Uncovering Performance Differences Among Backbone ISPs with Netdiff

Ming Zhang (Microsoft Research)

Joint work with

Ratul Mahajan (Microsoft Research)
Lindsey Poole, Vivek Pai (Princeton)

Benchmarking computer systems

- Benchmarking is common today for databases, web servers, compilers, file systems, ….
- Systematic quantification of system performance
- Enables comparison across competing vendors
  - Helps consumers make informed decisions
  - Helps vendors make relevant optimizations
  - The resulting transparency and competition helps the industry at large

Today benchmarks do not exist for Internet Service Providers (ISPs)

- ISPs' customers do not know which ISP is best for their traffic
  - Even though it has a big impact on application performance
- ISPs may have little incentive to improve end-to-end customer performance

Existing methods to compare ISPs

- Service-level agreements (SLAs)
  - Typically within an ISP's network
  - Not directly related to application performance
  - Hard to compare across ISPs
- Keynote systems
  - 1,200 nodes inside ISP PoPs
  - Measure individual path
    - Insufficient for systematic comparison

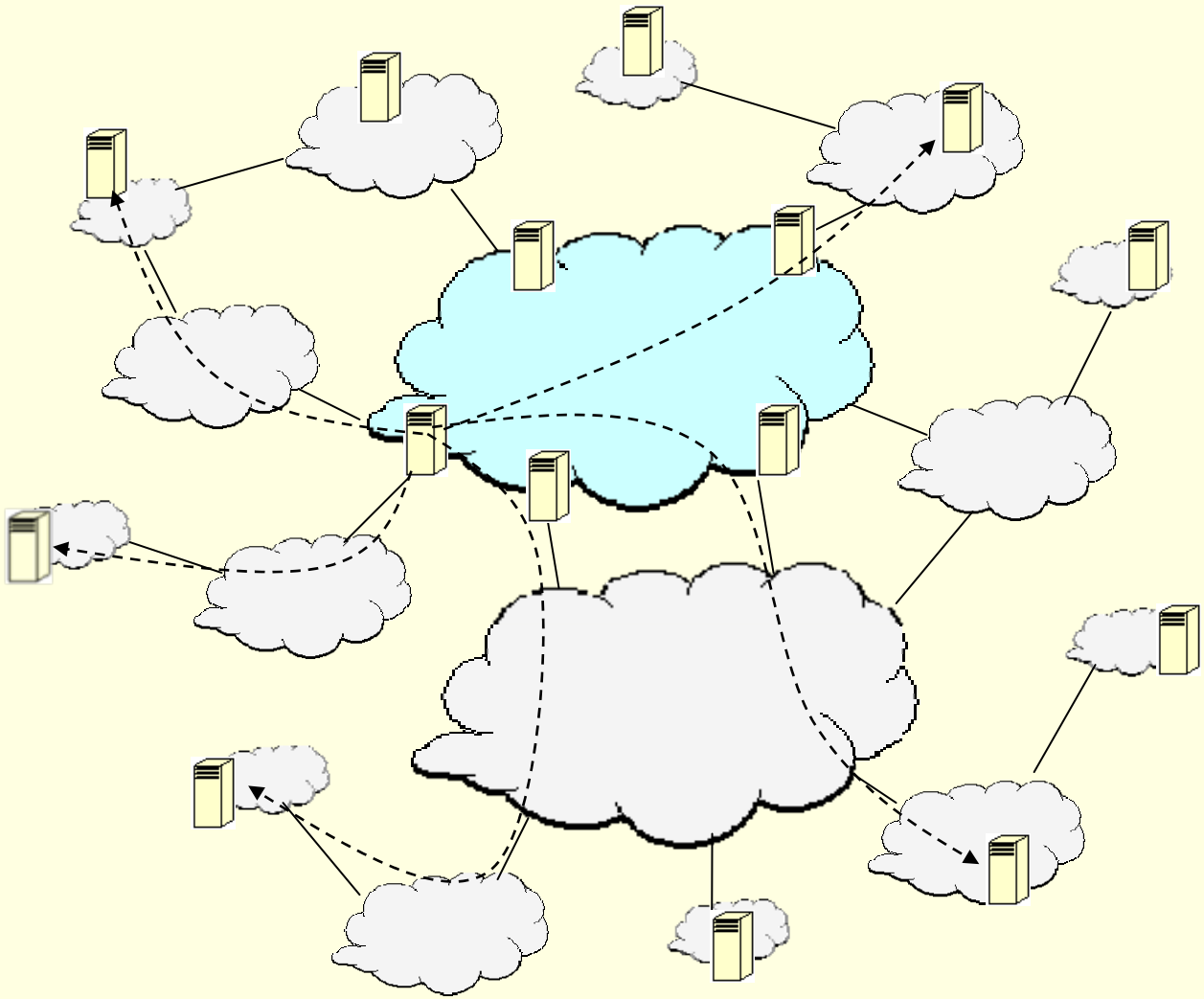
First step towards systematic ISP benchmarking

- Introduce the Netdiff system
  - An easily deployable system that does not require ISP cooperation
  - Provides detailed differences between backbone ISPs
  - Performance measure: path latency

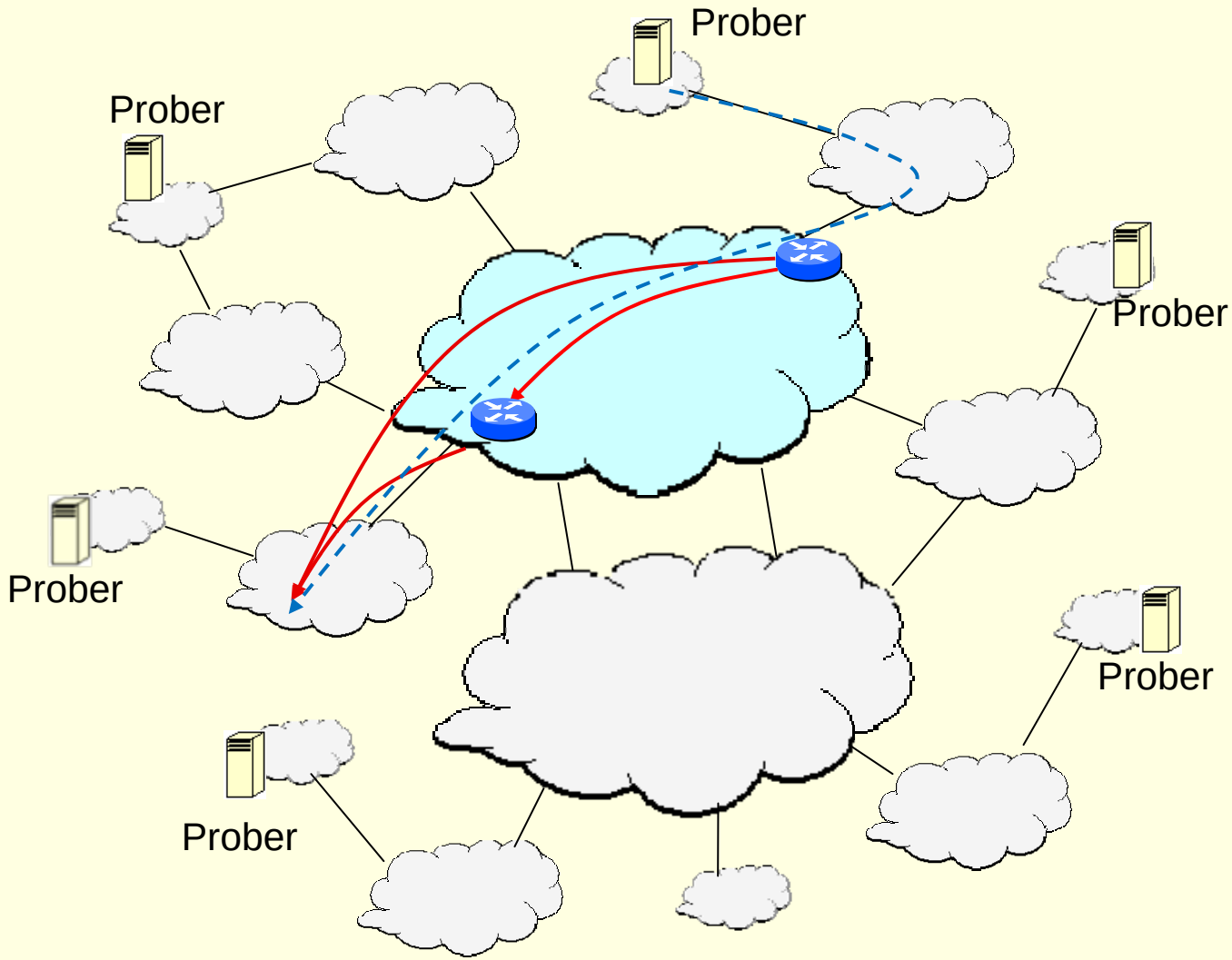
Requirements of ISP comparison

- Relevant to customers
  - Measure end-to-end path
  - Target destinations of interest
  - Compare based on workload
- Helpful to ISPs
  - Account for size and geographic presence
  - Internal *vs.* external
  - Bad Point-of-Presence (PoP) or destination

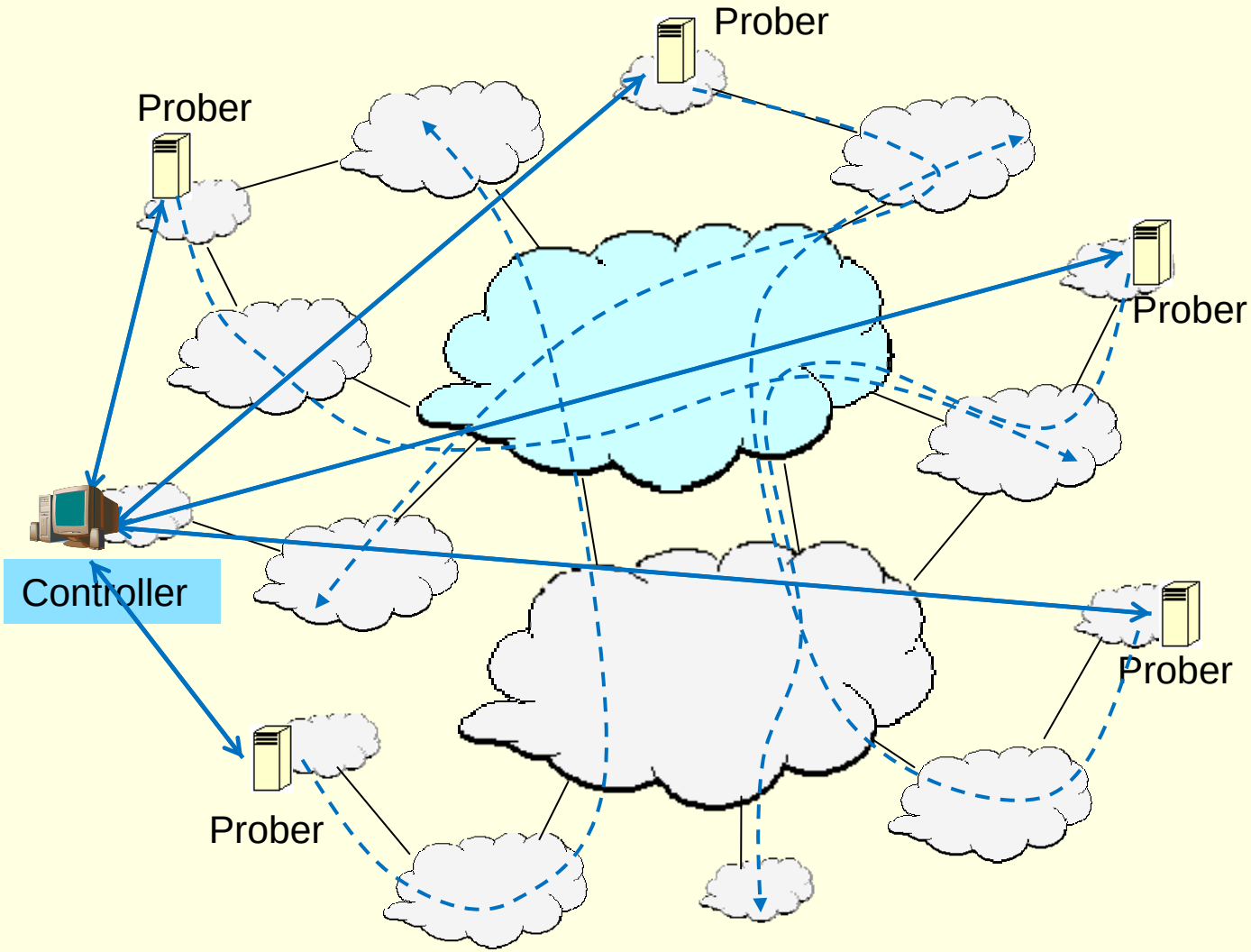
Ideal architecture



Netdiff architecture



Measurement process



Mapping IP to PoP/location

- Router IPs
  - DNS name contains PoP info
    - sl-gw2-sea-4-0-0.sprintlink.net -> Sprint in Seattle
  - Parse names with *undns*
- Destination IPs
  - MaxMind geo-location database
    - Based on user input
    - Accuracy: 99% country-level, 80% city-level in US

Deal with scalability

- Map ISP topology at large time scale
  - Probe to all prefixes
  - Track topological changes
- Measure ISP performance at small time scale
  - Probe a subset of prefixes
    - Each PoP to destination network
    - Each internal path between two PoPs
    - Probing load below threshold
  - Formulate as a set covering/pack problem
    - Solve using greedy algorithm

Overhead & coverage

ISP	# cities	# probes (K) (Reduction factor)	# dst. paths (% of total)	# int. paths (% of total)	# int. paths seen by Keynote (% of total)
AOL Transit	27	5 (5.5)	658 (0.05)	230 (32.8)	n/a
AT&T	113	86 (13.5)	13364 (0.24)	838 (6.6)	72 (0.6)
AboveNet	20	40 (7.6)	17258 (1.73)	277 (72.8)	n/a
British Telecom	32	11 (13.7)	4898 (0.31)	440 (44.3)	2 (0.2)
Broadwing	23	28 (9.5)	7655 (0.67)	149 (29.3)	n/a
Cogent	72	161 (10.2)	42620 (1.18)	1799 (35.2)	12 (0.2)
Deutsche Telekom	67	29 (7.7)	2266 (0.07)	129 (2.9)	0 (0.0)
France Telecom	25	31 (12.1)	6		n/a
Global Crossing			1		n/a
Level 3			7		30 (0.8)
NTT/Verio			2		6 (0.3)
Qwest			2		30 (1.1)
Savvis	41	56 (8.2)	1		20 (1.2)
Sprint	55	136 (13.1)	36366 (1.32)	1208 (40.7)	20 (0.7)
Tiscali	36	30 (8.8)	5483 (0.30)	325 (25.7)	0 (0.0)
VSNL(Teleglobe)	43	30 (14.4)	5500 (0.26)	524 (30.0)	6 (0.3)
Verizon	161	120 (11.5)	20507 (0.26)	2098 (8.1)	306 (1.2)
XO	47	7 (23.5)	1415 (0.06)	522 (24.1)	2 (0.1)

18 backbone ISPs, 20 min per ISP

5K to 23 K probes per ISP, 400x reduction

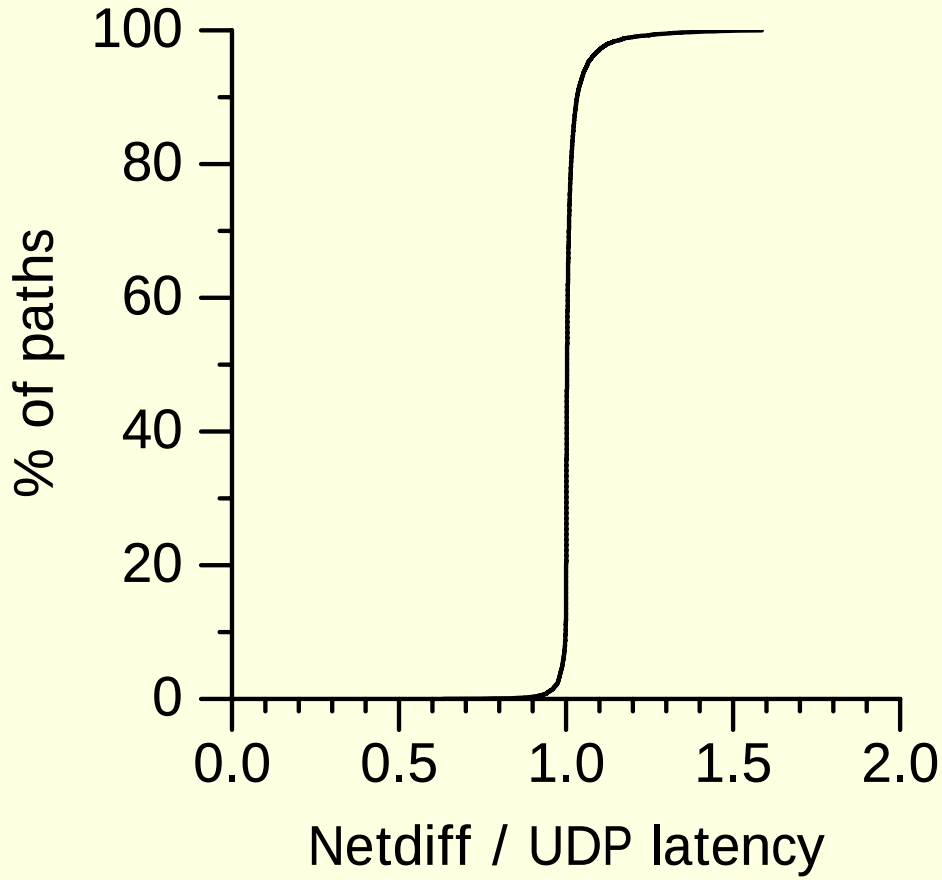
10x more internal paths

Dealing with noise and errors

- Sanity checks to guard against possible errors
  - Inferred ISP topology
  - IP-to-Geo mapping
  - Path asymmetry
  - Overloaded prober or access link

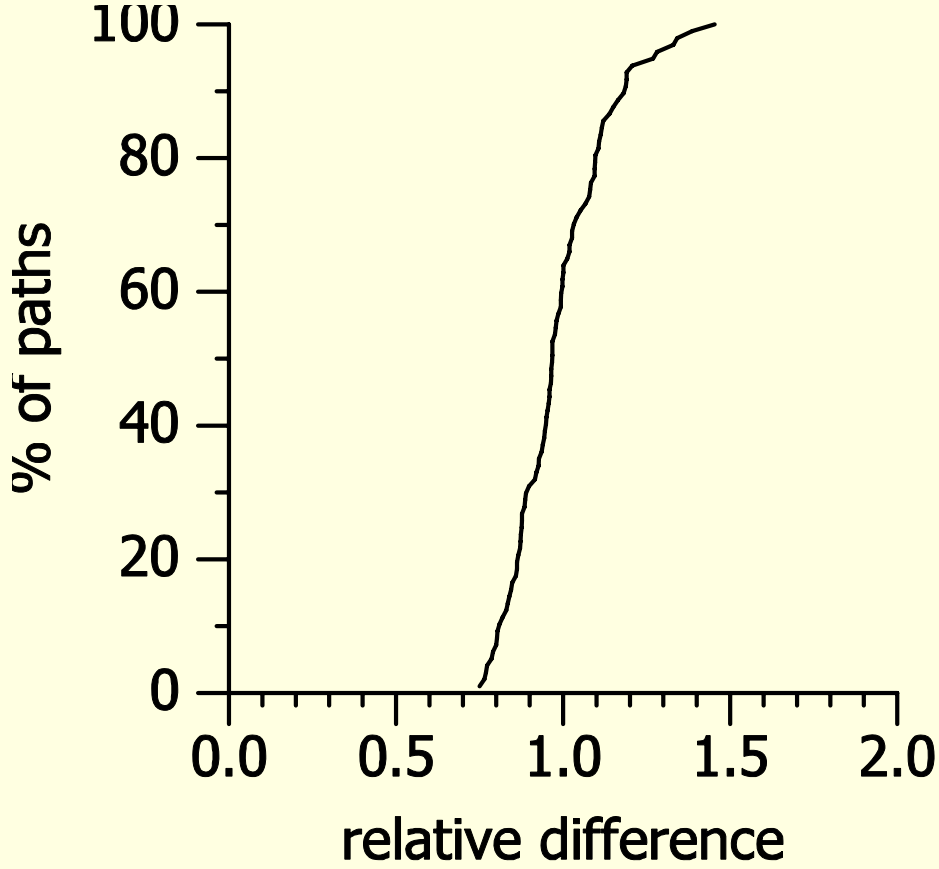
Noise due to ICMP?

- UDP on both directions *vs.* UDP on forward & ICMP on reverse

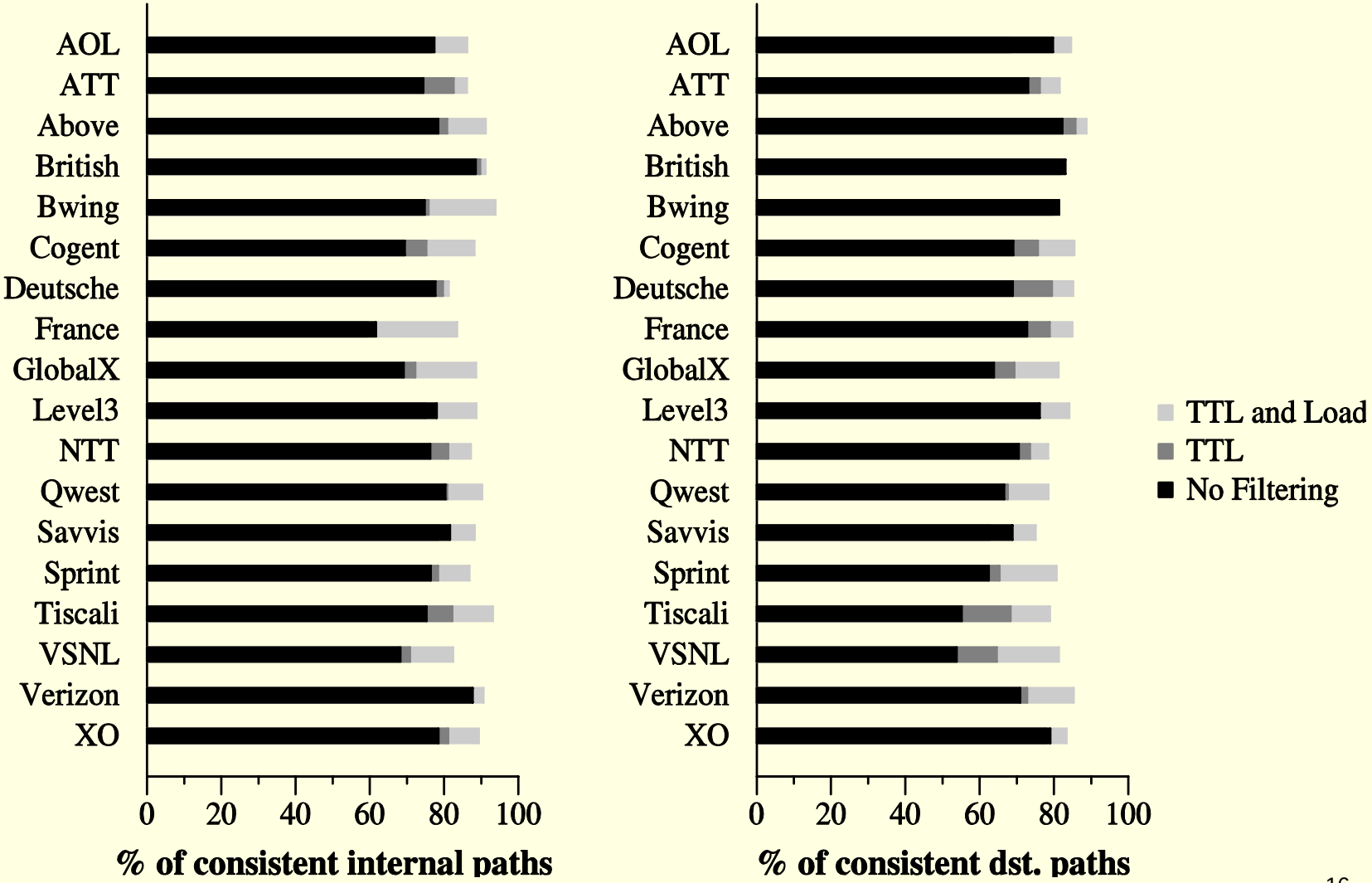


Validation with Keynote

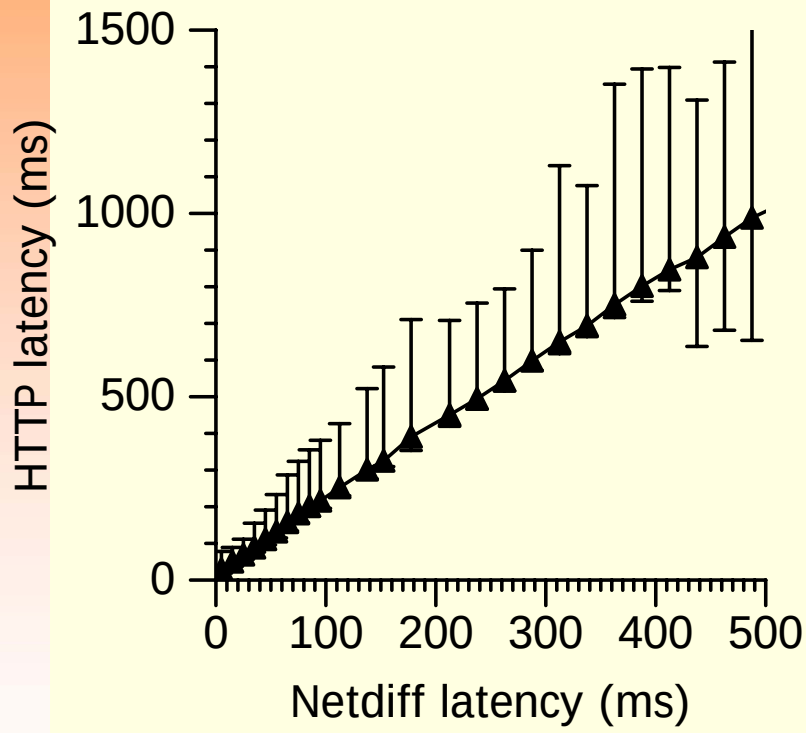
- Keynote double-ended *vs.* Netdiff single-ended

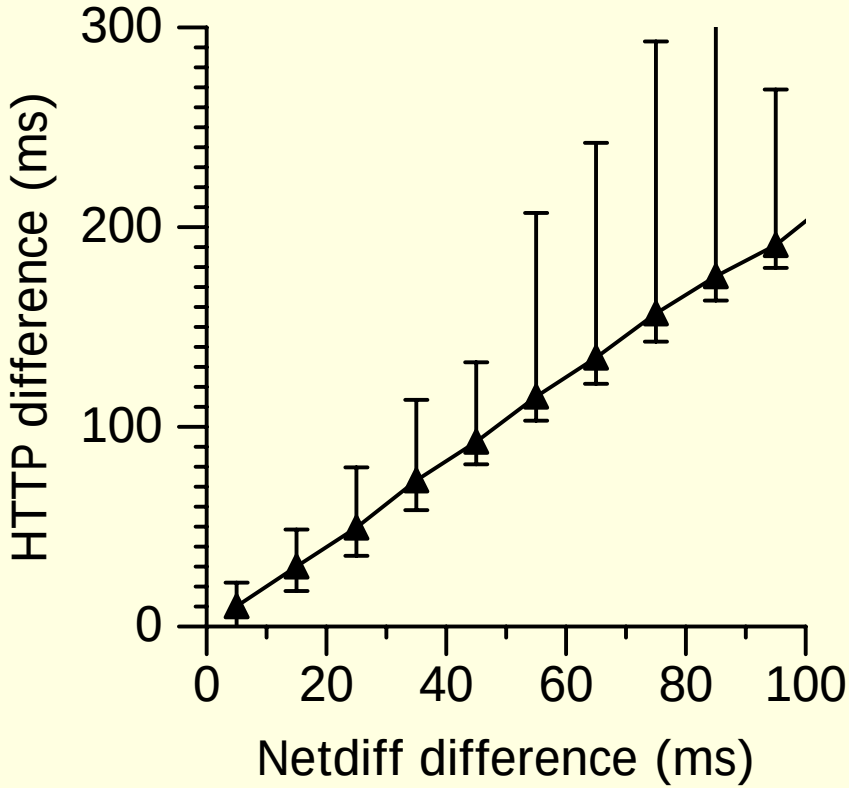


Noise due to path asymmetry/load?



Relevance to application





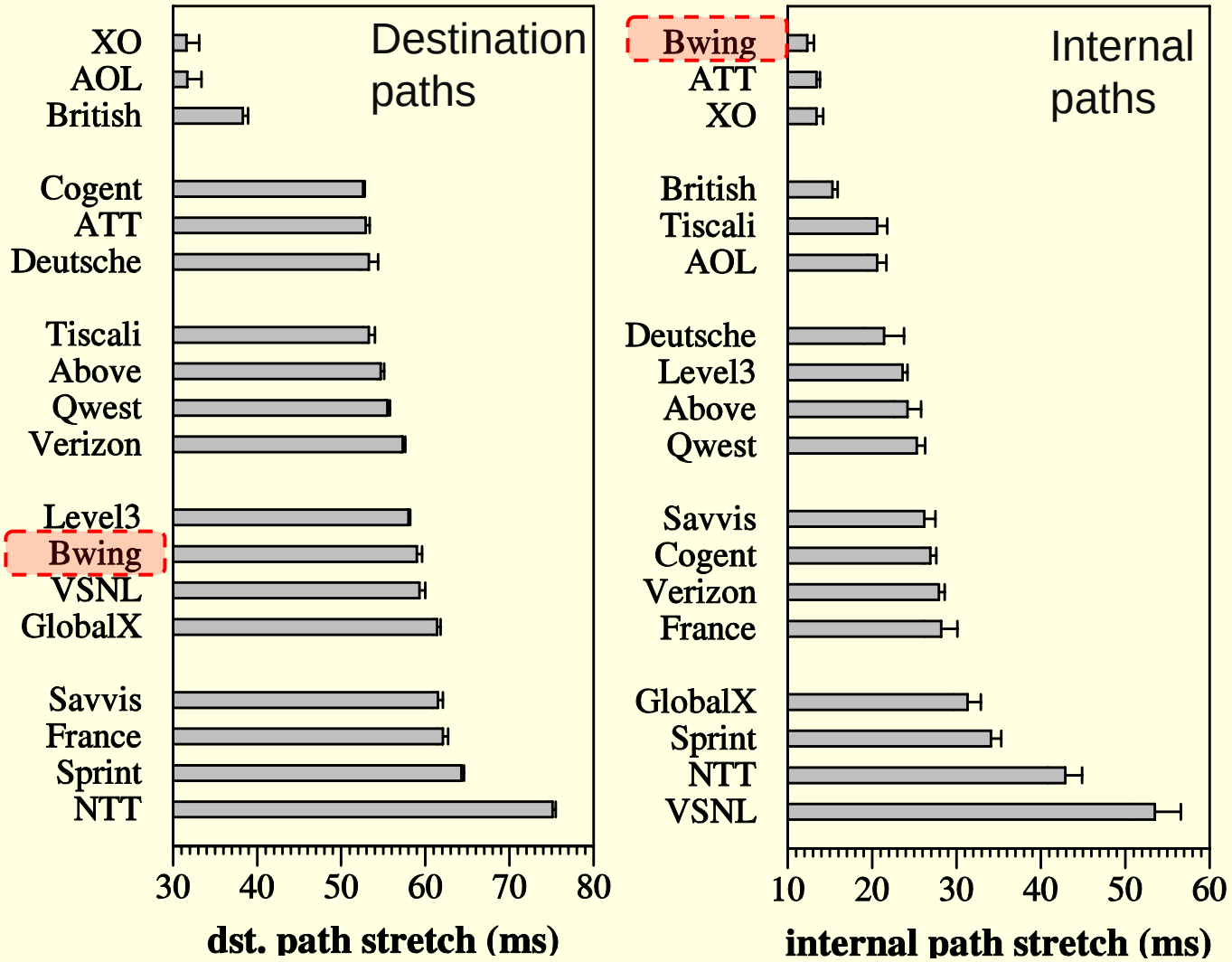
Current status

- PlanetLab: 300+ sites worldwide
- 18 backbone ISPs: AT&T, Sprint, BT…
- Since Feb 2007: 20 minutes per ISP

Comparison methodology

- Path of interest
  - Internal: PoP ⟶ PoP
  - Destination: PoP⟶ destination networks
- Path stretch
  - Measured latency minus optimal latency
  - Account for difference in path length
- Group paths based on length
  - Short: < 20 ms
  - Medium: 20 - 50 ms
  - Long: > 50 ms

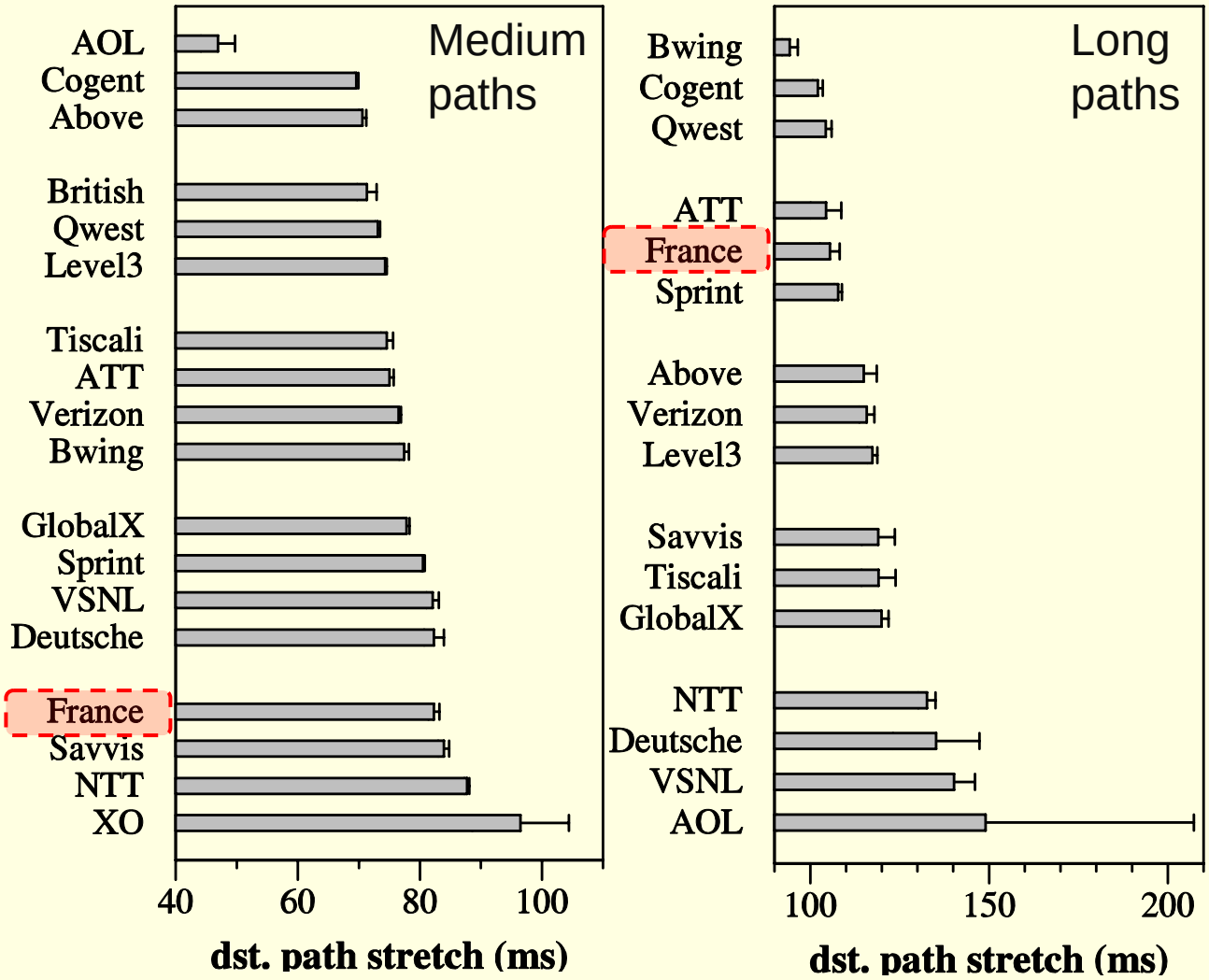
Overall comparison



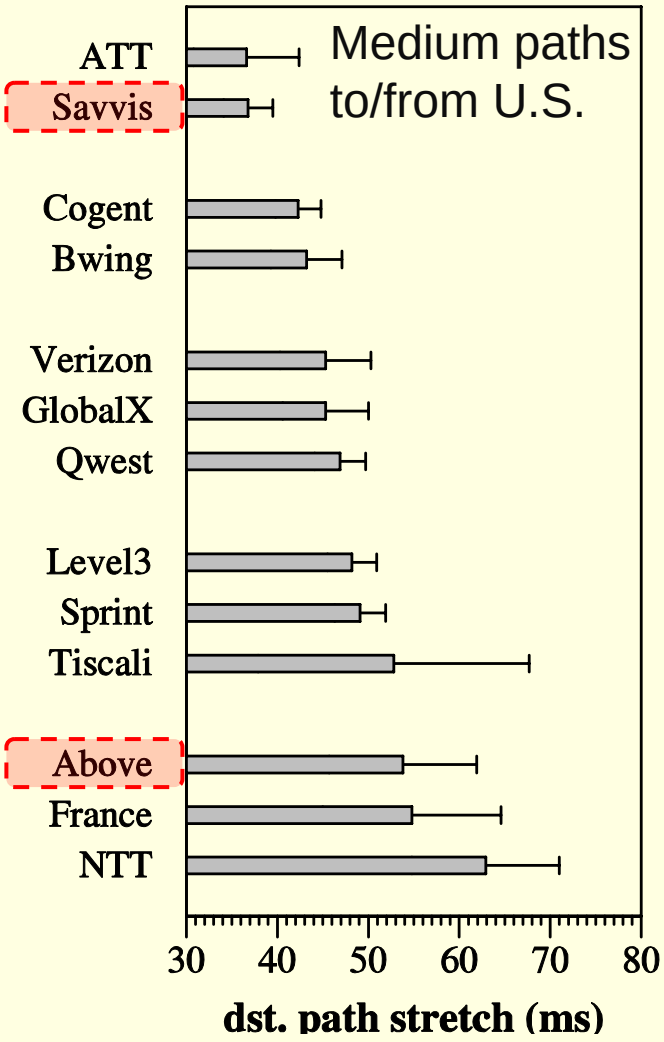
Detailed comparison

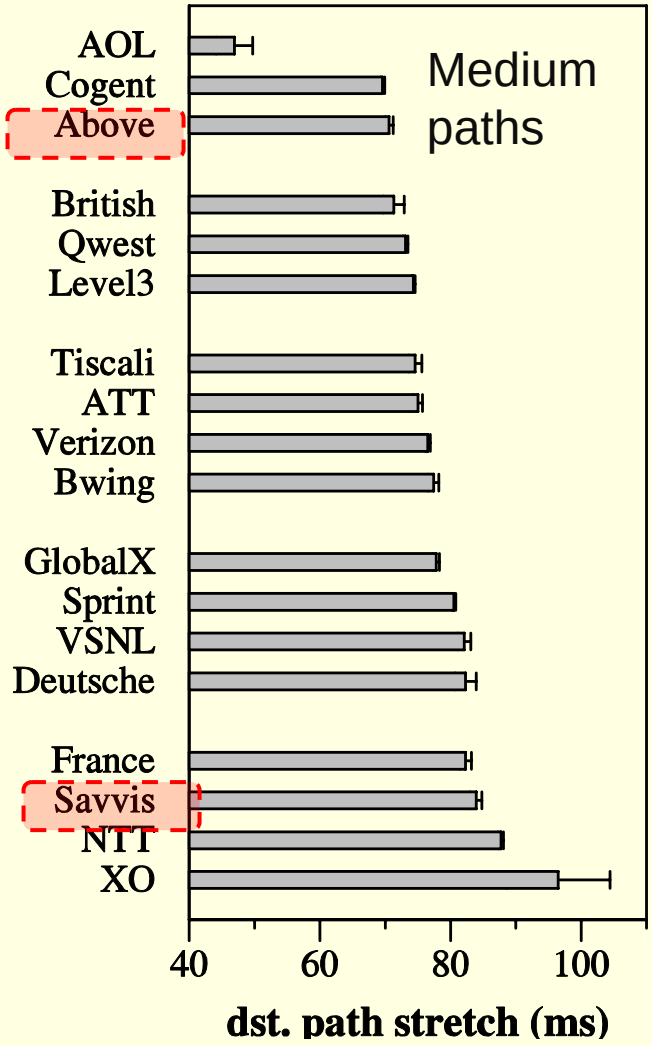
- Which ISP is best depends on workload
  - Distance
  - Geographical region
  - Direction
  - Destination
  - …
- Data available at http://netdiff.org
  - Customize your own comparison

Dependency on distance

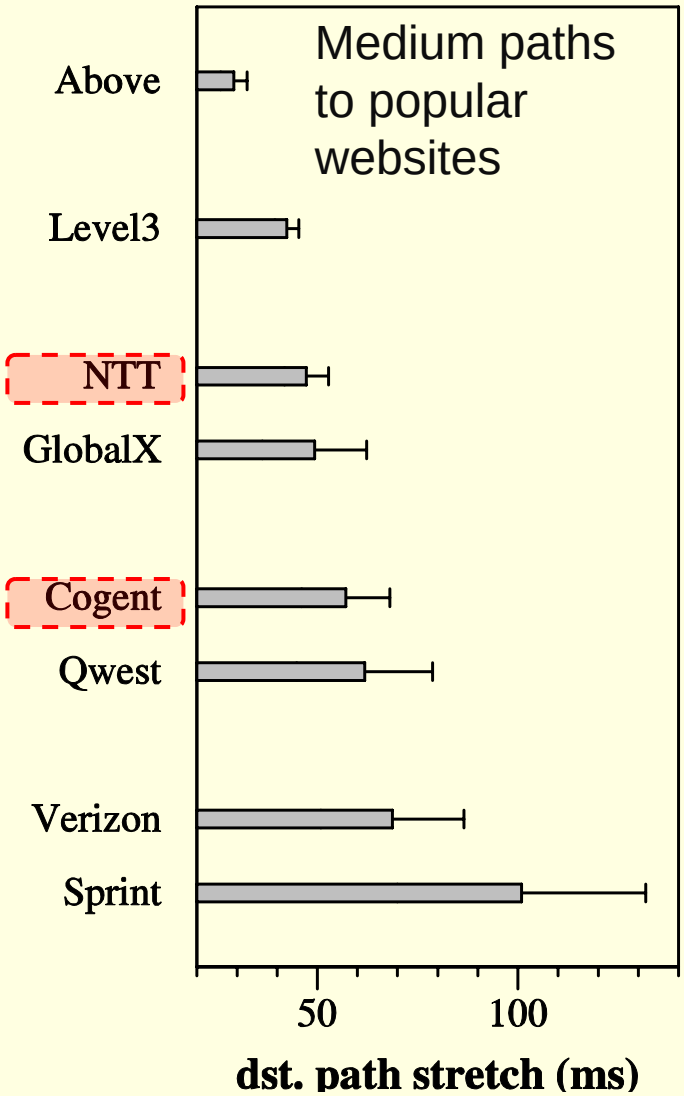


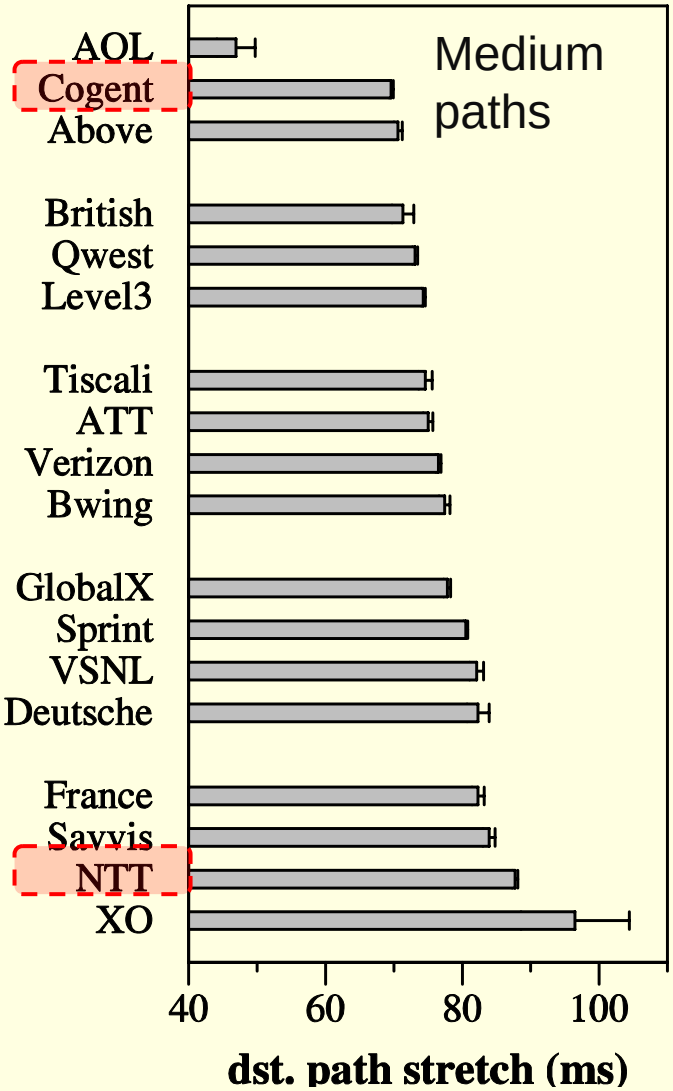
Dependency on geographical region





Dependency on destination





Future directions

- Extend Netdiff to other measures
  - Loss
  - Stability
  - Reachability
  - Network neutrality
- Explain differences in ISP networks

Conclusions

- Netdiff is the first system that provides detailed comparison among backbone ISPs
  - Easy to deploy, requires no ISP cooperation
  - Currently measures 18 backbone ISPs
- ISP ranking depends on workloads
  - Netdiff's detailed comparison is essential
- ISP benchmarking will help customers and improve overall network infrastructure

Dependency on direction