# The End-to-End Effects of Internet Path Selection

Stefan Savage
Andy Collins, Eric Hoffman, John Snell
Tom Anderson

Department of Computer Science and Engineering
University of Washington

# Motivation

- Routing is a black box
  - Packets follow some path chosen for you
  - That path has a delay and a loss rate
  - Maybe that was the best path… probably not
- Our goal:
  - Quantify and understand the impact of *path selection* on end-to-end performance

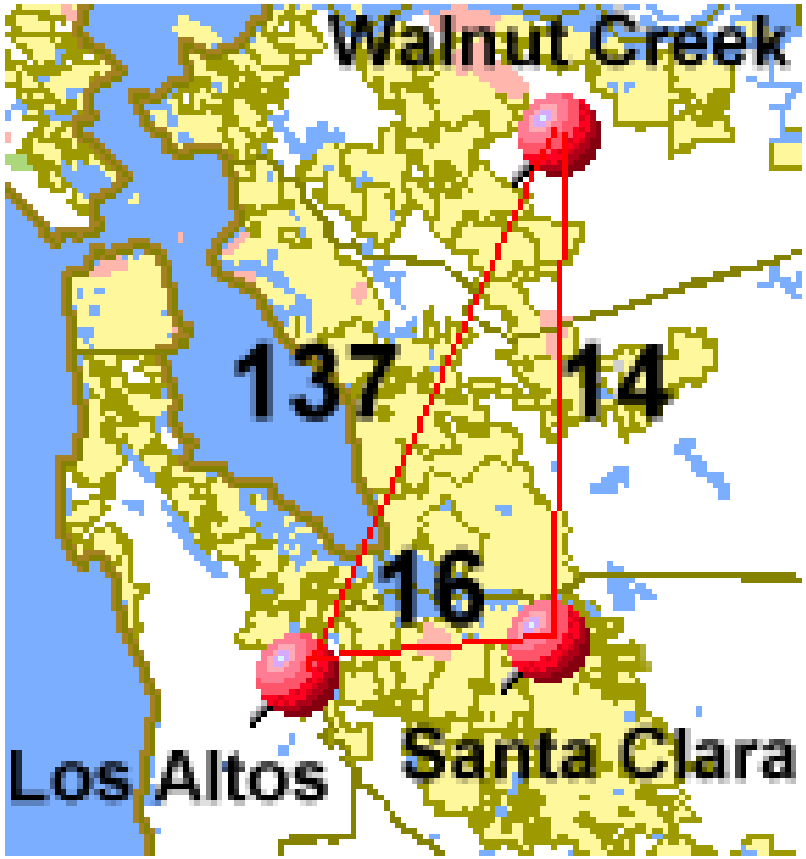
# Anecdotal evidence

Does path selection impact performance?

YES

(sometimes a lot)

How often, how much, and why?




Measured round-trip times (in ms)

# Overview

- Methodology
- Basic results
- Sources of delay
- Conclusions and open questions

# Quantifying the impact of path selection

- Basic metric:

  Let X = performance of default path

  Let Y = performance of best path

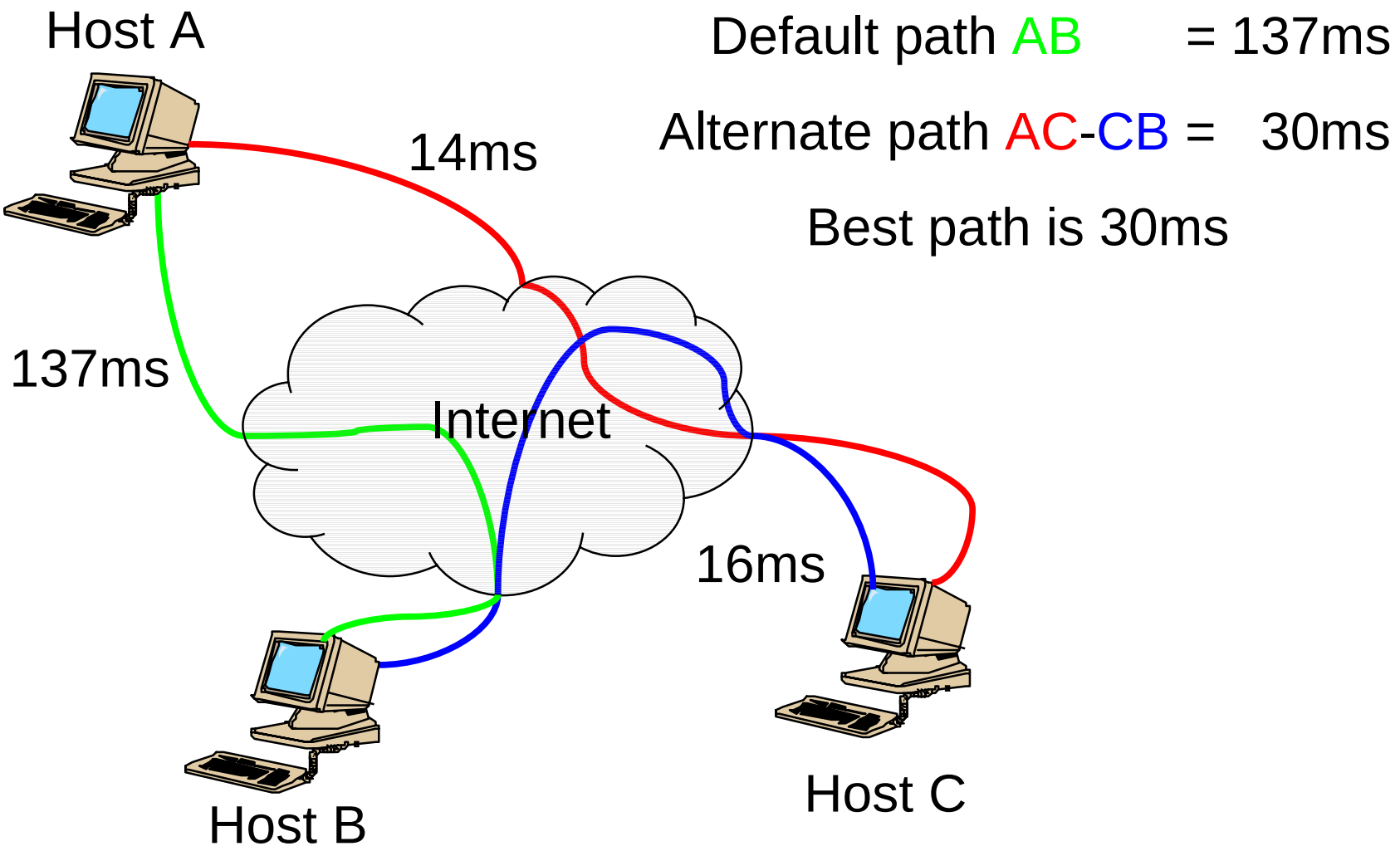
  Y-X = cost of using default path

- Technical problems
  - How to find the best path?
  - How to measure the best path?

# Approximating the best path

- Key idea
  - Use end-to-end measurements to extrapolate potential alternate paths
- Rough algorithm
  - Measure paths between pairs of hosts
  - Generate *synthetic* topology – full $N^2$ mesh
  - Find best alternate path through this graph
- Conservative approximation of *best* path

# Example:
# Lowest latency path from A to B



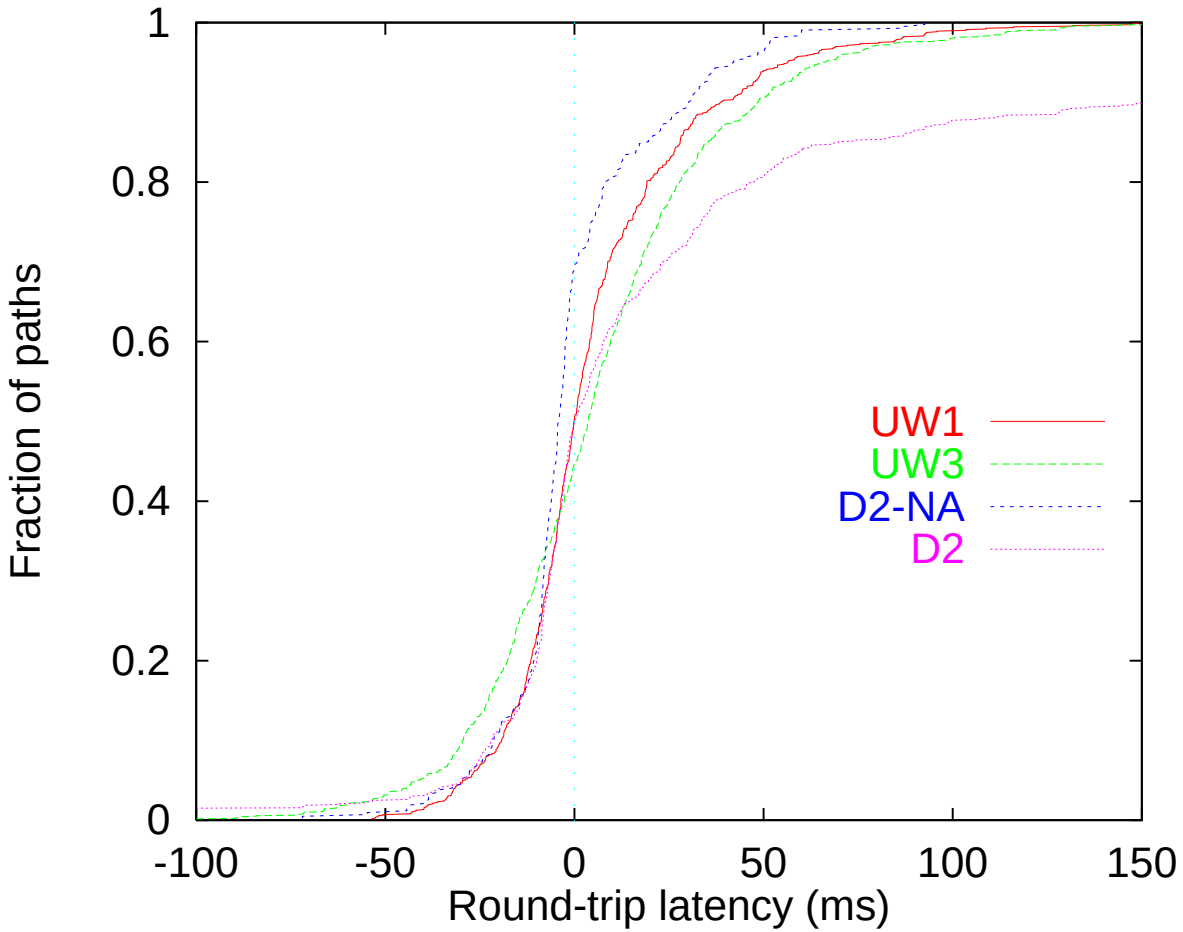
# Visualizing “routing efficiency”

- For each pair of hosts, calculate
  - Average round-trip time
  - Average loss rate
  - Average bandwidth
- Generate synthetic alternate paths based on long-term averages
- For each pair of hosts, graph difference between default path and best alternate

# Input trace data

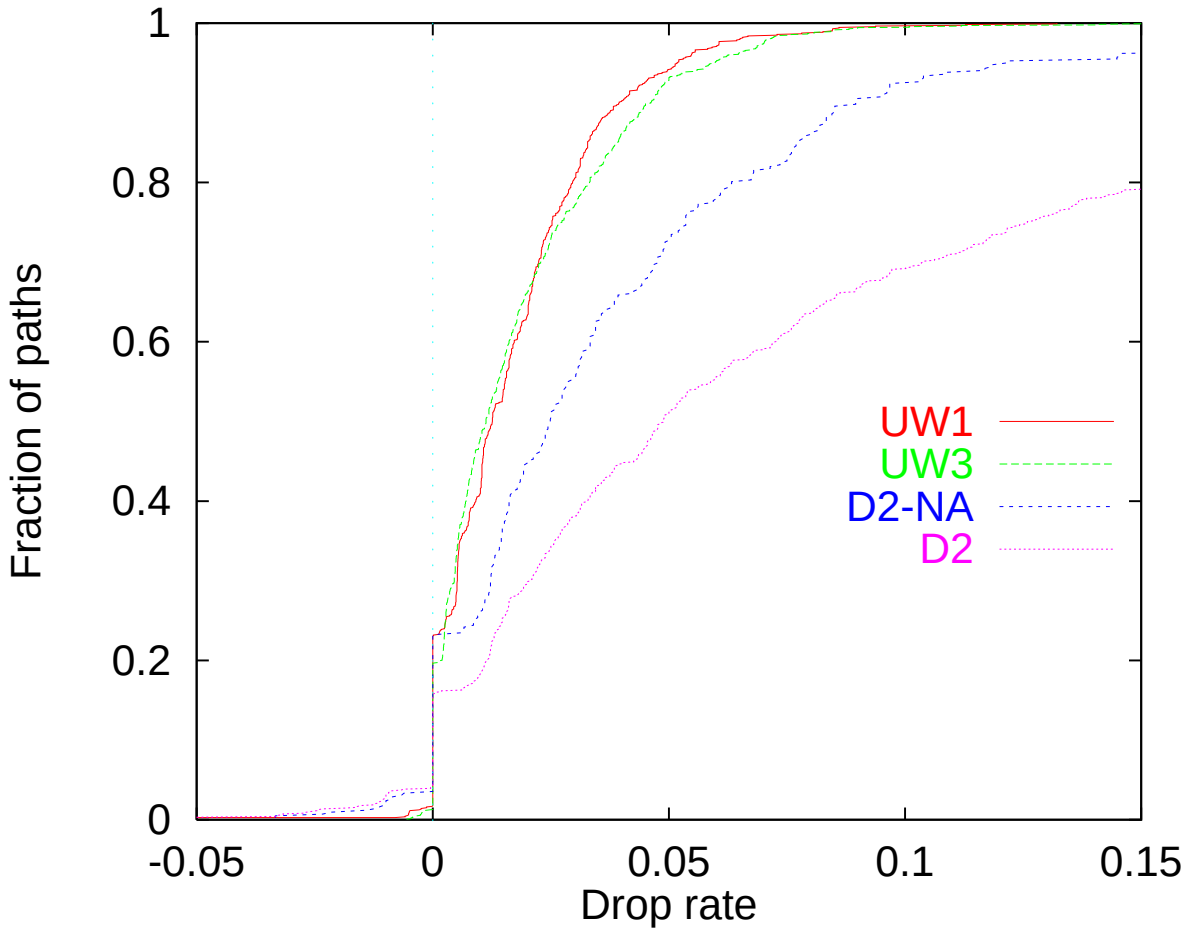
- Traceroute between hosts
  - Random requests: *UW1, UW3, UW4A, D2*
  - Synchronized requests: *UW4-B*
- Analysis of TCP transfers between hosts
  - Random requests: *N2*
- Full descriptions of datasets in our paper, [Paxson96], and [Paxson97]

# Round-trip time



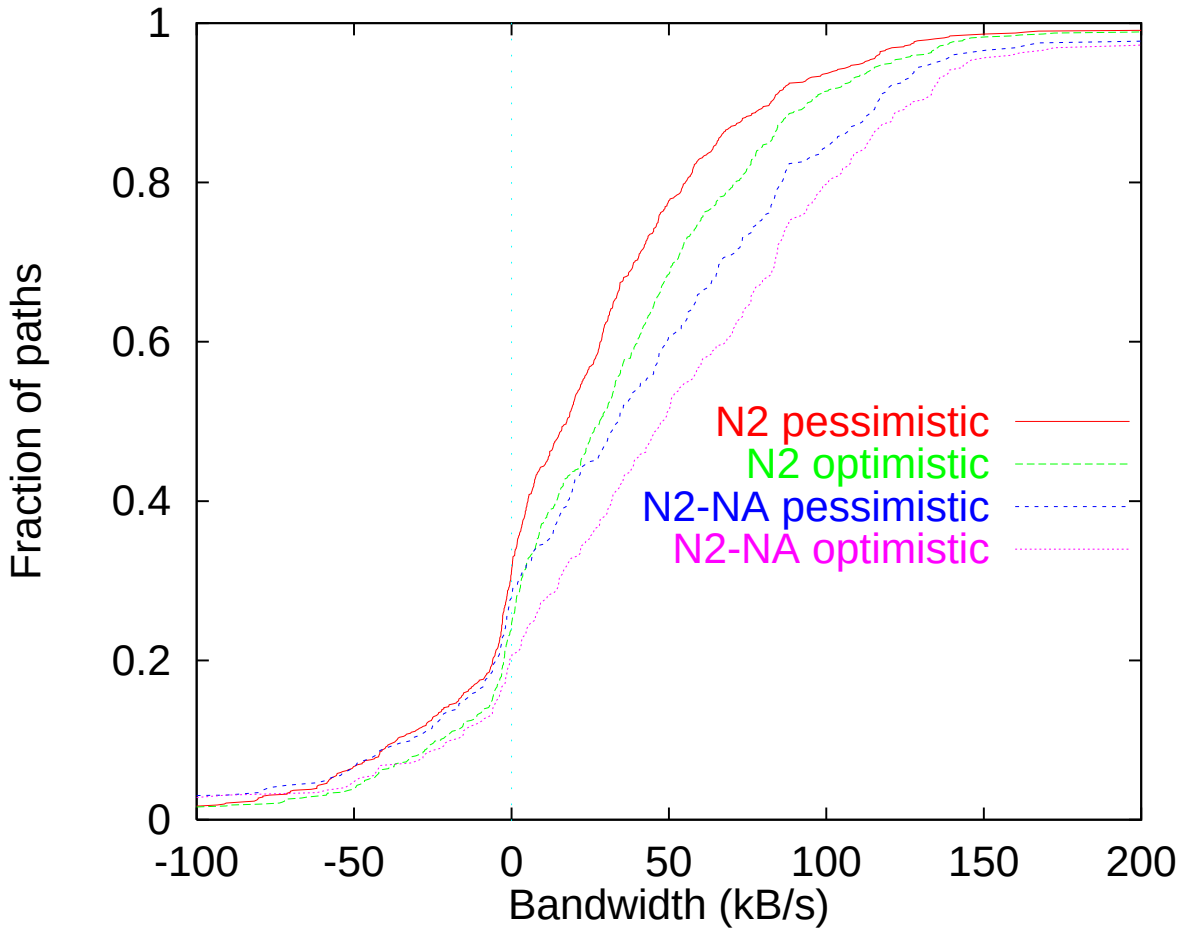

30%-55% of default paths have longer round-trip times

# Loss rate



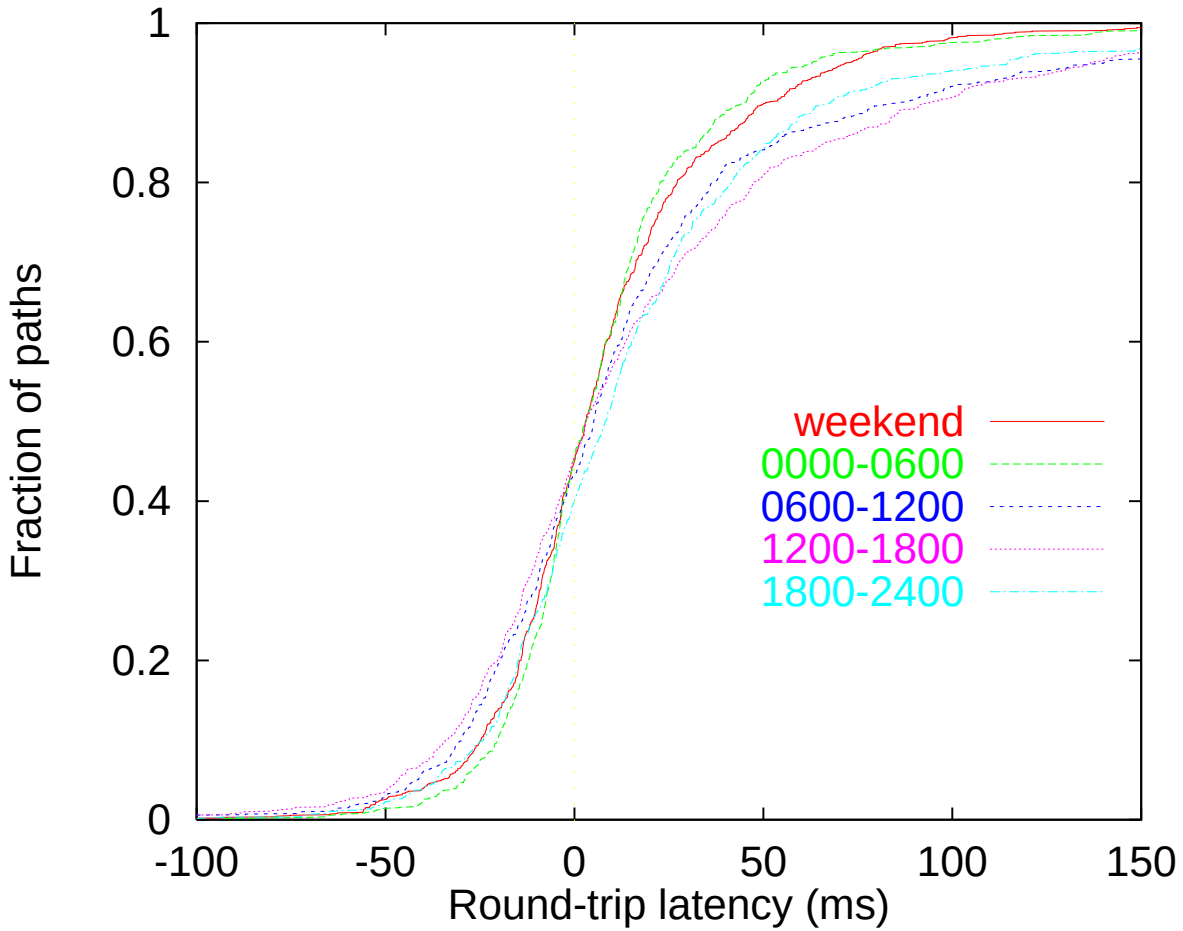

75%-85% of default paths have higher loss rates

# Bandwidth



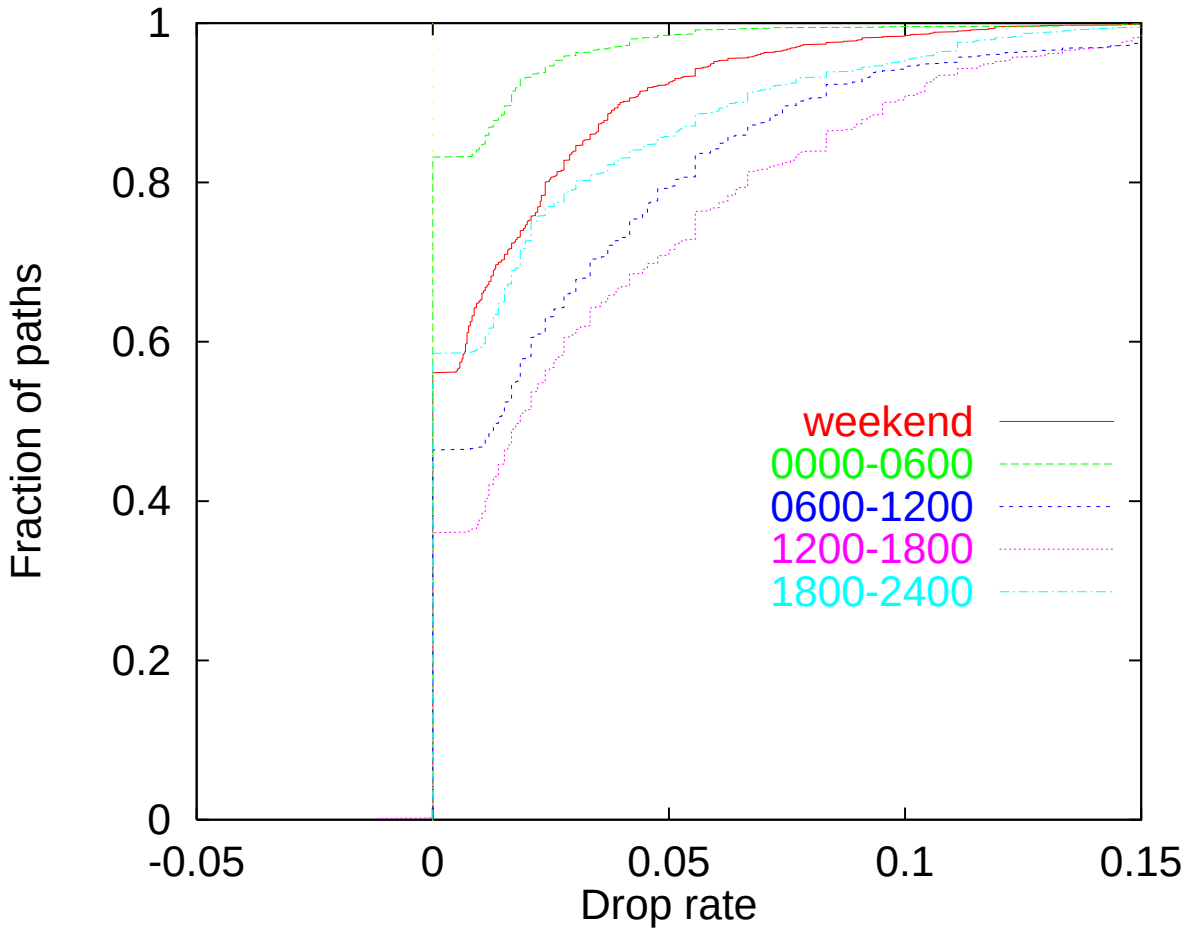

70%-80% of default paths have lower bandwidth

# Time-of-day variation (latency)




Effect stronger during “peak” hours

# Time-of-day variation (loss)




Even stronger peak effect for loss

# Quick summary of results

- The default path is usually not the best
  - True for latency, loss rate, and bandwidth
  - **In spite** of synthetic end-host transiting
- Many alternate paths are much better
- Effect stronger during peak hours

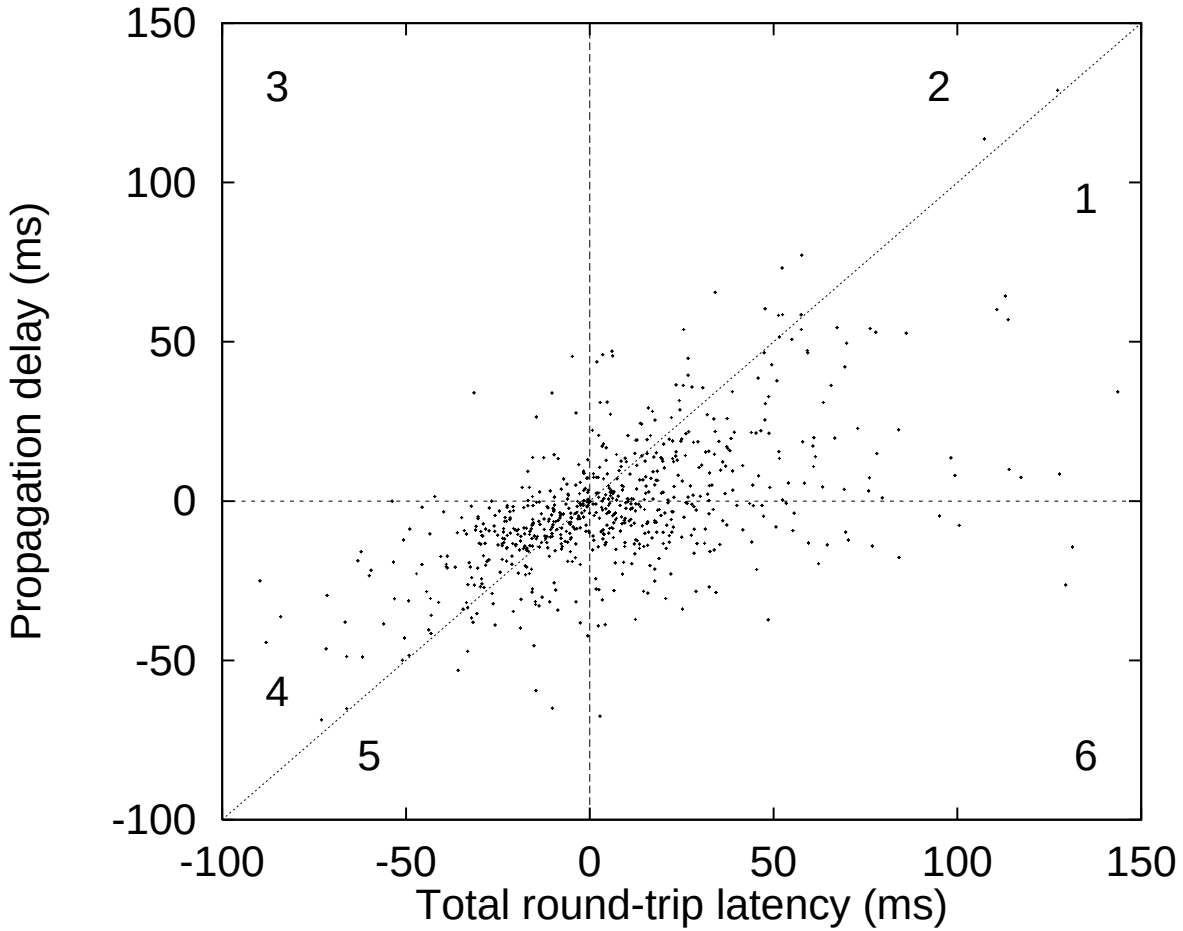
# Quick summary of results

- The default path is usually not the best
  - True for latency, loss rate, and bandwidth
  - **In spite** of synthetic end-host transiting
- Many alternate paths are much better
- Effect stronger during peak hours
- Better paths can be shorter, less congested, or both

# What makes a better path better?

- Possibilities
  - Avoids congested queues
  - Shorter propagation delay
- Answer seems to be: **both**
- Visualizing propagation and congestion
  - Estimate propagation delay ($10^{th}$ percentile)
  - Queuing delay = RTT – propagation delay
  - Graph improvement in propagation delay vs improvement in RTT

# Propagation delay vs congestion



# Why path selection isn't "perfect"

- Technical reasons
  - Single-path routing
  - Non-topological route aggregation
  - Coarse routing metrics (AS_PATH)
  - Local policy decisions
- Economic reasons
  - Disincentive to offer transit
  - Minimal incentive to optimize transit traffic

# Open questions

- How do these results scale to large populations?
- Which routing mechanisms have the biggest impact on performance?
- Can we improve performance without violating policy?  (guess: yes)
- Is there a role for application-level “value-added routing”?

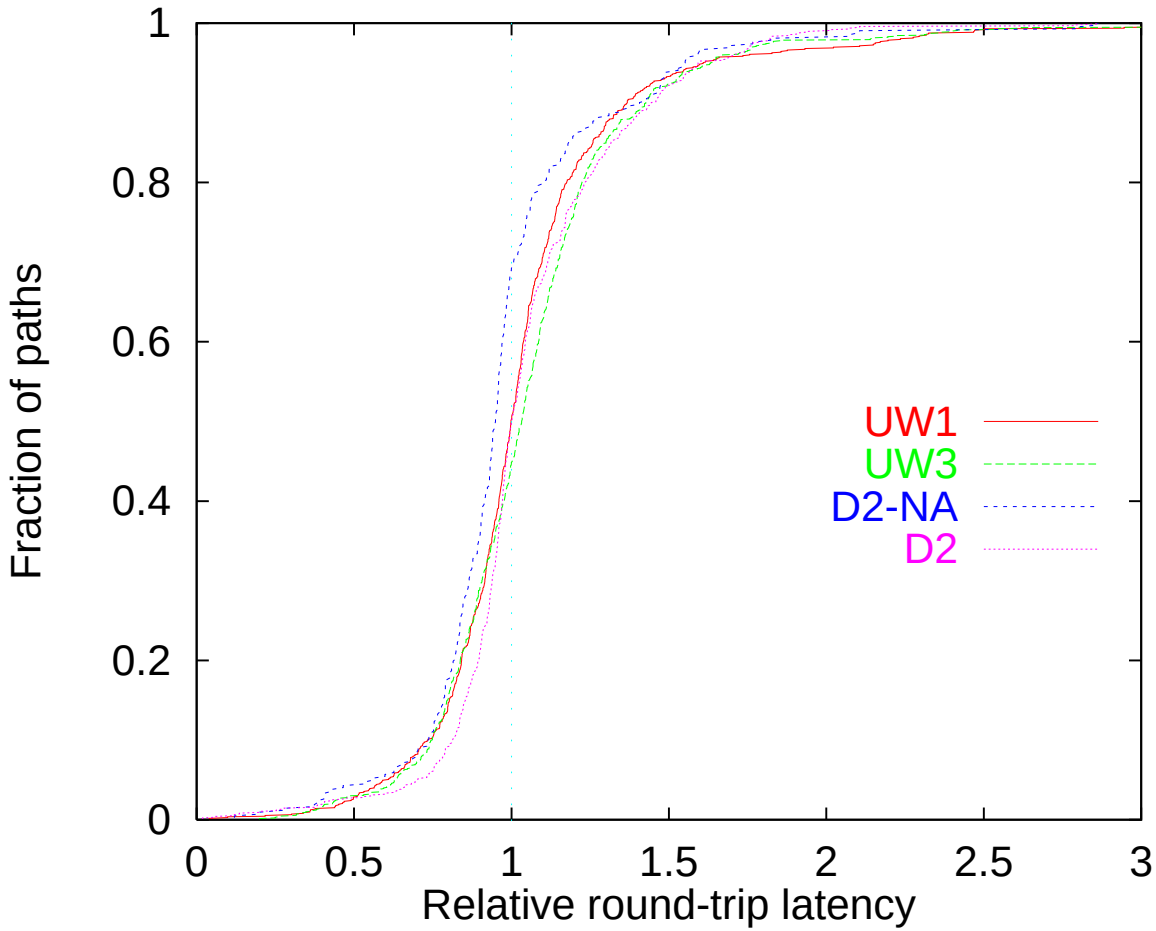
# Conclusions

- We can roughly quantify the impact of path selection on performance
- Routing is clearly a significant part of the end-to-end performance equation

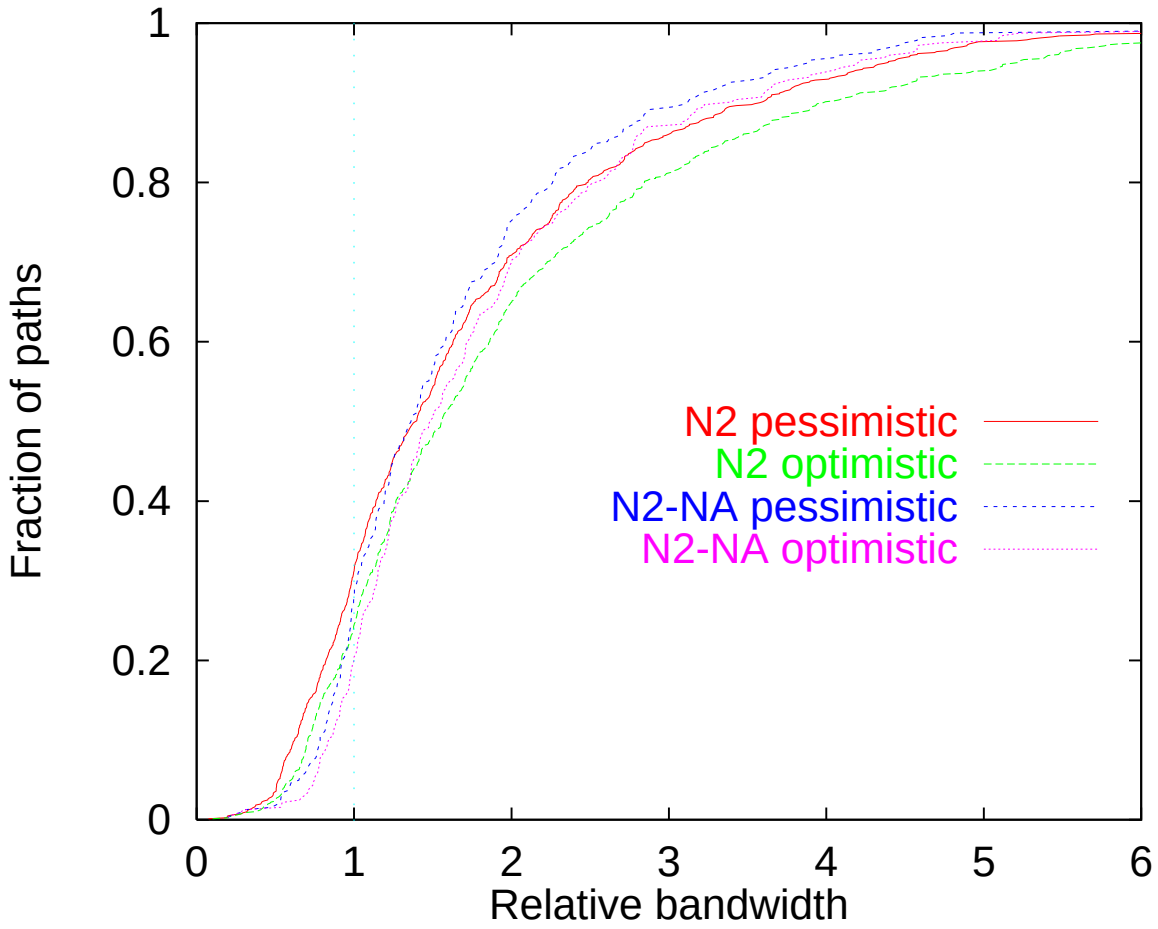
# Calculating synthetic path metrics

- Round-trip time: AB + BC = ABC
- Loss rate: 1-((1-AB)*(1-BC)) = ABC
- Bandwidth
  - Pessimistic: same as above
  - Optimistic: MAX(AB,BC) = ABC
  - Solve using [Mathis97] approximation for TCP bandwidth
- Use distribution convolution for medians

# Relative round-trip latency



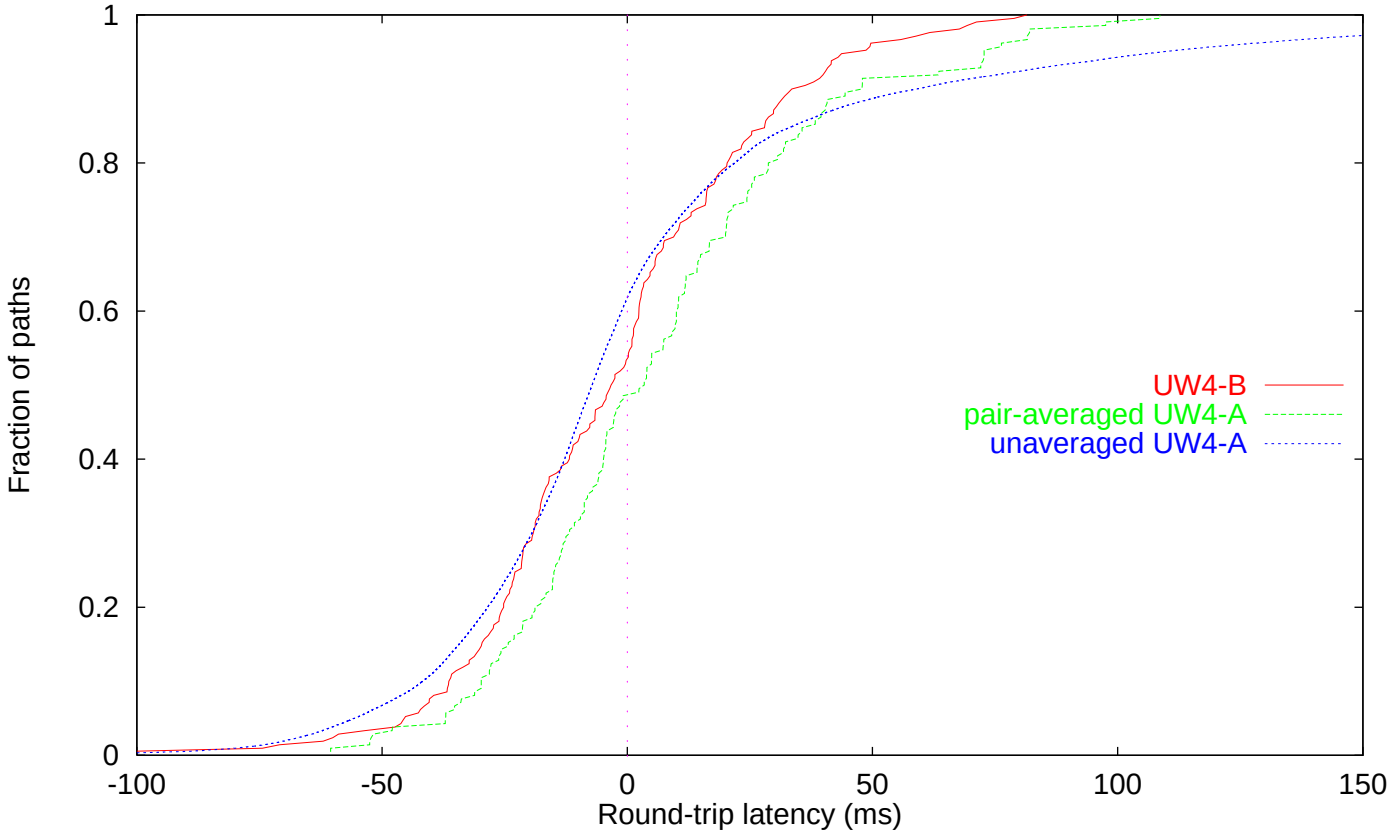
# Relative bandwidth



# Robustness

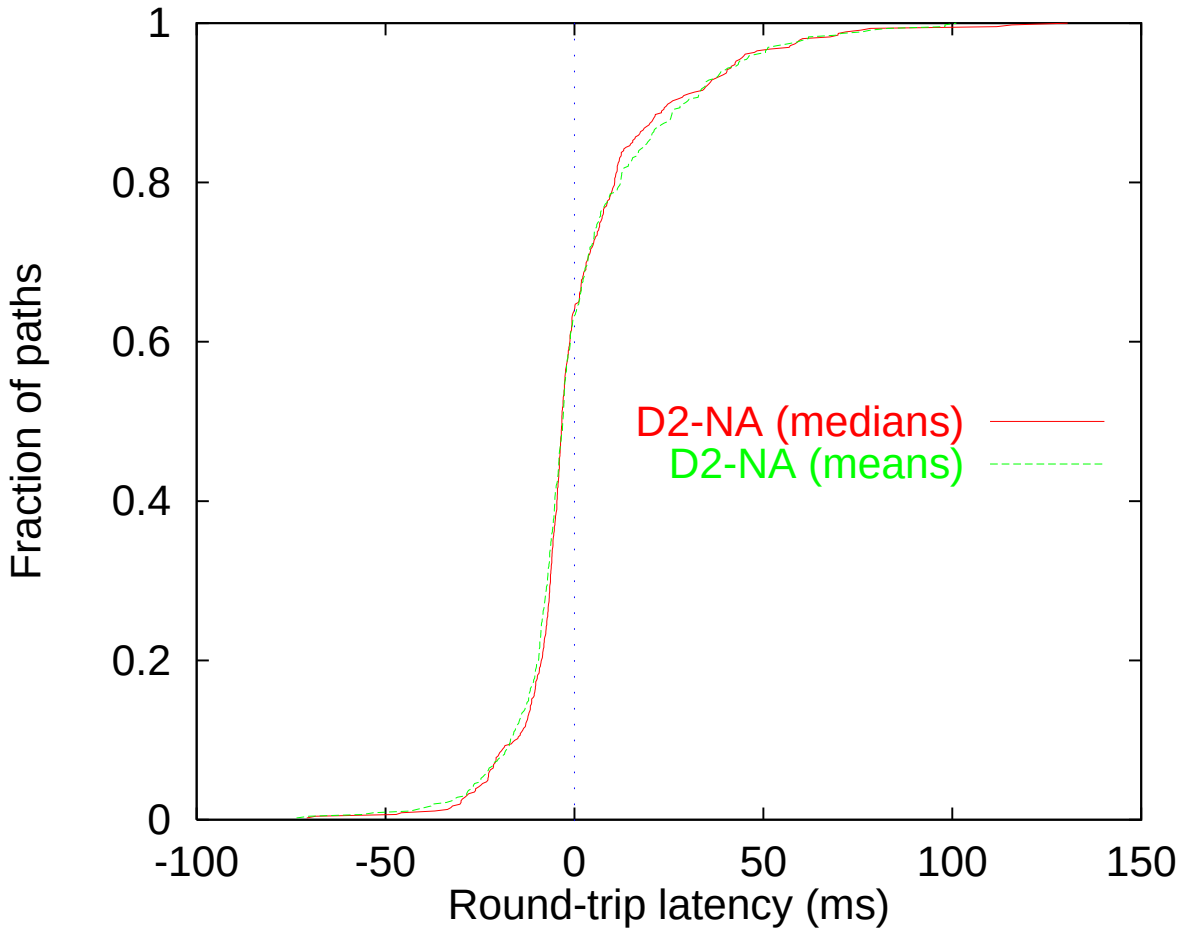
- High variability, but results not accounted for by random variation
- Equivalent results using mean and median
- Basic findings same across multiple time scales
- Results not explained by small number of hosts or networks

# Short time scales (latency)

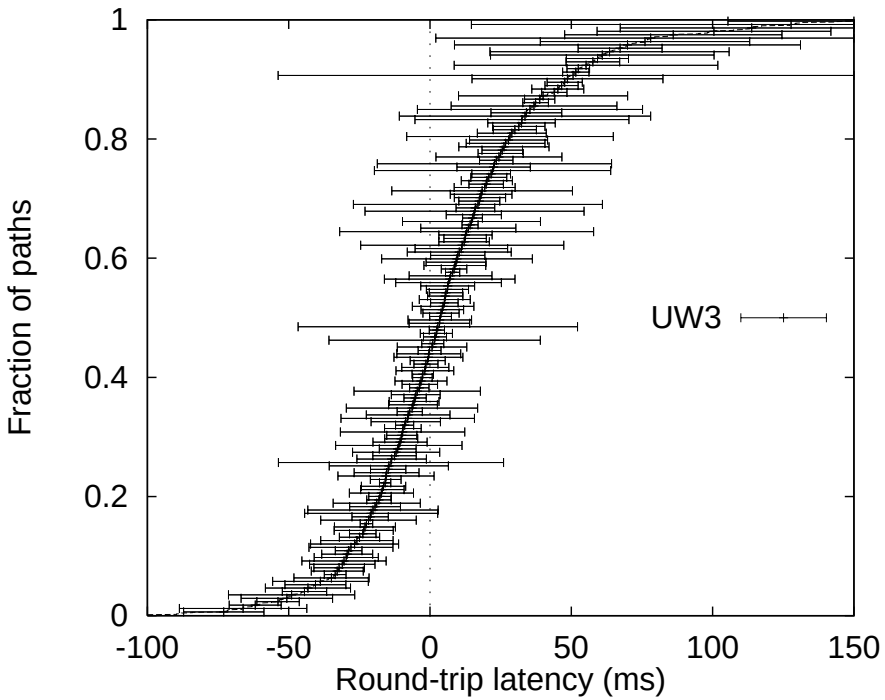


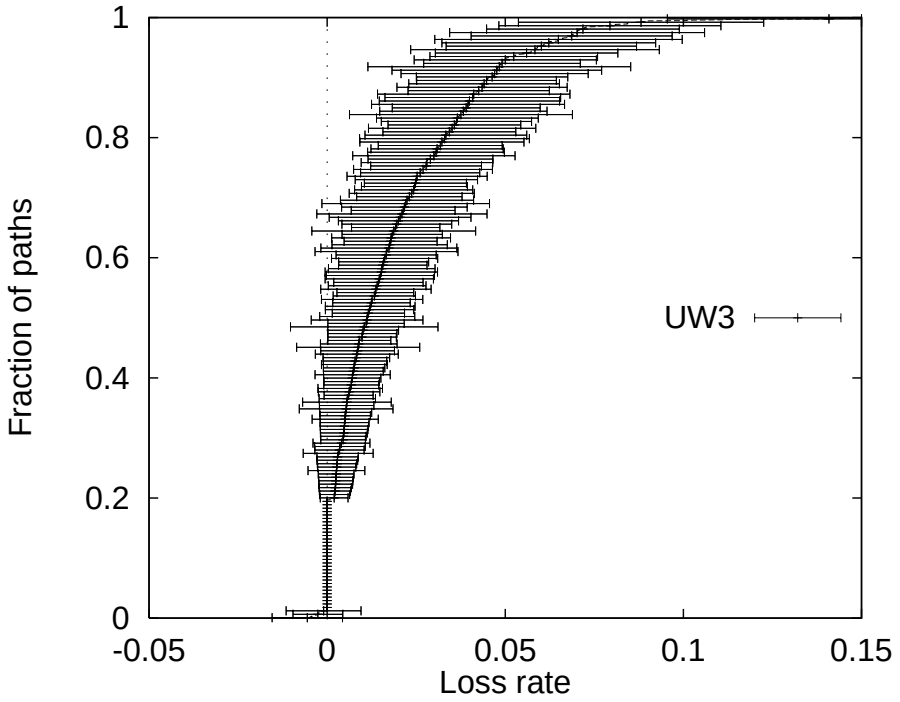

Same opportunities exist at short time scales

# Mean vs Median

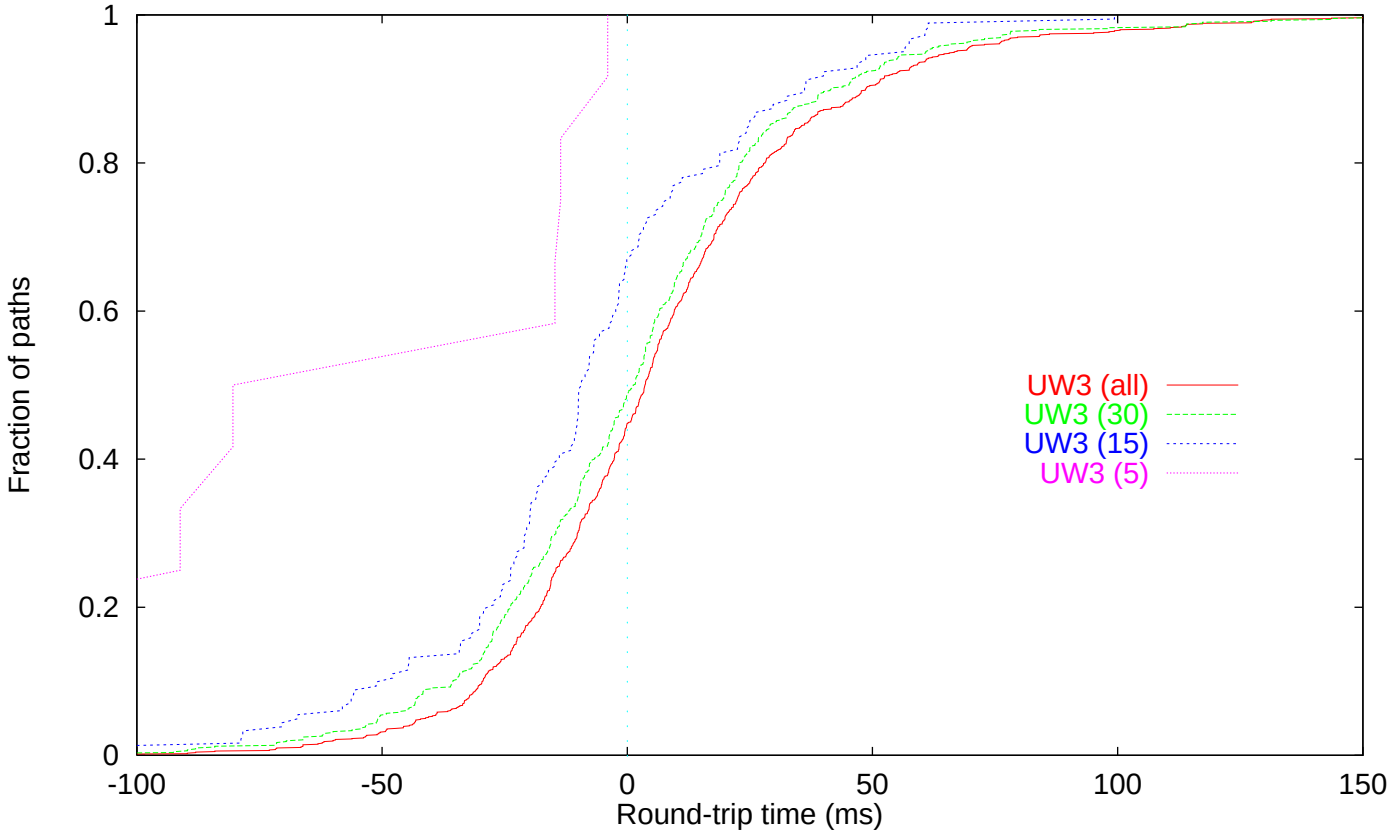


# Confidence Intervals

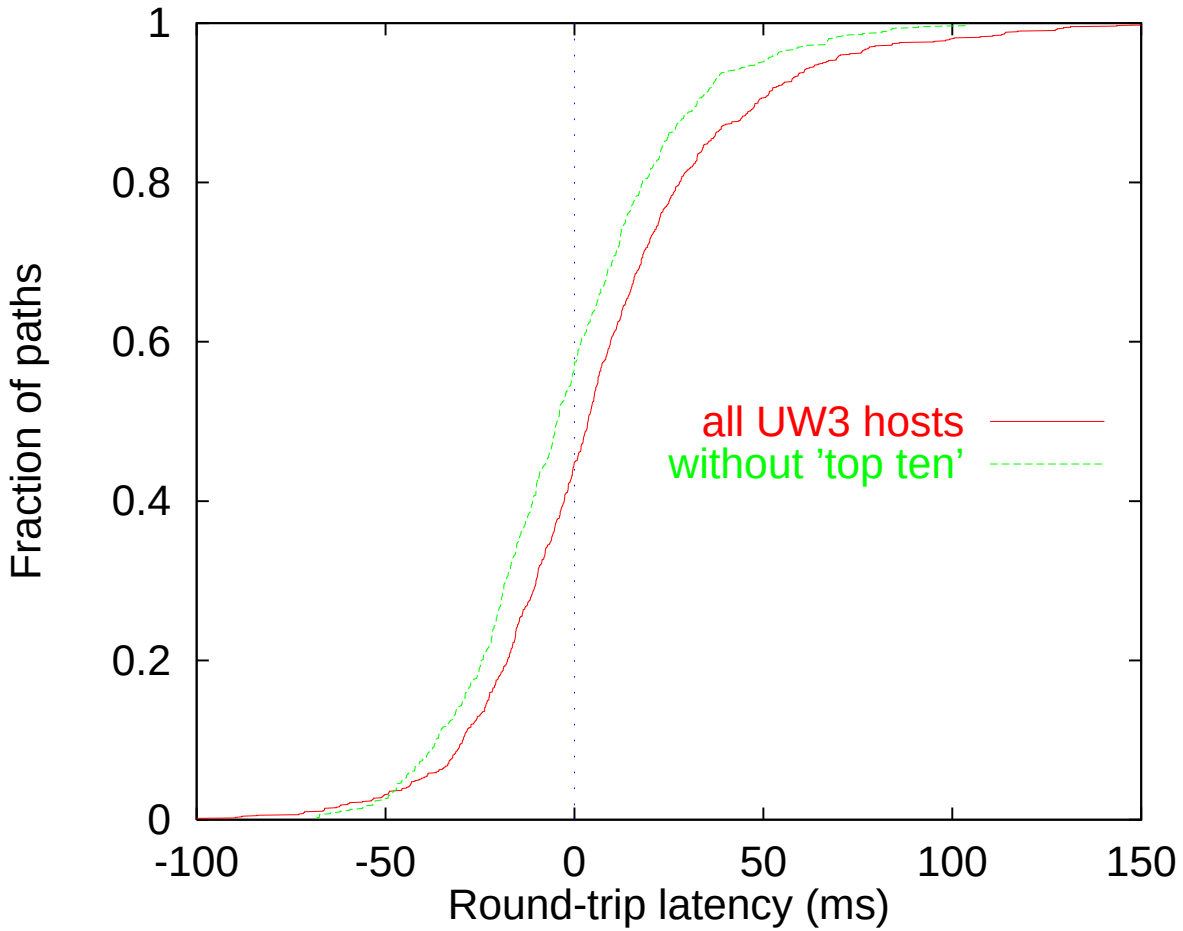


# Importance of sample size



# Are there a few bad or good hosts?

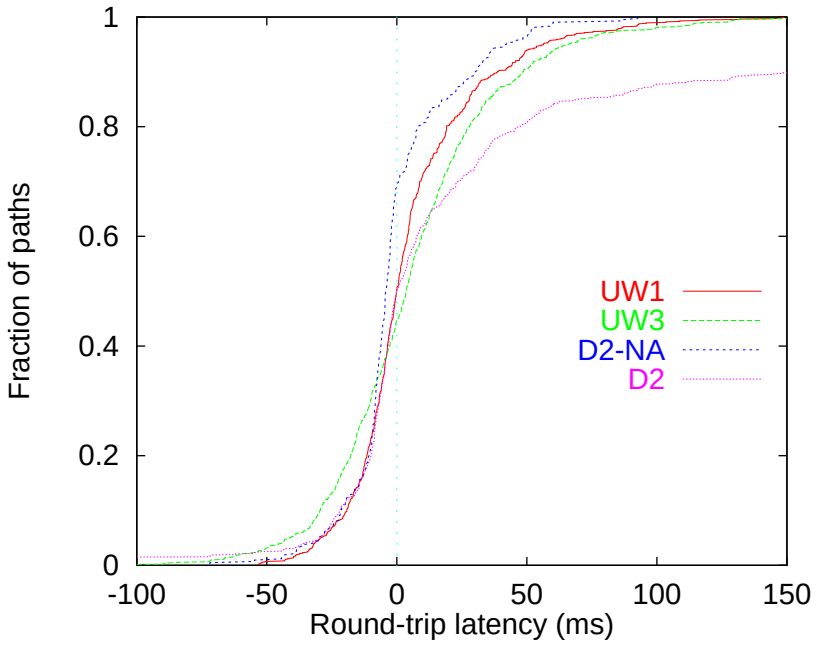


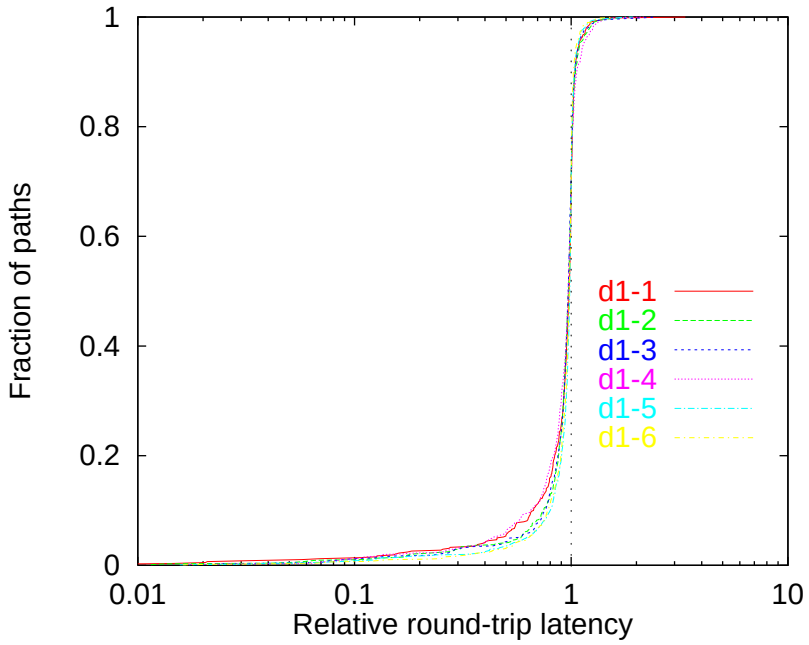
# Datasets characteristics

| Name | Year | Duration (days) | Hosts | Number of measurements |
|---|---|---|---|---|
| UW1 | 1998 | 34 | 36 | 54034 |
| UW3 | 1998 | 7 | 39 | 94420 |
| UW4-A | 1999 | 14 | 15 | 216928 |
| UW4-B | 1999 | 14 | 15 | 9169 |
| D2 (NA) | 1995 | 48 | 33 (22) | 35109 (14896) |
| N2 (NA) | 1995 | 44 | 31(20) | 18274 (7582) |

# Traceroute server placement (UW1)

# Our results vs IDMaps results



# Resolving differences with IDMaps results

- Presentation differences
  - Relative vs absolute metrics
  - Log scale vs linear scale
  - Trim tail vs follow tail
- Real differences
  - Mean vs min RTT metric
  - Error in IDMaps SPF algorithm