

# A DNS REFLECTION METHOD FOR GLOBAL TRAFFIC MANAGEMENT

JIN LI (MICROSOFT RESEARCH)

06.25.2010

With Cheng Huang, Nick Holt, Y. Angela Wang,  
Albert Greenberg, Keith. W. Ross

# Outline

---

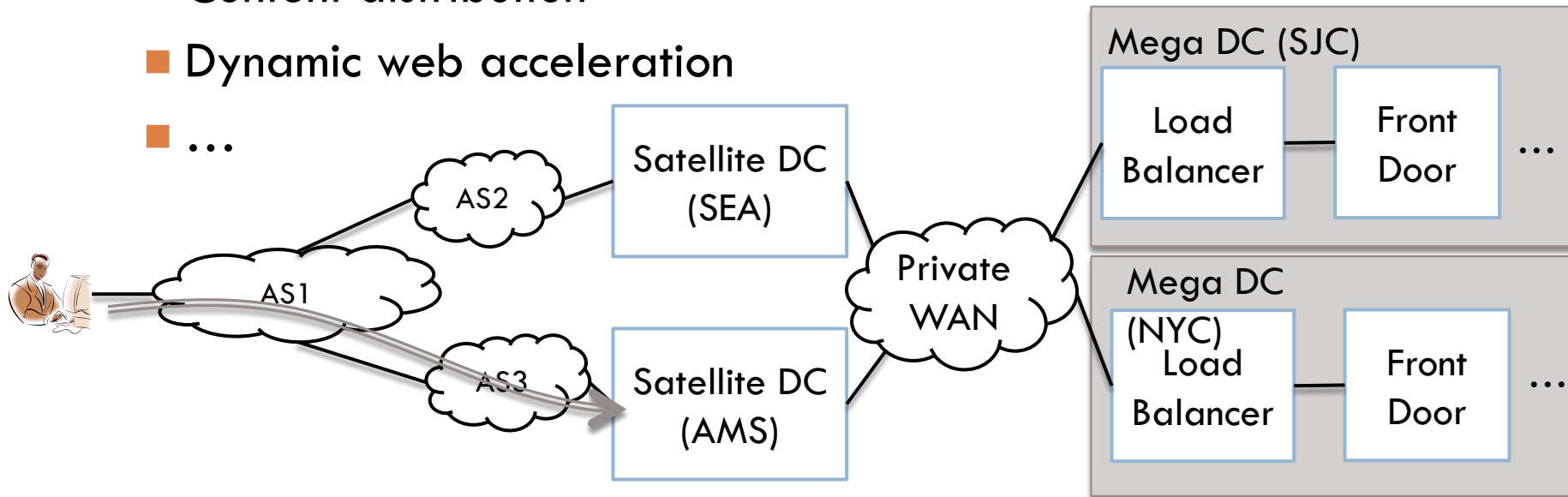
- Introduction
- DNS based GTM
- GTM optimization, LDNS population & reachability
- DNS reflection
- Results
- Conclusion

# Global Traffic Management (GTM)

## □ The GTM problem

▣ For any service running in N satellite data centers, which data center should be selected to serve a particular client to achieve the best (latency and/or throughput) performance?

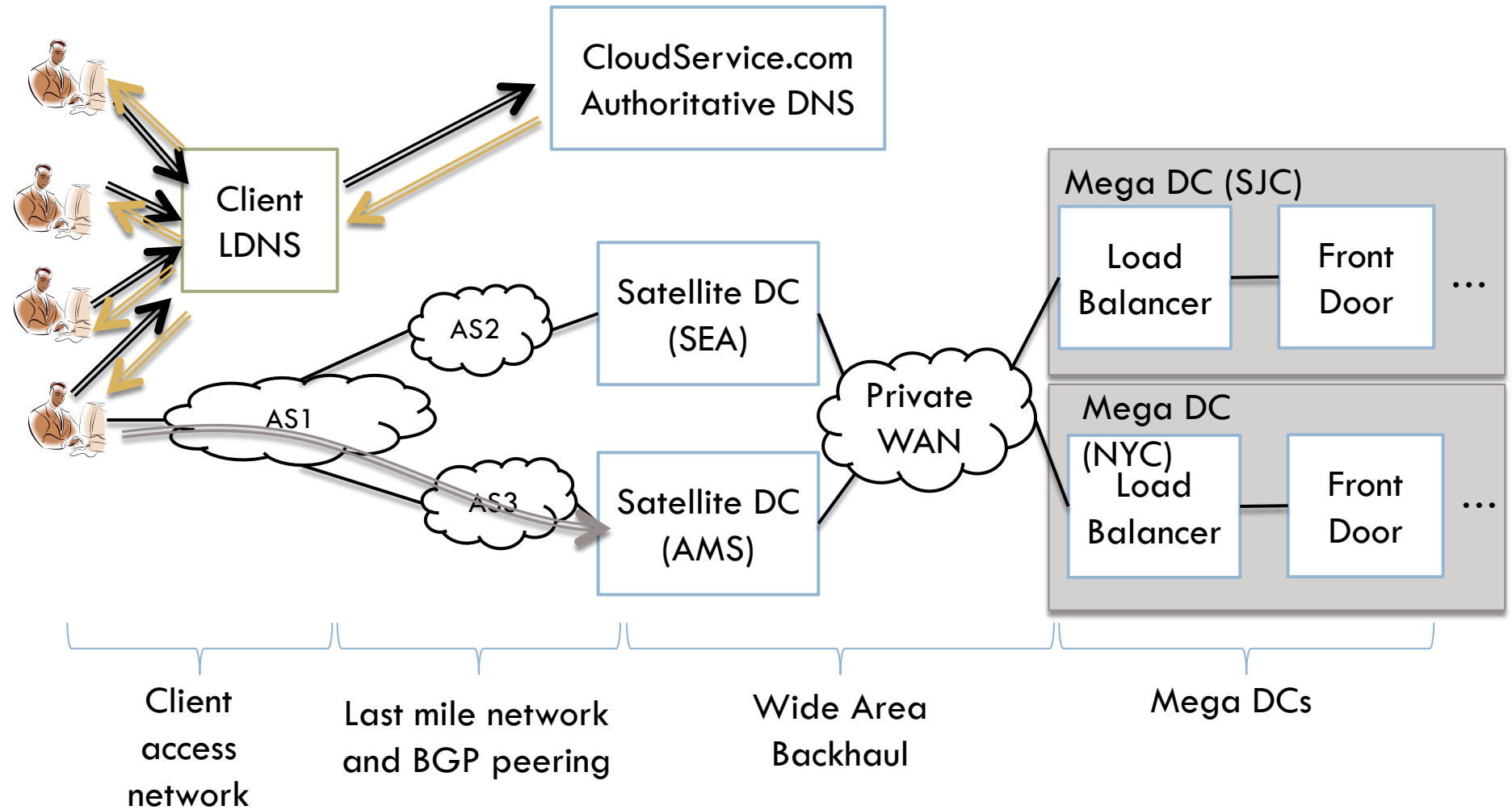
- Cloud storage
- Content distribution
- Dynamic web acceleration
- ...



# Global Traffic Management (GTM)

- Practical GTM solutions: how to redirect?
  - HTTP redirection
  - URL rewriting
  - DNS-based GTM
    - All clients resolve the same hostname (e.g., gtm.CloudService.com)
    - GTM returns the IP of the best DC
      - Based on clients' Local DNS servers (LDNS)
      - GTM never sees clients' IPs
    - Most common ← highly scalable
      - What we deal with in this paper

# DNS based GTM



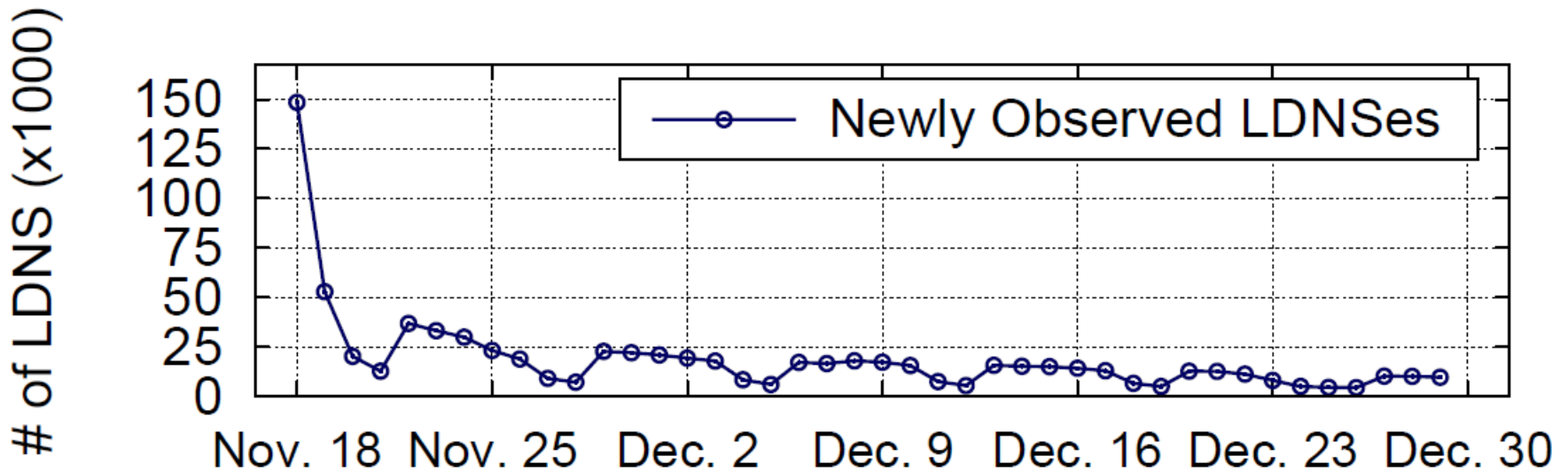
# DNS-based GTM solutions

- Geography-based GTM
  - ▣ Decision based on geographic location
  - ▣ Mapping from location to DC
- Anycast-based GTM
  - ▣ Serving clients from the anycast-closest DC
  - ▣ Anycast (BGP)-closest  $\neq$  latency-closest

# DNS-based GTM solutions

- Passive measurement
  - ▣ Monitor performance between IP prefix and DCs
    - Most clients directed to the best DC
    - Some clients (randomly selected) directed to probe other DCs
    - Traces captured at DCs to infer performance
  - ▣ Major problem
    - Performance of the selected clients is degraded
      - LDNS caching will affect subsequent clients and can be very bad
- Active probing
  - ▣ Most often used by CDN
  - ▣ See next slide for reachability

# Reachability of LDNS



- 6 week's logging of 5% NCSI DNS traffic
  - 795K unique LDNSes in 10,012 cities and 229 countries



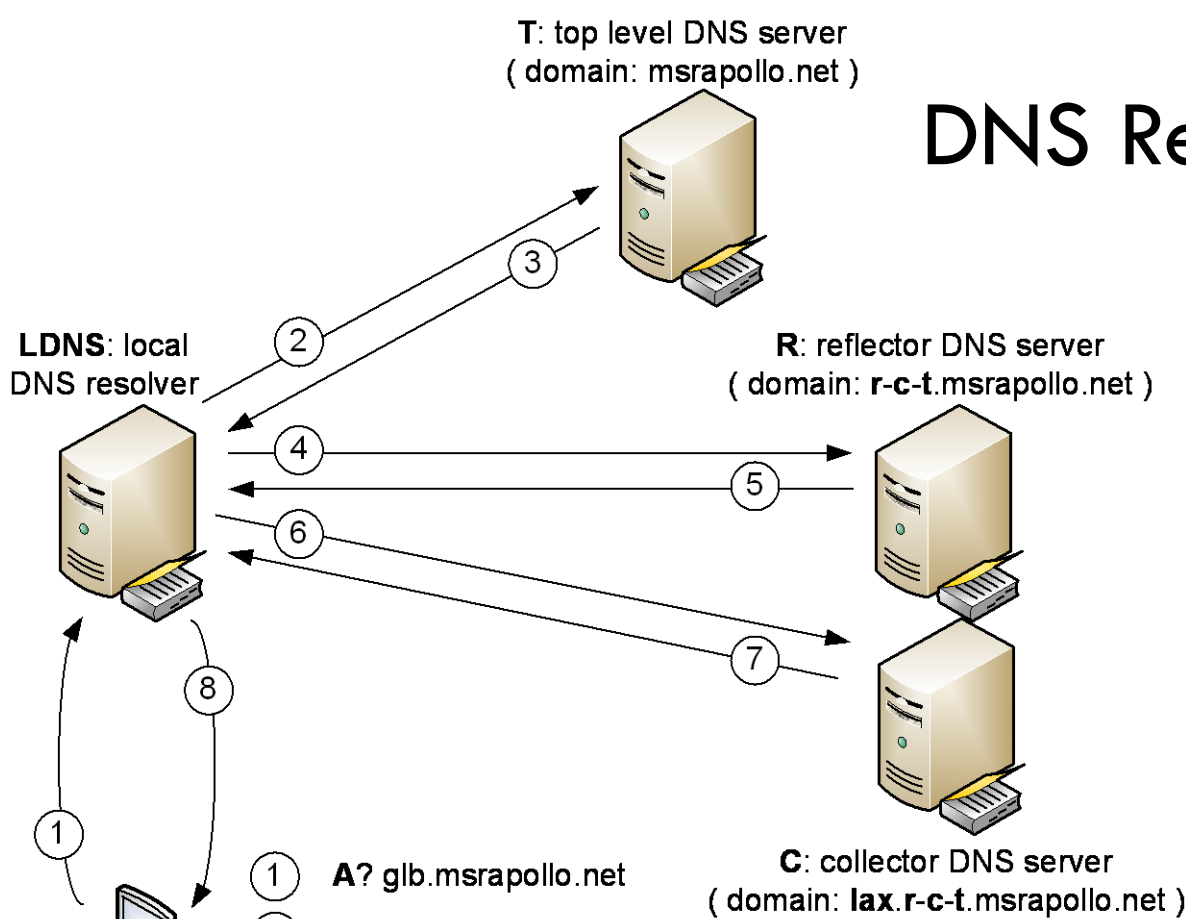
# Reachability of LDNS

- Monitor performance between LDNS and DC
  - Active probing
    - 49% Ping-able
    - another 6% respond to DNS probe queries
  - For the rest 45% – **passive measurement w/ DNS traffic**
    - Trigger DNS query from arbitrary LDNS to measure any target DC through **DNS Reflection**
    - Passive measurement → no LDNS query, no measurement
    - Universal → applicable to any LDNS
    - Minimize performance impact → always serve clients with the optimal DCs
    - Achieve high accuracy

# Our proposal – DNS Reflection

- **GTM using DNS Reflection**
  - ▣ Minimize performance impact
  - ▣ Achieve high accuracy
- How does it work?
- How does the solution fare with existing ones?

# DNS Reflection Method



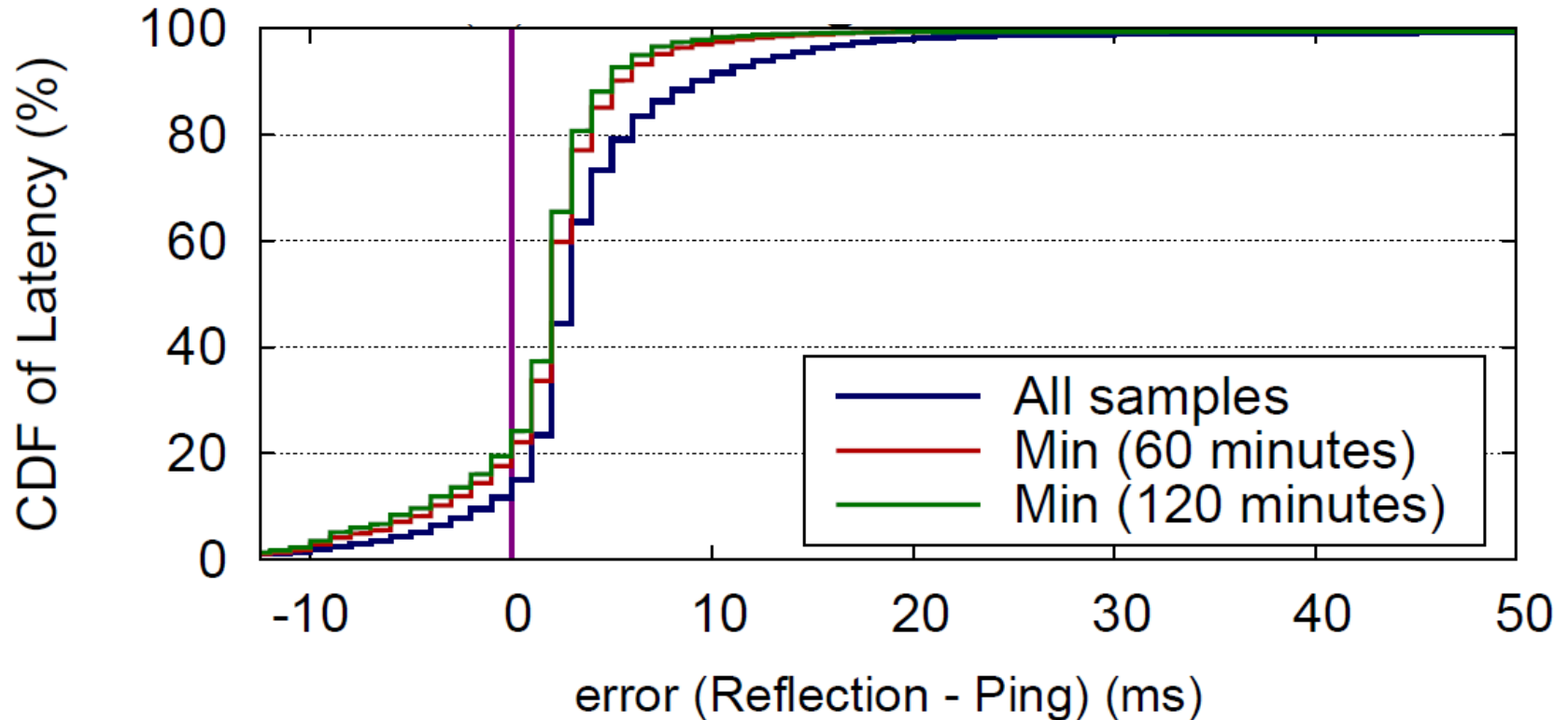
- 1 **A?** glb.msrapollo.net
- 2 same as (1)
- 3 **CNAME:** rand.lax.reflector-collector-target.msrapollo.net  
**NS:** ns.reflector-collector-target.msrapollo.net  
**NS\_ADDRESS:** reflector
- 4 **CNAME:** rand.lax.reflector-collector-target.msrapollo.net
- 5 **NS:** ns.lax.reflector-collector-target.msrapollo.net  
**NS\_ADDRESS:** collector
- 6 same as (4)
- 7 **NS\_ADDRESS:** target
- 8 same as (7)

□ The key is to trigger a LDNS query the same DC twice

# Evaluation

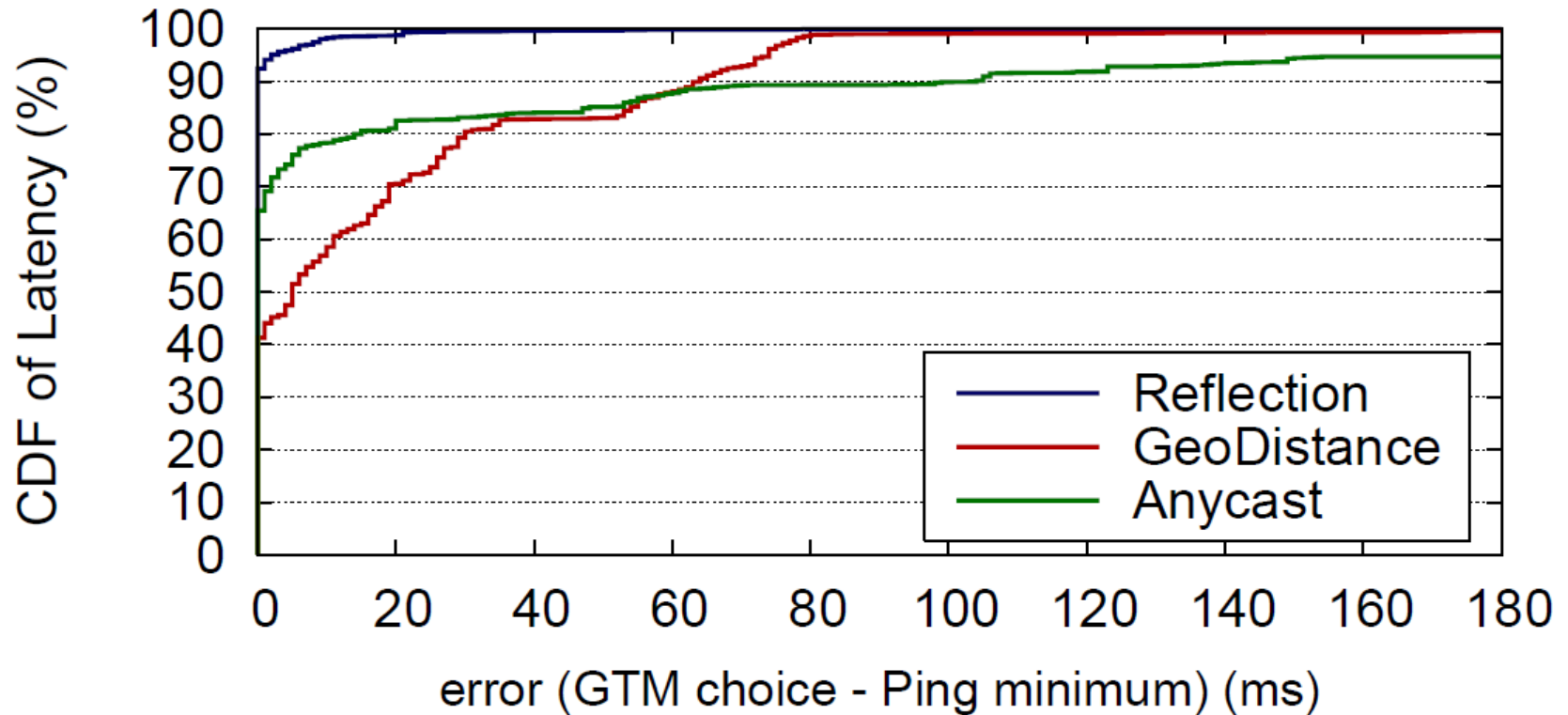
- How accurate is the measurement? How good is reflection based GTM fare with geography & anycast based GTM?
- Prototype deployed on 17 DCs in the Microsoft global data center network
- 162 (out of 274) PlanetLab nodes
  - ▣ LDNS co-locates with client (240)
  - ▣ LDNS responds to Ping (162)

# Accuracy – DNS Reflection vs. Ping



- DNS reflection matches very well with Ping
  - 6 ms away from Ping

# GTM Performance



- GTM using DNS Reflection is very close to optimal
  - 2 ms away from optimal vs. 74 ms (geography) and 183 ms (anycast)

# Conclusion

- DNS-based GTM is most commonly used
- Active probing suffers from limited reachability
- Passive measurement by redirecting clients to sub-optimal DCs degrade performance and affect subsequent clients
- DNS Reflection method
  - ▣ Cause a minimal performance impact
  - ▣ Achieve high accuracy