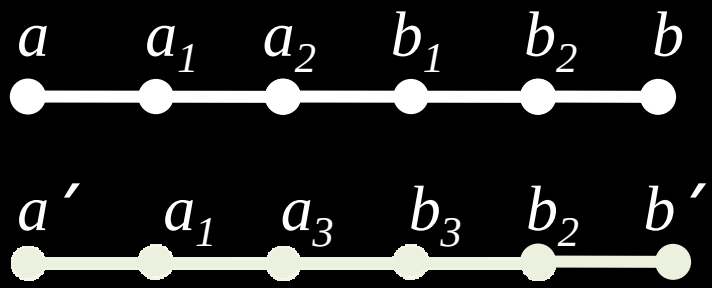


Choose shortest path with low ***unsure length*** and high ***frequency index***
Accuracy of 60% reported

Step 3. Stitching path segments



:A::B:

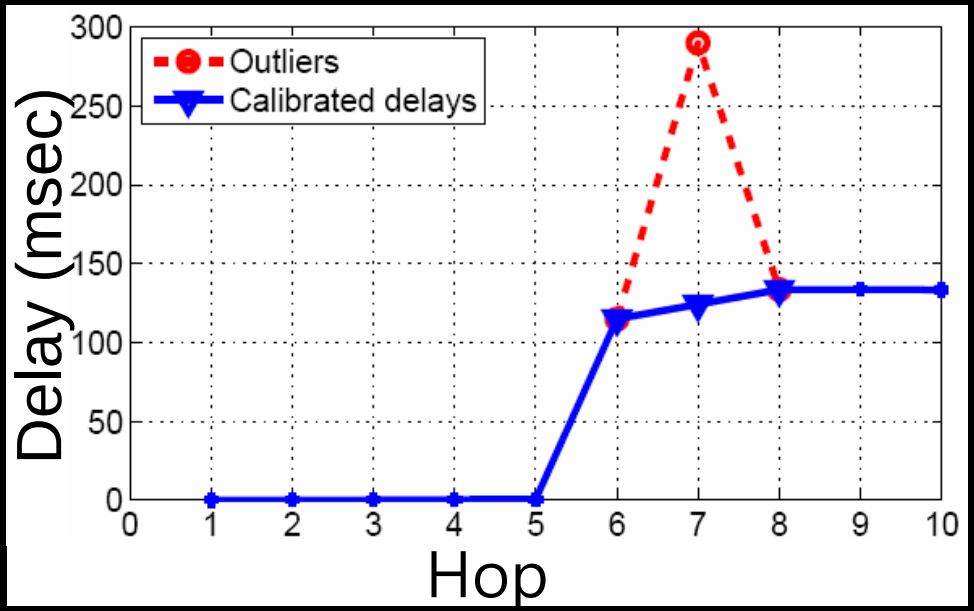


$d_A + d_{AB} + d_B$

$d'_A + d'_{AB} + d'_B$

Sources of error – *traceroute*

- Dynamic nature of the Internet
  - » Record all reported measurement per path segment.
  - » Report the most recent or median of the past known history.
- Non-decreasing delay principle



When Path Stitching produces no stitched path

Case #1: No path segments in source/destination AS

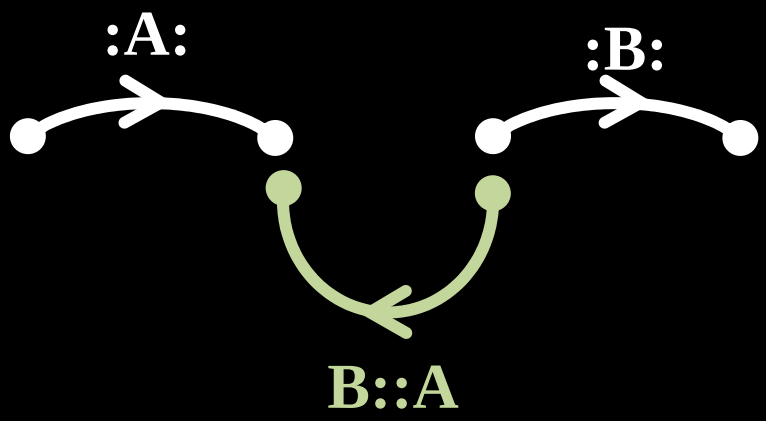
- The source or the destination is not in the same AS with any measurement data

Data type	*Total AS*	*Transit AS*	*Stub AS*
Ark	14,378	4,418	***9,960***
BGP	28,244	4,847	***23,397***

- For 90% of undiscovered AS in Ark, the *traceroute* did not reach to AS
- ASes not covered by Ark accounts for only 110M or 5.8% of IP addresses in BGP

Case #2: No segments in the middle of inferred AS path

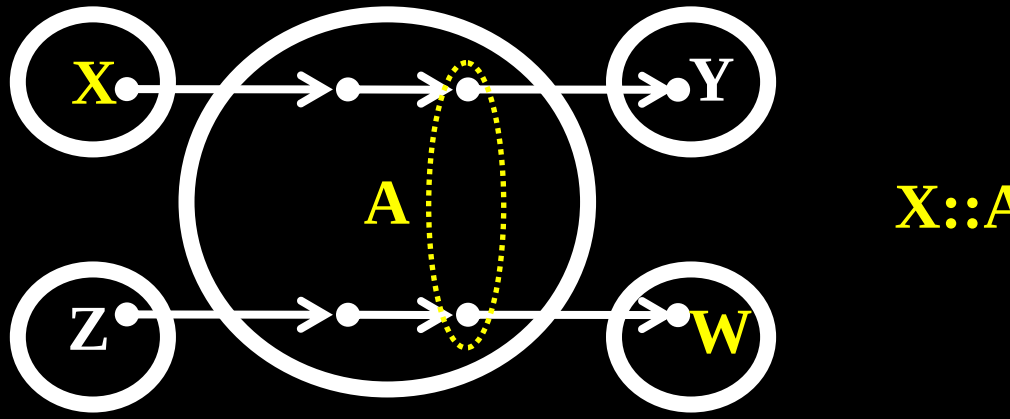
- No inter-domain path segment
  - Incorporating the reverse inter-domain segments



- No intra-domain path segment
  - No solution yet

Case #3: Segments does not rendezvous at the same address

For all ASes along the path has segments, but they do not rendezvous at the same address

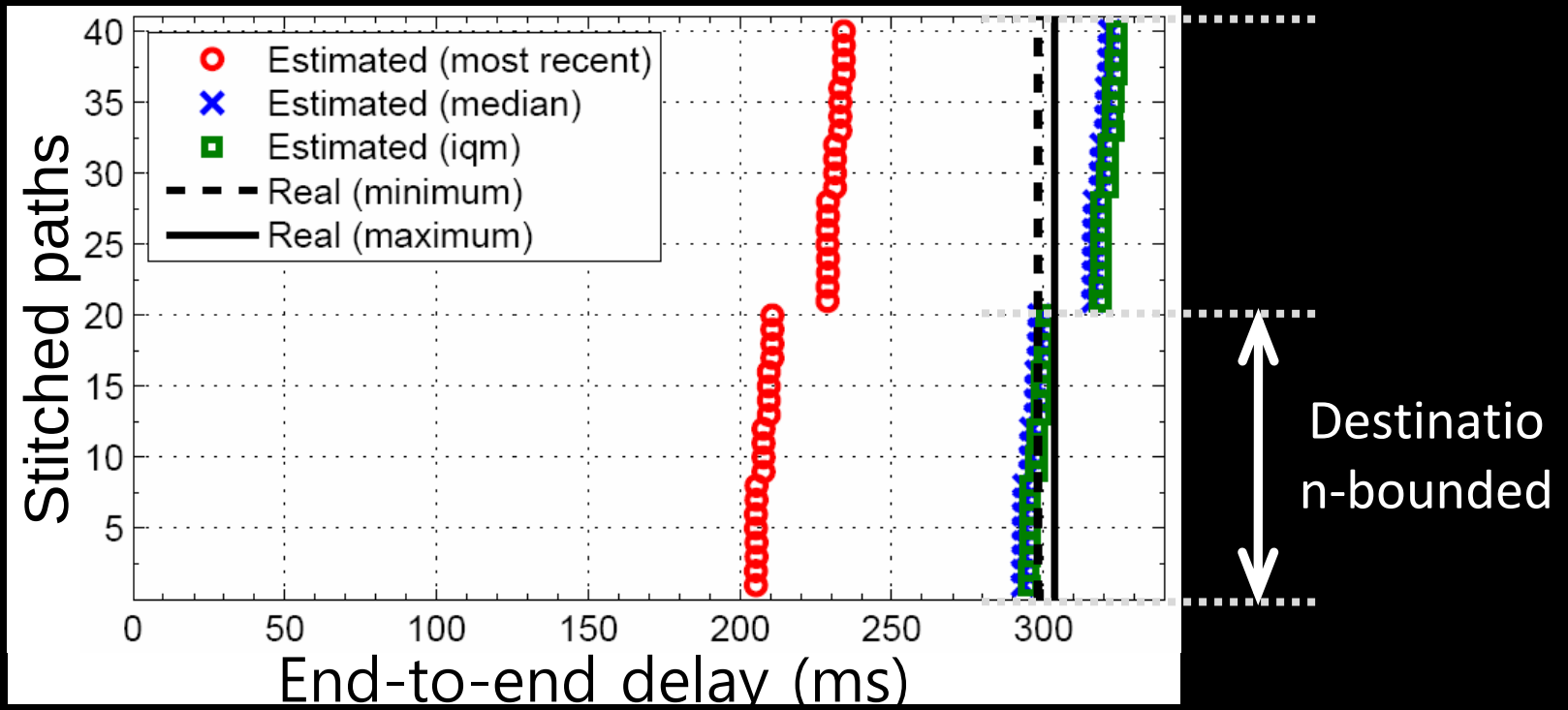


- Clustering heuristics:
  - Clustering IP addresses of ***the same router***
  - Clustering IP addresses ***in a single Point-of-presence (PoP)***
  - Clustering two ending points based on their ***IP prefix proximity***

When Path Stitching produces multiple stitched paths

Same destination-bound preference

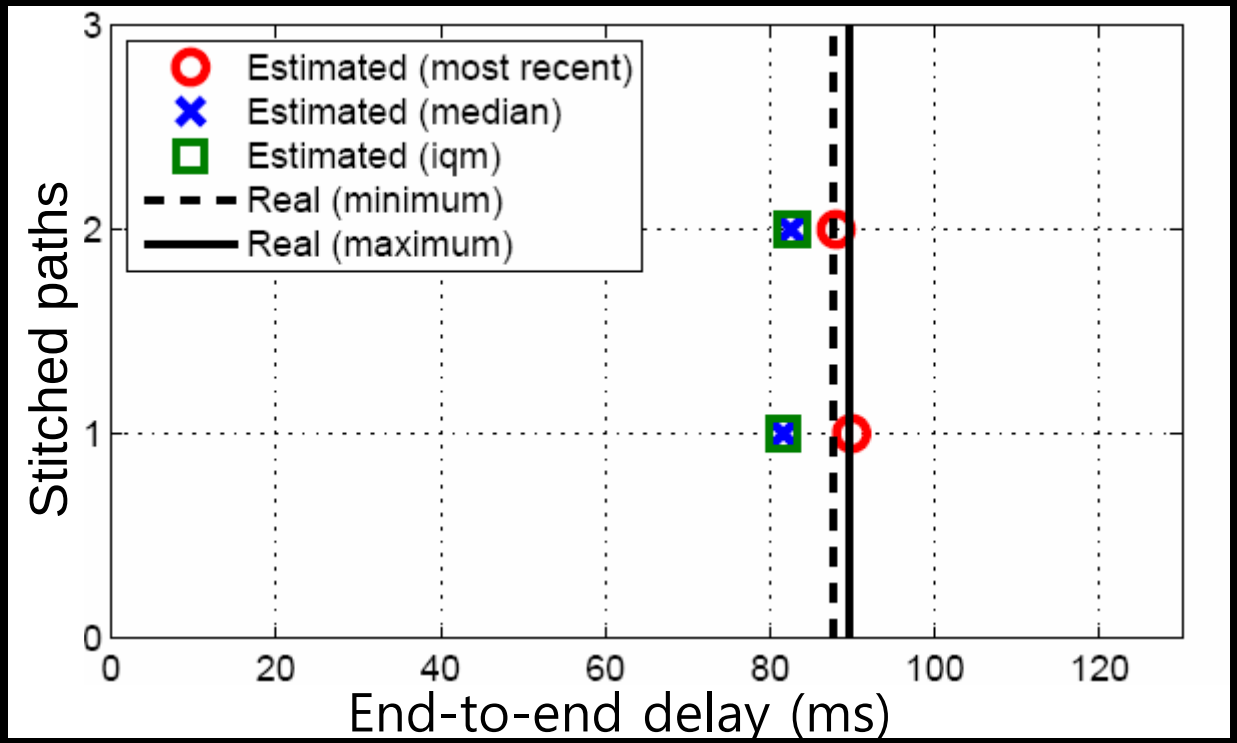
- planetlab2.xeno.cl.***cam.ac.uk***
  → pl1-higashi.ics.es.***osaka-u.ac.jp***



» Preference ***to the same destination-bound*** path segments

Closeness to source and destination

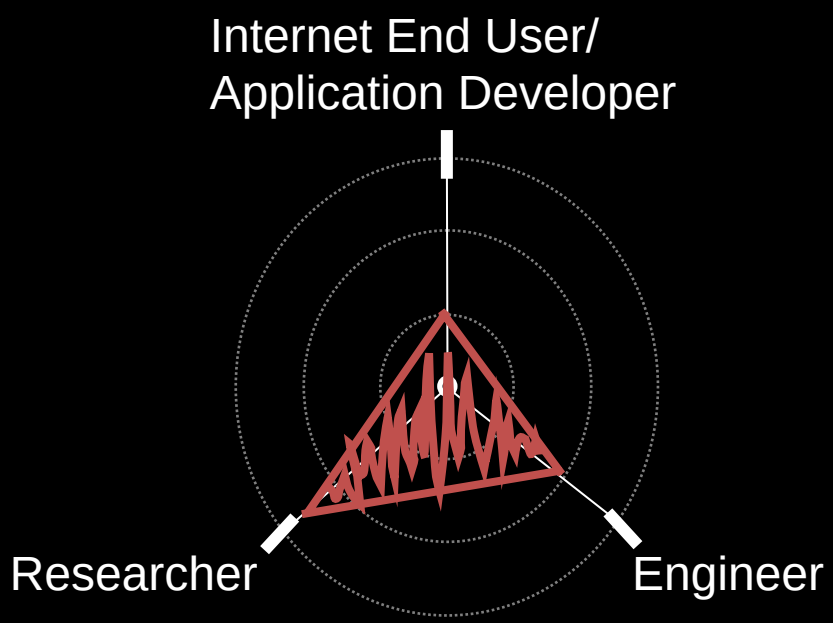
- Planetlab2.csil.***mit.edu***
  → planet2.scs.***stanford.edu***

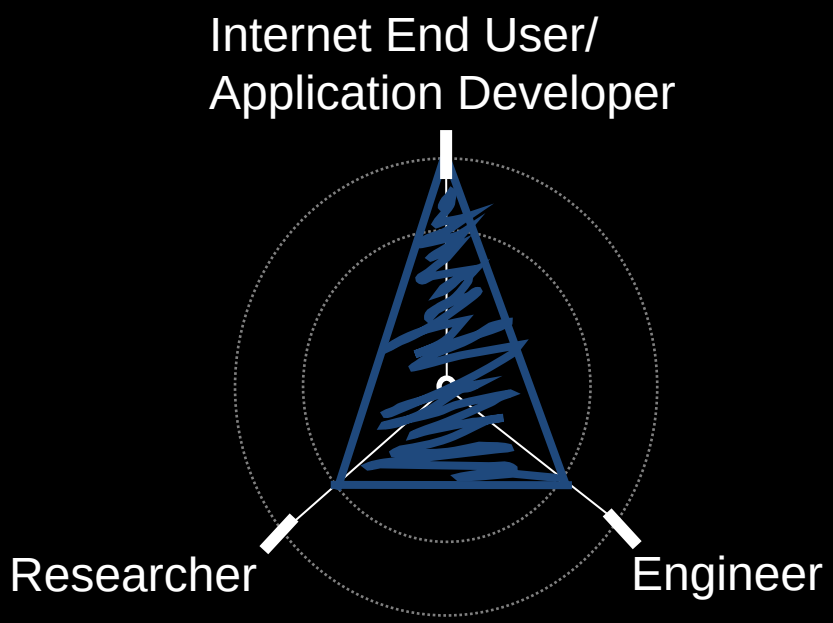


- In 20 % of Ases, delay difference within an AS is > 100 ms.
  » Preference ***to the closest points*** in source and destination ASes

Build and deploy DNS-like System in the real world

For Whom the Technology's Exploited





Previous works:
NetQuest, Vivaldi, IPlane

Our approaches:
SIBÍLLA:

Requirements

- Planet-scale coverage
- Provides historical/ recent/ real-time data
- High accuracy
- Real-time query processing
- Programmability

Data gathering method

Uniform data representation

Measurement retrieval

Programmable interface

SIBÍLLA: Unified Internet Looking Glass

System-relate Requirements

- Supports distributed query processing
- Quality control
- Security
- Political issues
- Deployment

- ***Drafting behind the domain name system (DNS)***
  - Distributed database system which's been deployed all over the world.
  - ***TTL***-based caching mechanism
  - ***gethostbyname()***

Conclusions

- ***Path stitching***
  - Internet forwarding path and latency prediction by combining *traceroutes* and *BGP data*
- Our approach *uses* existing measurement data and *do no* additional active measurements
- Evaluation results are preliminary, but promising

Thank you!

- Any question?
- For more question:

  dklee@an.kaist.ac.kr