iPlane: An Information Plane for Distributed Services

Harsha V. Madhyastha, Tomas Isdal, Michael Piatek
Colin Dixon,Thomas Anderson, Arvind Krishnamurthy
University of Washington
Arun Venkataramani
University of Massachusetts Amherst

Motivating Example: BitTorrent

- Default BitTorrent
  - Client contacts central tracker
  - Tracker returns random subset of peers

- Tracker needs to predict performance between peers
  - Can return peers that will provide best performance to the client

Related Work

- Simple and elegant approaches exist for estimating latency
  - Embed all end-hosts into a Euclidean space (e.g., GNP, Vivaldi)
  - Not extensible to other metrics (e.g., bandwidth)

- We pursue a structural approach
  - Use measurements of the Internet's structure to predict end-to-end route
  - Compose properties of links on predicted route

Our Work: iPlane

- Design and implement *iPlane*
  - System that predicts path properties on the Internet between arbitrary end-hosts
  - Predict multiple metrics along unmeasured paths
- Demonstrate utility of *iPlane*
  - Accurate enough to be useful for distributed services

Challenges in building iPlane

- How do we …
  - build a structured atlas of the Internet?
  - predict routing between arbitrary end-hosts?
  - measure properties of links in the core?
  - measure links at the edge?

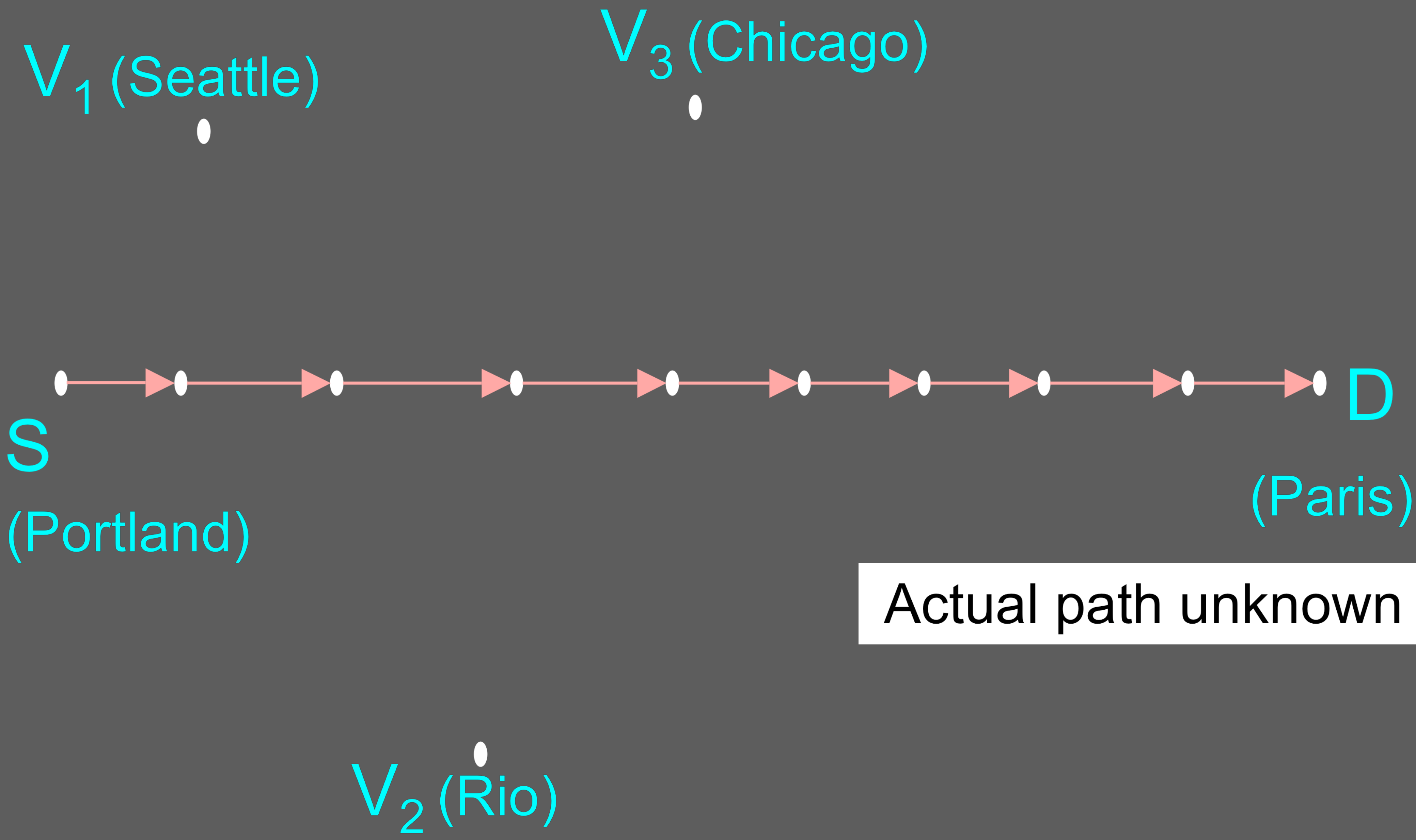
Build a Structural Atlas of the Internet

- Use PlanetLab + public traceroute servers
  - Over 700 geographically distributed vantage points
- Build an atlas of Internet routes
  - Perform traceroutes to a random sample of BGP prefixes
  - Cluster interfaces into PoPs
  - Repeat daily from vantage points

Challenges in building iPlane

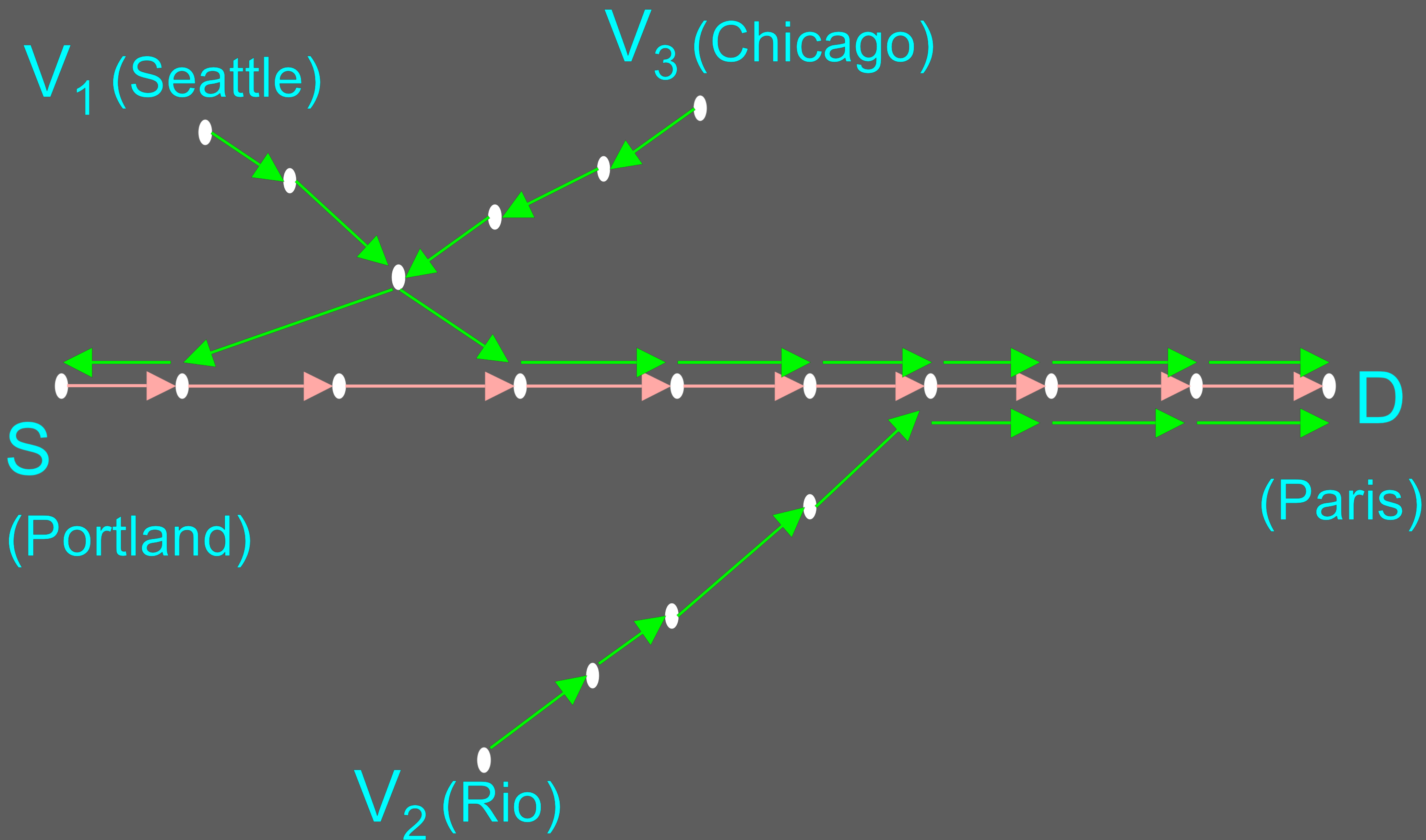
- How do we …
  - ✓ build a structured atlas of the Internet?
  - predict routing between arbitrary end-hosts?
  - measure properties of links in the Internet?
  - measure links at the edge?

Model for Path Prediction

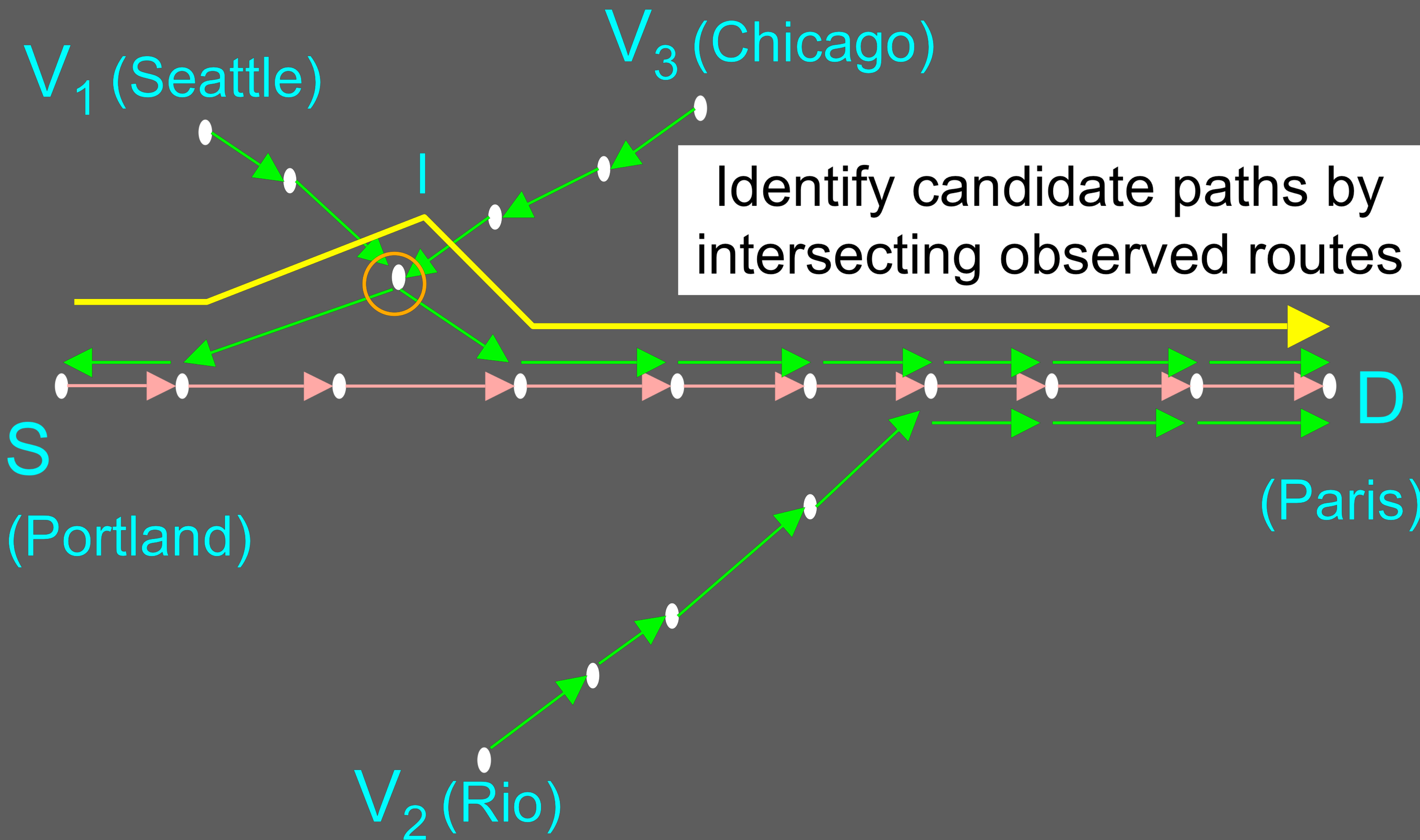


Actual path unknown

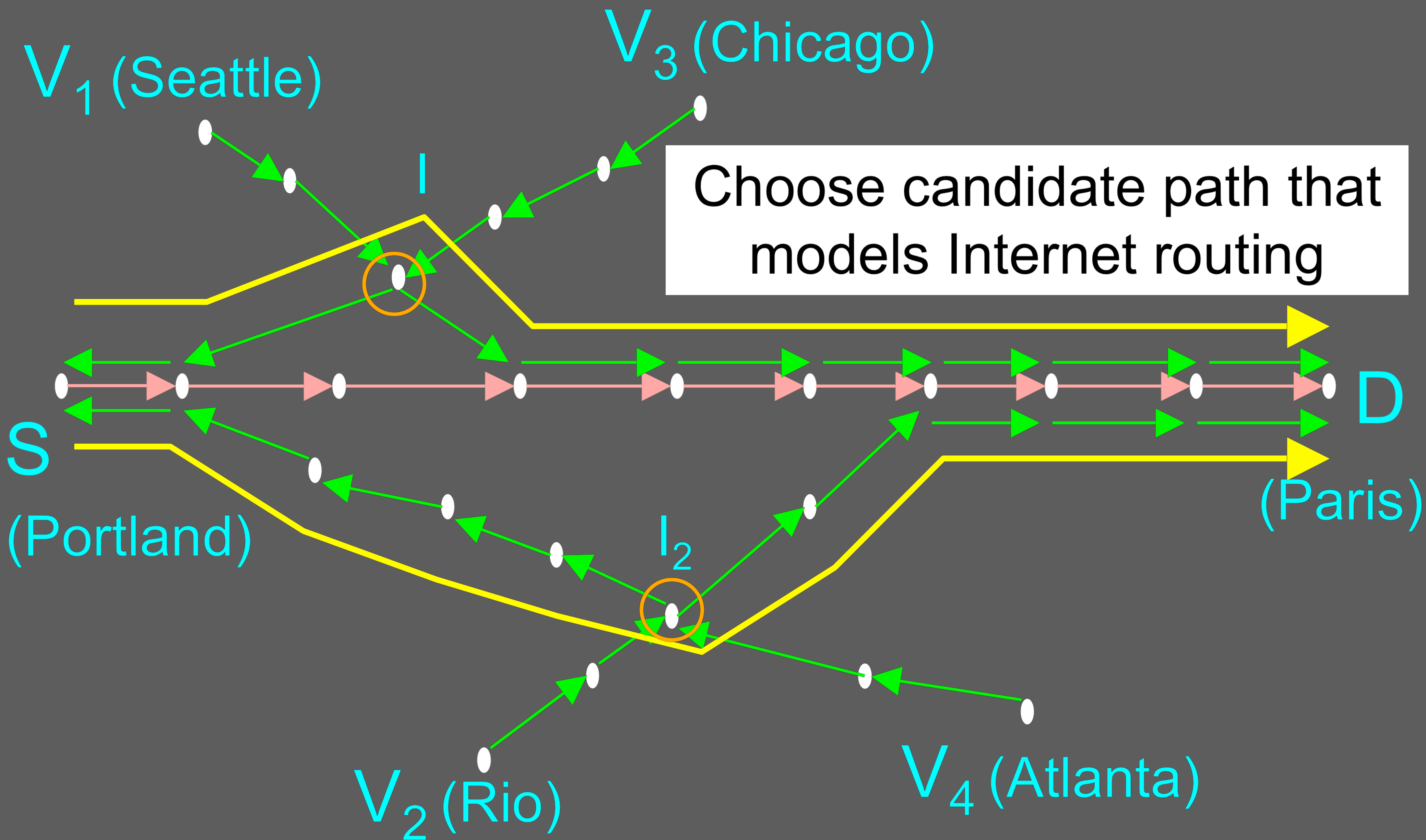
Model for Path Prediction



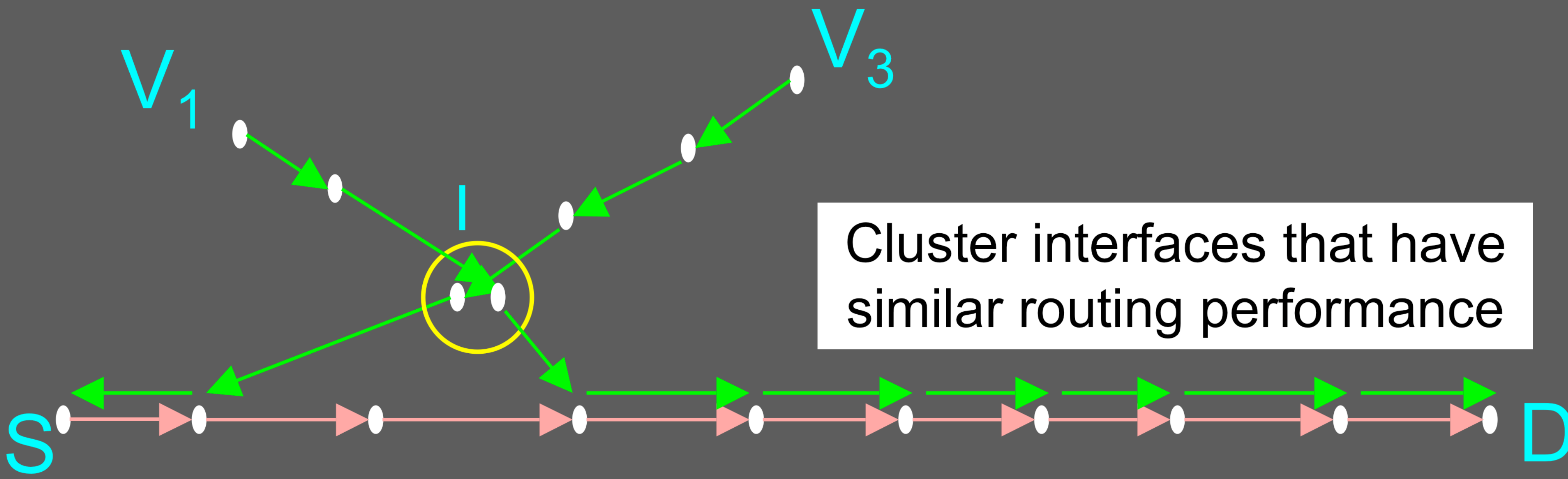
Model for Path Prediction



Model for Path Prediction



Can Miss Intersections

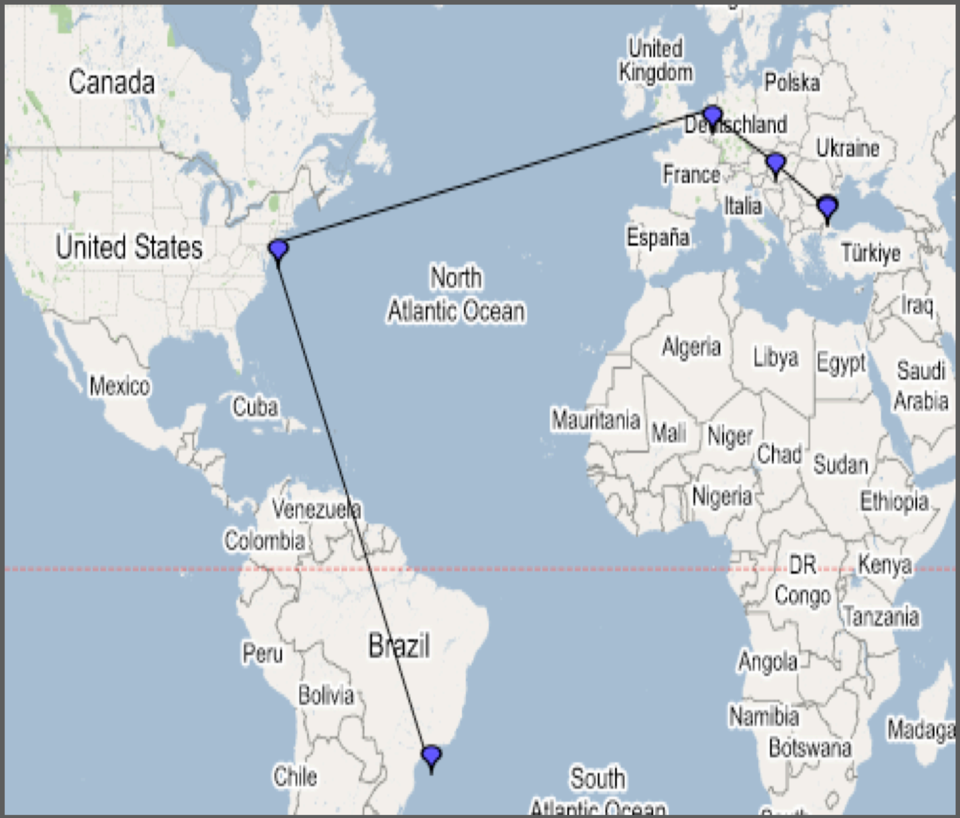


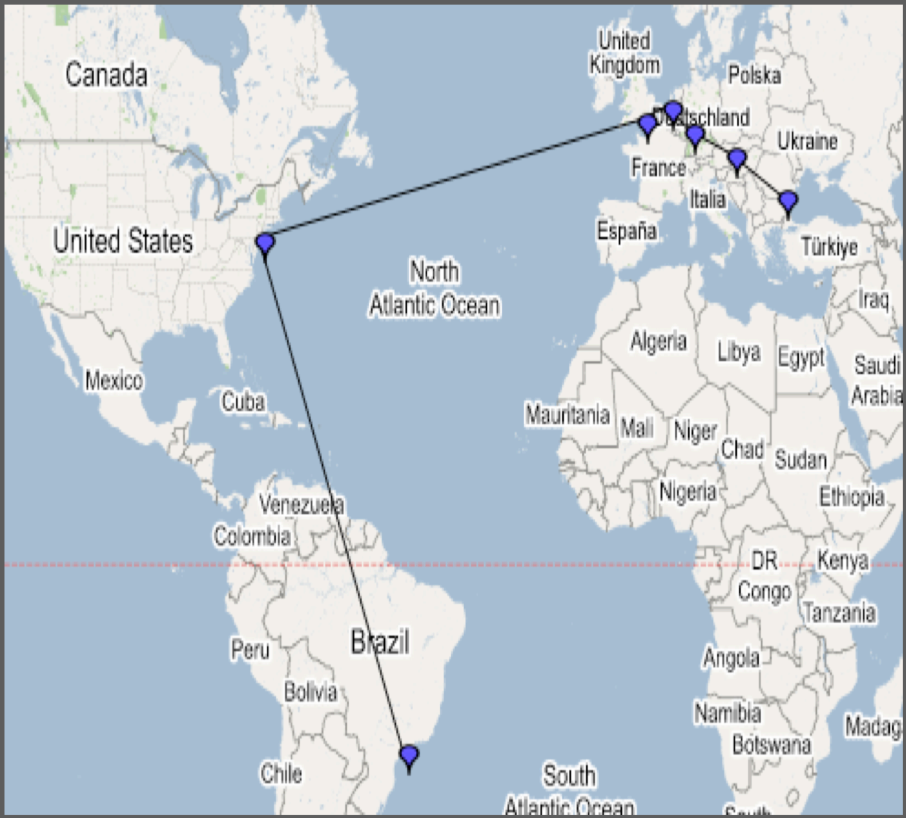
- Helps in reuse of measurements without loss of accuracy
- Fewer links to be measured

Cluster Interfaces into PoPs

- Interfaces on the same router use the same routing table
- Routers at the same location within an AS will have similar routing tables
  - Discover locations based on DNS names
    - Invalidate inferred locations if incorrect
  - Discover co-located interfaces
    - Nearby interfaces have similar reverse paths back to each vantage point

Example of Path Prediction

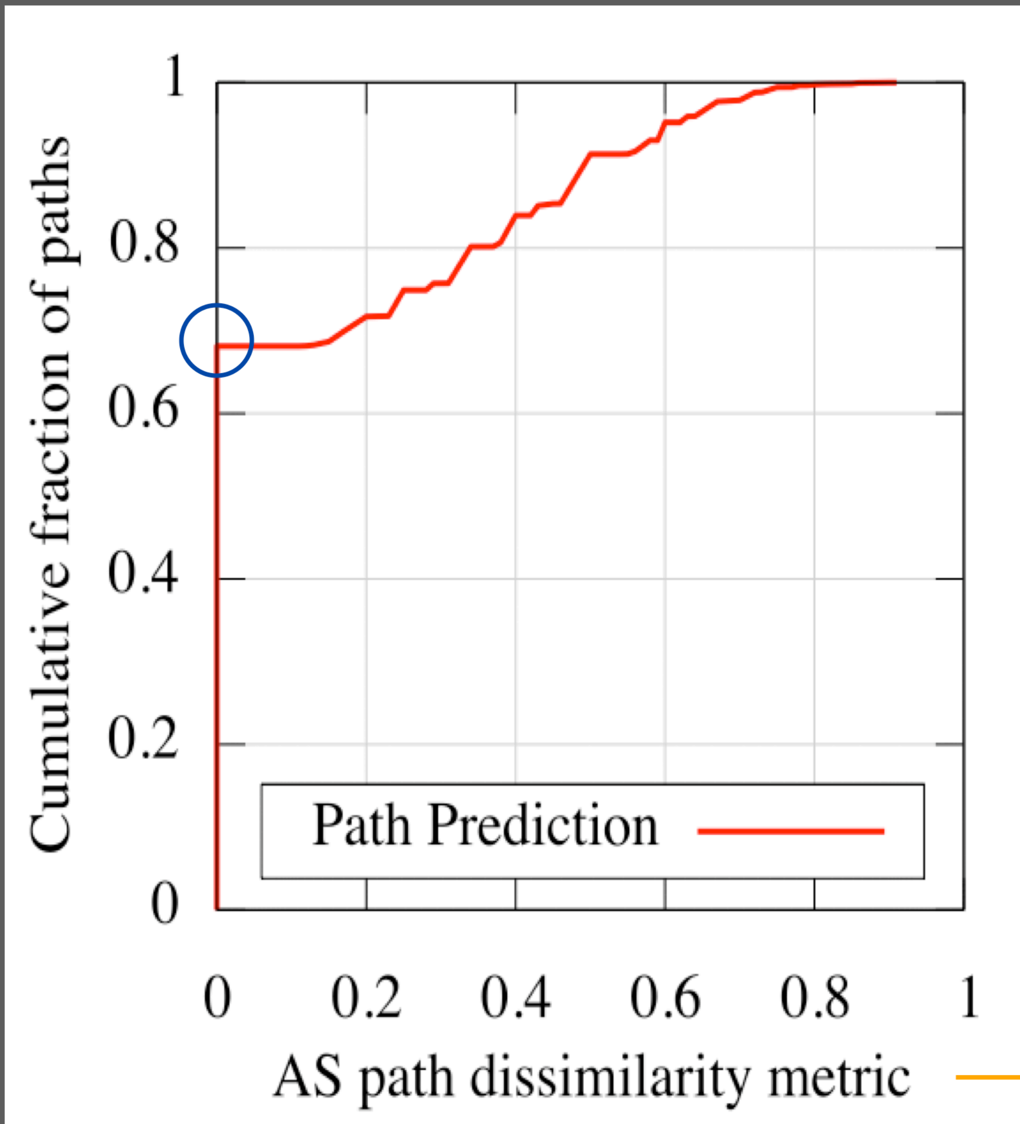




Actual path: RTT 298ms

Predicted path: RTT 310ms

Does Path Prediction work?



- Used atlas measured from PlanetLab to predict paths from public traceroute servers
- 68% of path predictions are perfect

$$1 - \frac{|\text{Intersection of ASes}|}{|\text{Union of ASes}|}$$

Predicting Path Properties

- To estimate end-to-end path properties between arbitrary *S* and *D*
  - Use measured atlas to predict route
  - Combine properties of
    - Links in the core along predicted route
    - Access links at either end

Latency	Sum of link latencies
Loss-rate	Product of link loss-rates
Bandwidth	Minimum of link bandwidths

Challenges in building iPlane

- How do we …
  - ✓ build a structured atlas of the Internet?
  - ✓ predict routing between arbitrary end-hosts?
  - – measure properties of links in the core?
  - – measure links at the edge?

Measuring Links in the Core

- Only need to measure inter-cluster links
- Objectives
  - Probe each link mostly once
  - Distribute probing load evenly across vantage points
  - Probe each link from closest vantage point

- Frontier Search algorithm selects paths that cover all links
  - Parallelized BFS across PlanetLab nodes
- To span atlas measured from 200 PlanetLab sites
  - Each node has to measure around 700 links

Challenges in building iPlane

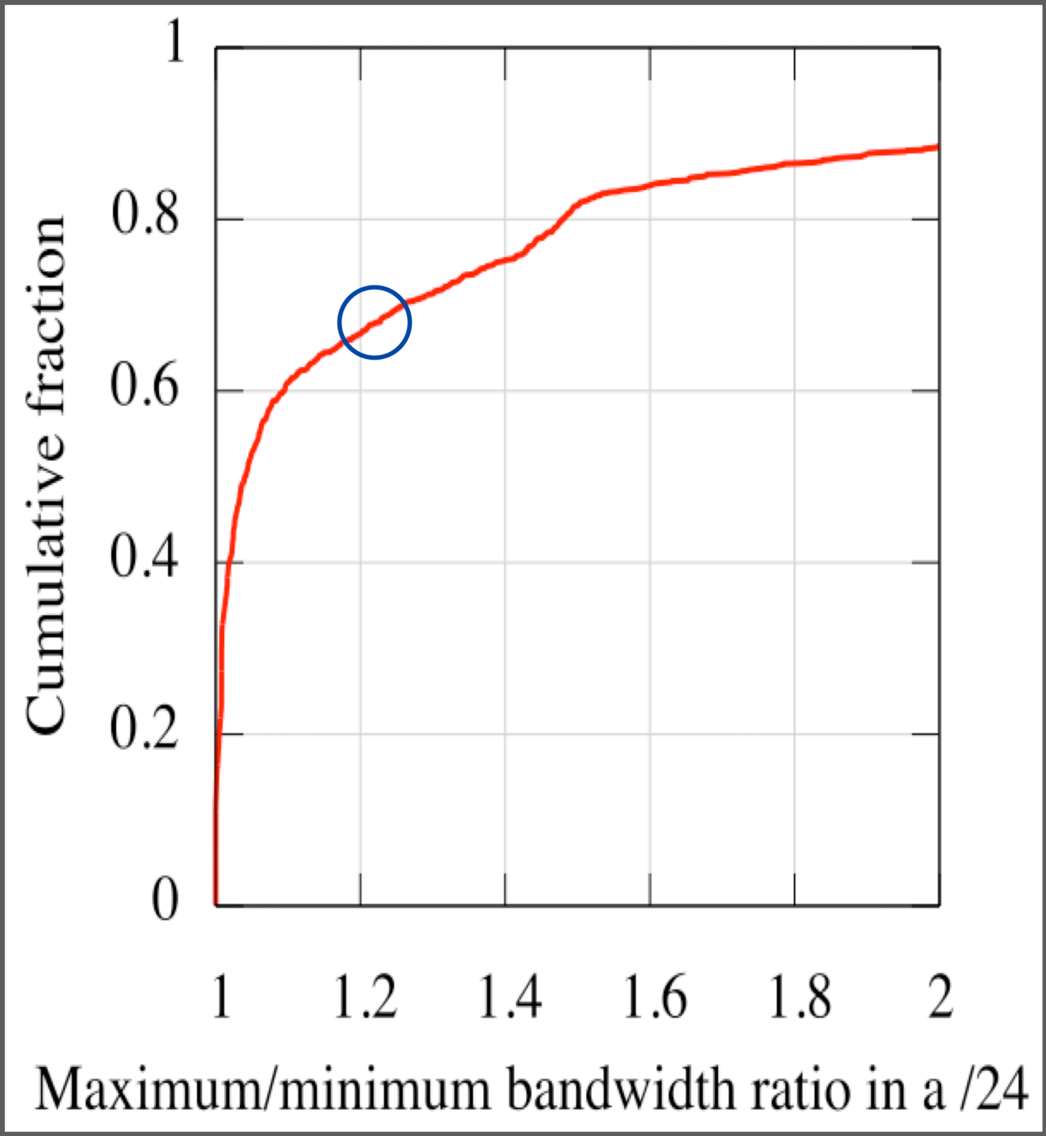
- How do we …
  - ✓ build a structured atlas of the Internet?
  - ✓ predict routing between arbitrary end-hosts?
  - ✓ measure properties of links in the core?
  - – measure links at the edge?

Measuring the Edge

- Participate in BitTorrent swarms
  - Popular application: wide coverage of end-hosts

- Passively monitor TCP connections to measure access link properties
  - Will not raise alarms

Reusability of Measurements



- Measurements to multiple addresses in the same /24 within 20% of each other in 66% of cases
- Reuse bandwidth measurements within a /24 prefix

Finally done building iPlane!

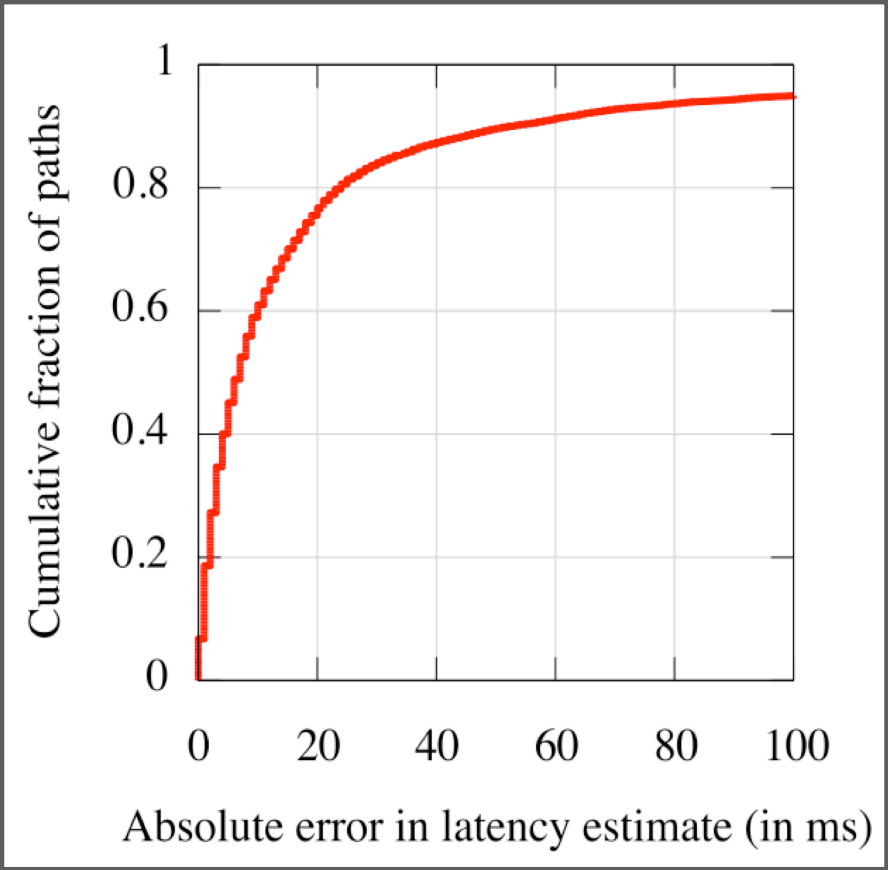
- How do we …
  - ✓ build a structured atlas of the Internet?
  - ✓ predict routing between arbitrary end-hosts?
  - ✓ measure properties of links in the core?
  - ✓ measure links at the edge?

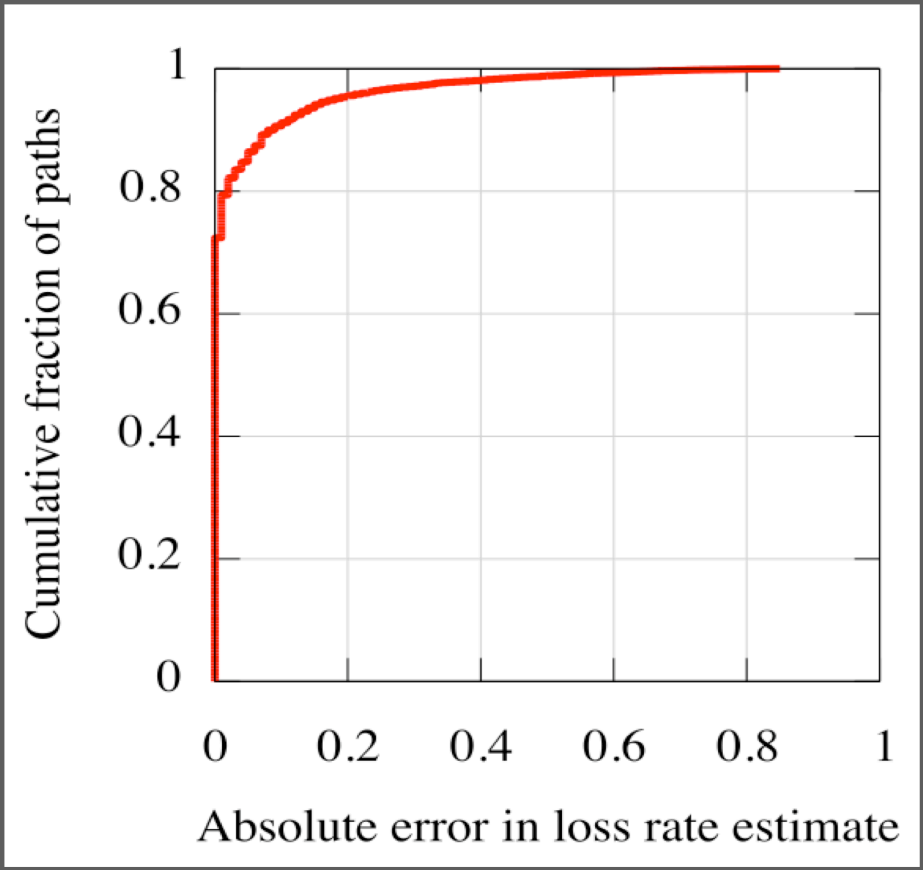
Finally done building iPlane!

- How do we …
  - ✓ build a structured atlas of the Internet?
  - ✓ predict routing between arbitrary end-hosts?
  - ✓ measure properties of links in the core?
  - ✓ measure links at the edge?

Does the combination of all this work?

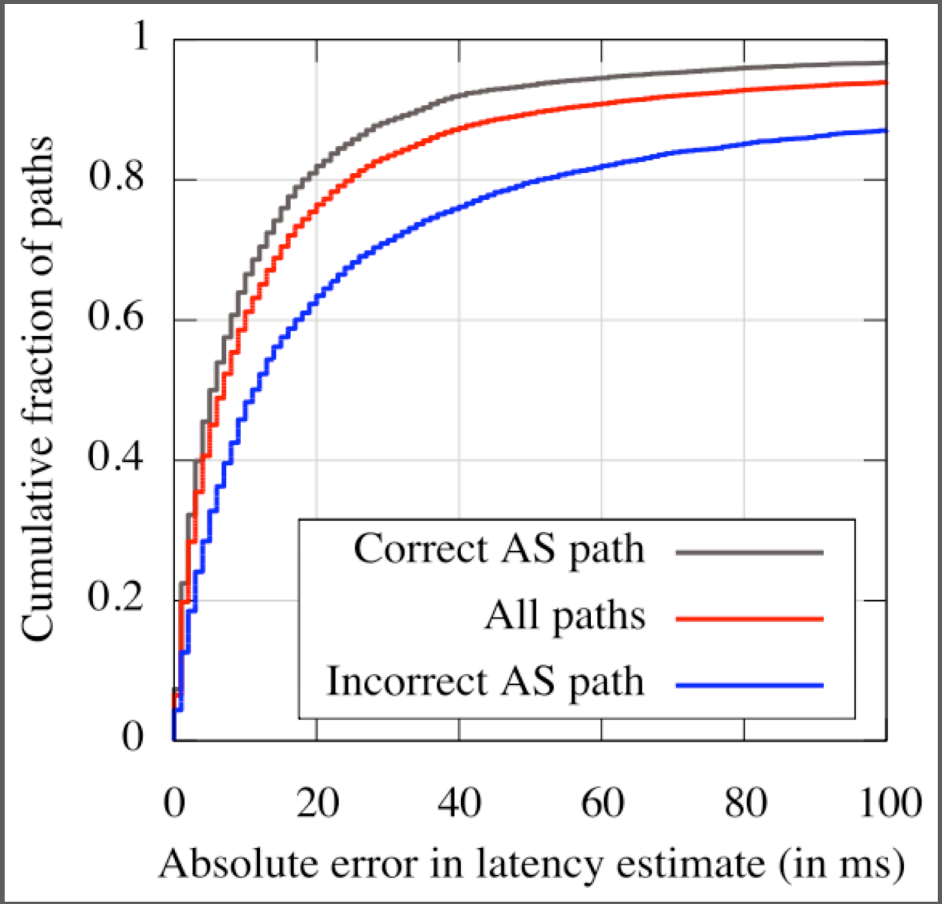
Accuracy of Predictions

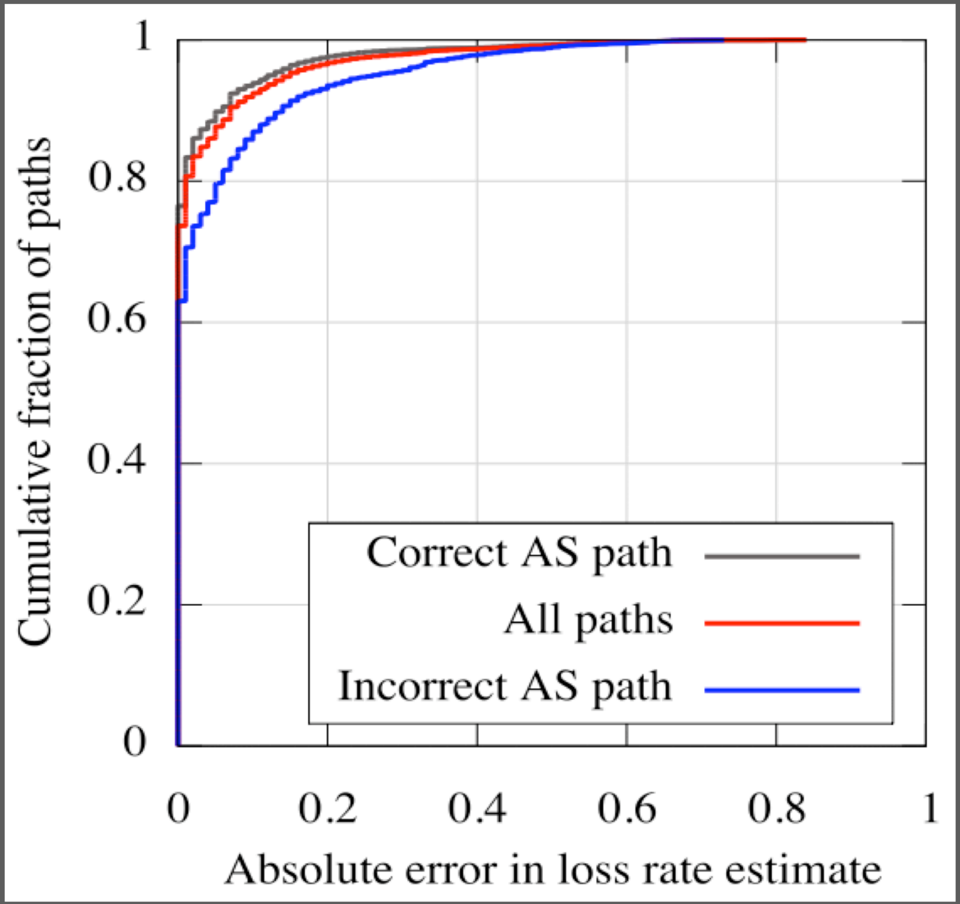




- For paths between all pairs of PlanetLab nodes
  - Latency estimates within 10ms for 61% of paths
  - Loss-rate estimates within 2% for 82% of paths

Room for Improvement





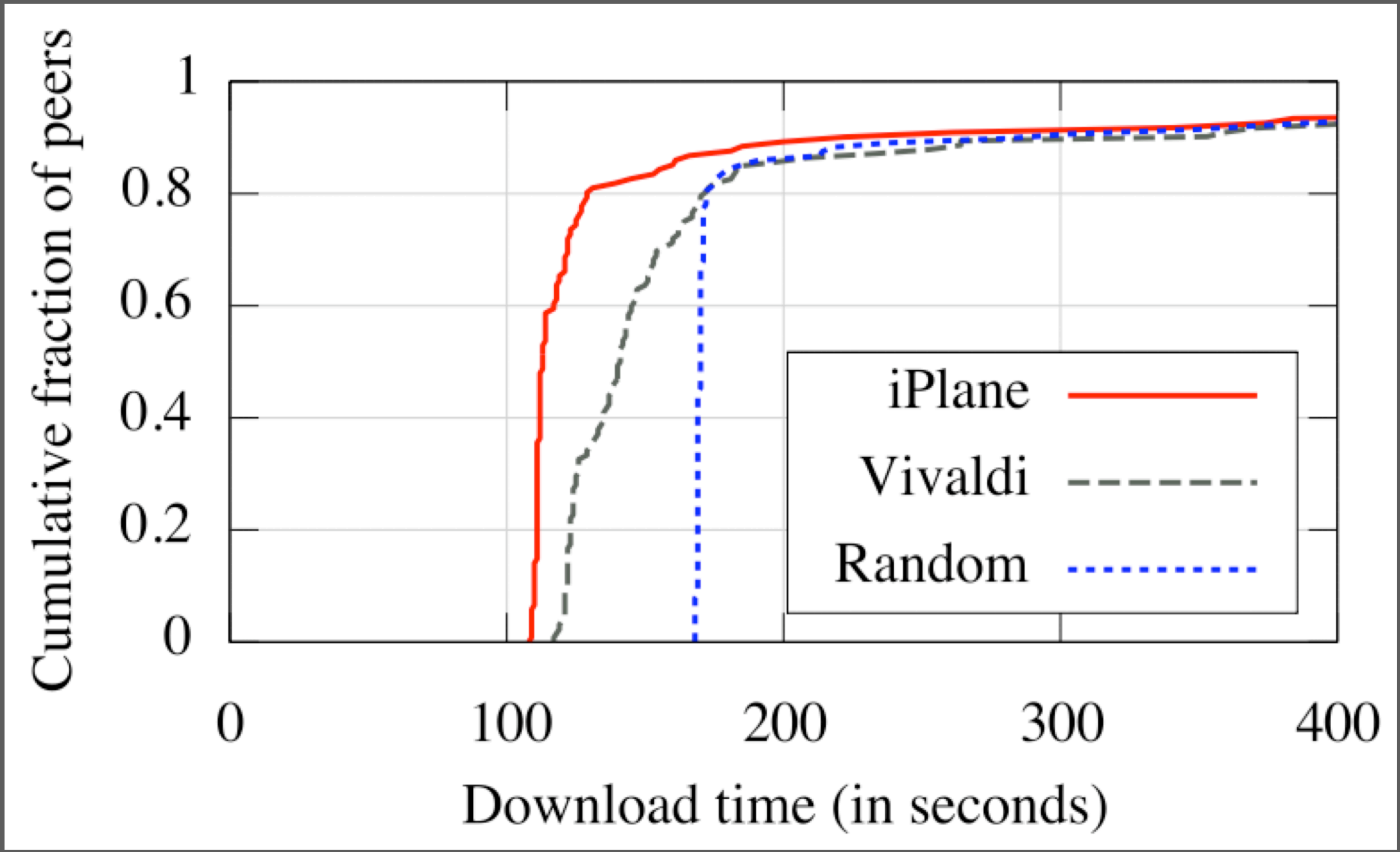
- Estimates are likely to improve with better mapping and path prediction techniques

Improving Distributed Services

- Used *iPlane*'s predictions to improve 3 apps
  - BitTorrent
    - Select peers that provide good performance
  - CDN
    - Direct each client to best performance replica
  - VoIP
    - Choose detour nodes to bridge hosts behind NATs

- *Refer to paper for CDN and VoIP experiments*

Improving BitTorrent



- 150 nodes participated in a swarm for a 50 MB file
  - 80% of peers do better than default BitTorrent

Conclusions

- We have implemented *iPlane*: an information plane
  - Maps the Internet's structure to predict multiple path properties between arbitrary end-hosts
- Demonstrated utility of *iPlane* in helping distributed applications deliver better performance

- Traces gathered by *iPlane* available at

http://iplane.cs.washington.edu

Future Work

- Refresh selected portions of the atlas more often
- More accurate model for path prediction
- Account for routing asymmetry in measuring link properties