Understanding BGP Next-hop Diversity

Jaeyoung Choi[1], Jong Han Park[2], Pei-chun Cheng[2], Dorian Kim[3], Lixia Zhang[2]

[1] Seoul National University, jychoi@mmlab.snu.ac.kr
[2] UCLA, {jpark,pccheng,lixia}@cs.ucla.edu
[3] NTT Communications Inc., dorian@blackrose.org

Global Internet Symposium 2011

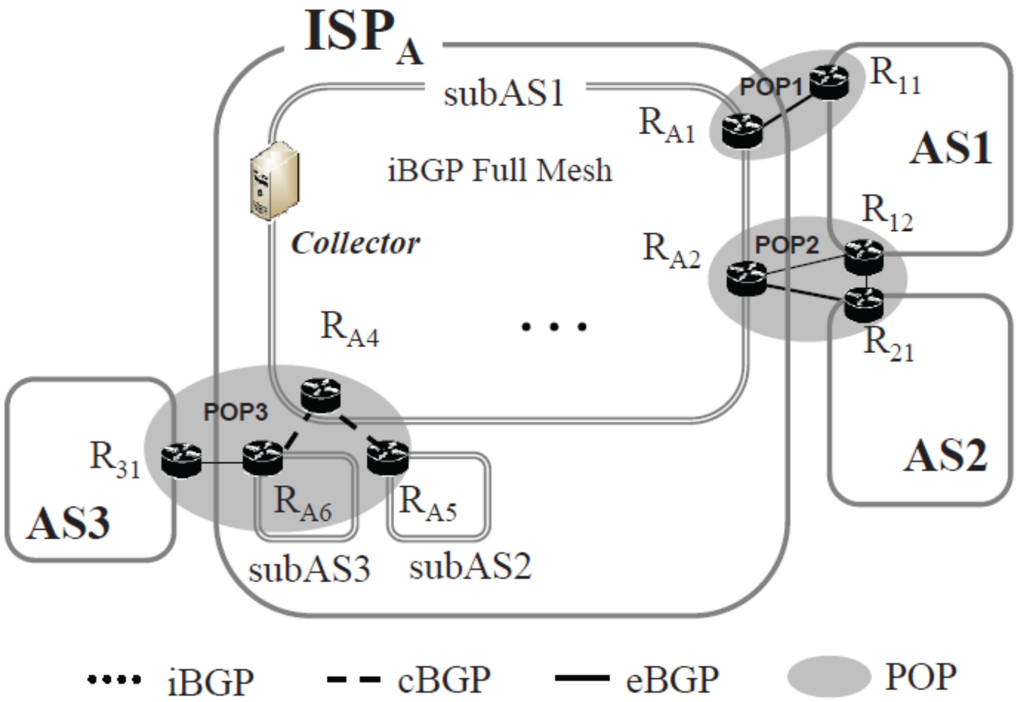
Why this work

- High path diversity is important
  - Increase network robustness for failures
  - Increase flexibility in traffic engineering and load balancing
  - Decrease convergence time when topology changes

- IETF efforts to increase path diversity
  - WG Diverse-path, Add-Path, Best-External
  - Proposes several ways to modify BGP to support multiple paths

- What is the *existing path diversity* in the operation networks? How does it change over time?
  - Are the modifications necessary?
  - What is the effective way to modify BGP to support multiple paths?

High level description of measurement ISP



- ISP_A (Tier-1 ISP in the Internet) using AS confederations
  - Backbone sub-AS with more than one hundred i-BGP routers in a full-mesh
    - Spreads across 14 countries and 3 continents
    - Most prefixes are announced directly to one of the routers in the backbone sub-AS
- A collector is placed in the backbone sub-AS to passively collect i-BGP data

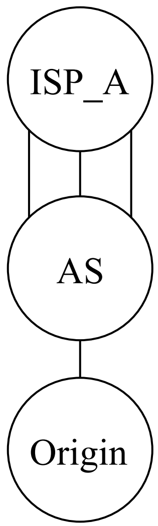
Quantifying next-hop diversity

- Why next-hop diversity?
  - ISP's concern on path diversity is confined to next-hop diversity within their network
    s

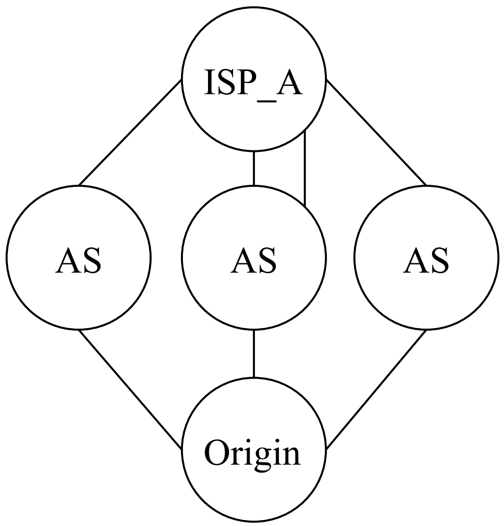
- Dataset
  - Routing table snapshots during one month of July 2009
  - Filter out internal prefixes and potential bogon prefixes
    - Filter out prefixes with their length smaller than 8 or greater than 24

Defining next-hop diversity

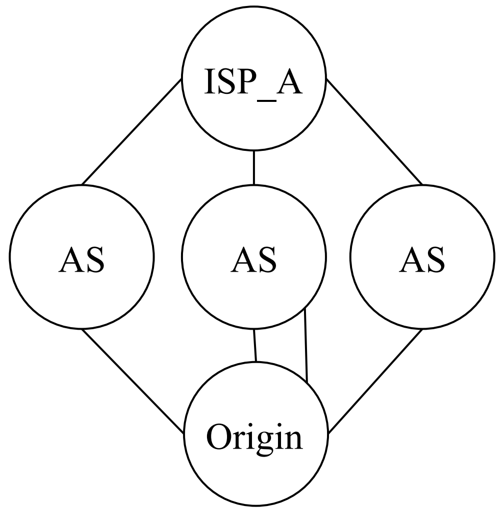
- Next-hop {AS,POP,router} diversity
  - The number of unique next-hop *routers* along with their *geographical locations (i.e. city)* and next-hop *ASes* for a prefix



- Next-hop router Div.: 3
- Next-hop AS diversity: 1

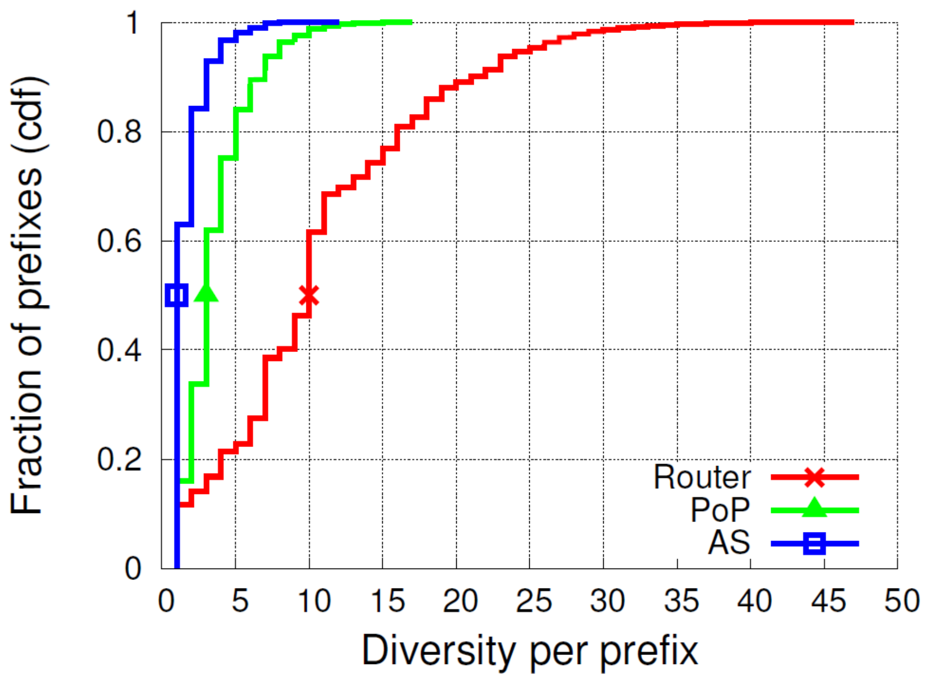


- Next-hop router Div.: 4
- Next-hop AS diversity : 3



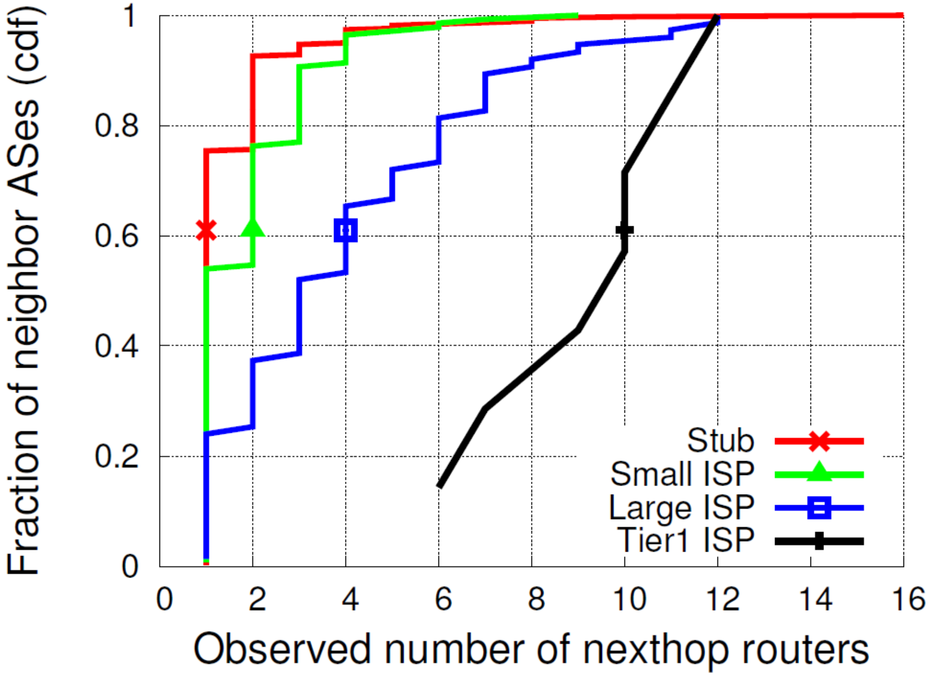
- Next-hop router Div.: 3
- Next-hop AS diversity : **3, not 4**

Existing next-hop diversity of ISP_A



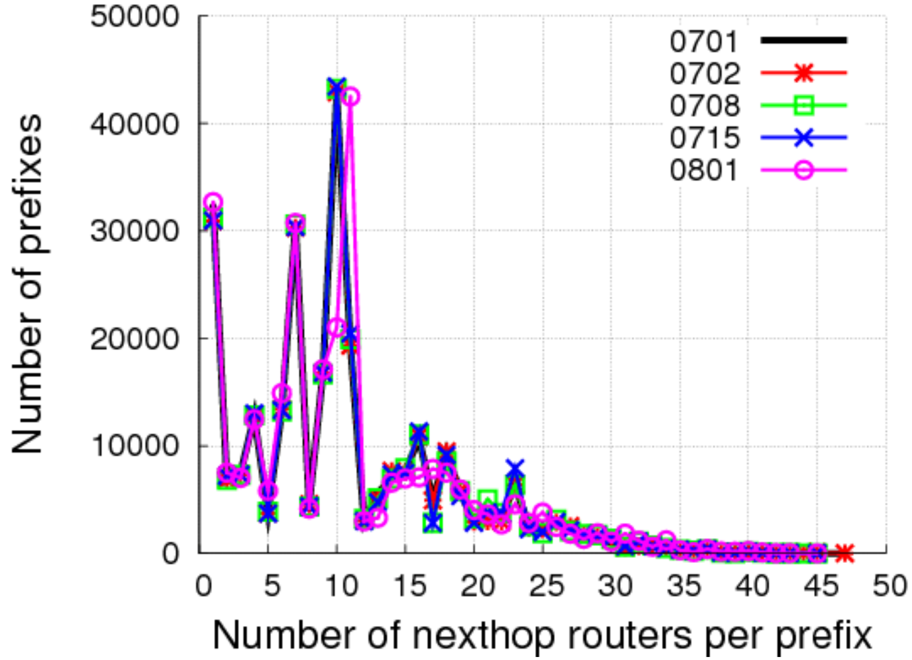
- Based on RIB on July 1st 2009 (276,712 prefixes in total)
  - Majority of prefixes (more than 11% and 18%) can be reached via more than 2 next-hop r outers and POPs
  - More than 60% of prefixes can be reach via only one next-hop AS
  - A small number of prefixes have a very high degree of next-hop router

Neighbor AS type and diversity



- 4 types of AS: (1) stub, (2) small ISP, (3) large ISP, and (4) Tier-1
- In general, ISPs with larger size tend to have more peering sessions across different routers and POPs
  - ISP_A and other Tier1s have 6 to 12 next-hop routers
  - ISP_A and large ISPs have 1 to 12 next-hop routers
- Stub AS with high connectivity (ex: UltraDNS, Amazon, Akamai)

Is our observation representative?



- Additional measurements performed on RIBS from 4 different dates
  - The number of next-hop routers are very similar across all measurements
  - Additional analysis on other results confirm our previous observation

What are the impacting factors of next-hop diversity?

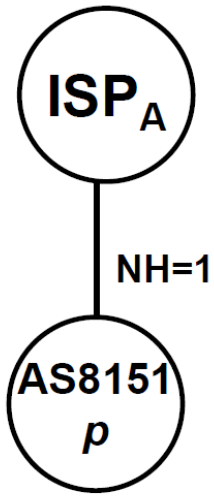
- Impacting factor analysis through case studies

- Major factors impacting next-hop diversity in ISP_A
  - ISP’s path preference (policy)
  - Number of peering routers with its neighbor ASes
  - Lack of geographical presence

Low diversity

- Two explanations
  - (1) Only one path exists
  - (2) BGP selects and propagate only the best path and hides the rest

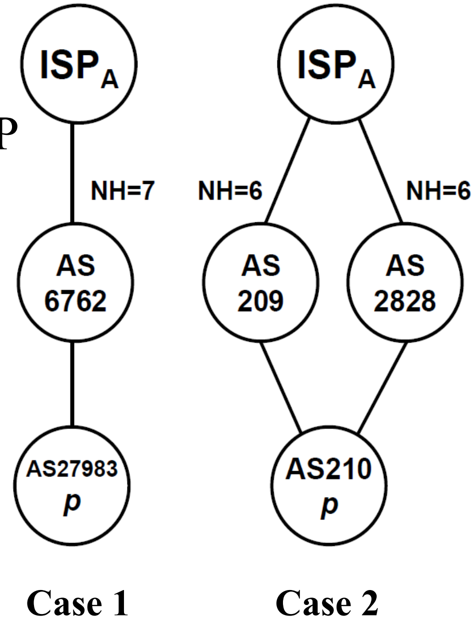
- Our further investigation confirms the latter
  - For most of prefixes, multiple paths do exist based on Cyclops (http://cyclops.cs.ucla.edu/)
  - Network operator may be able to increase their diversity by adjusti ng tunable parameters of BGP (ex: weight, local-pref)



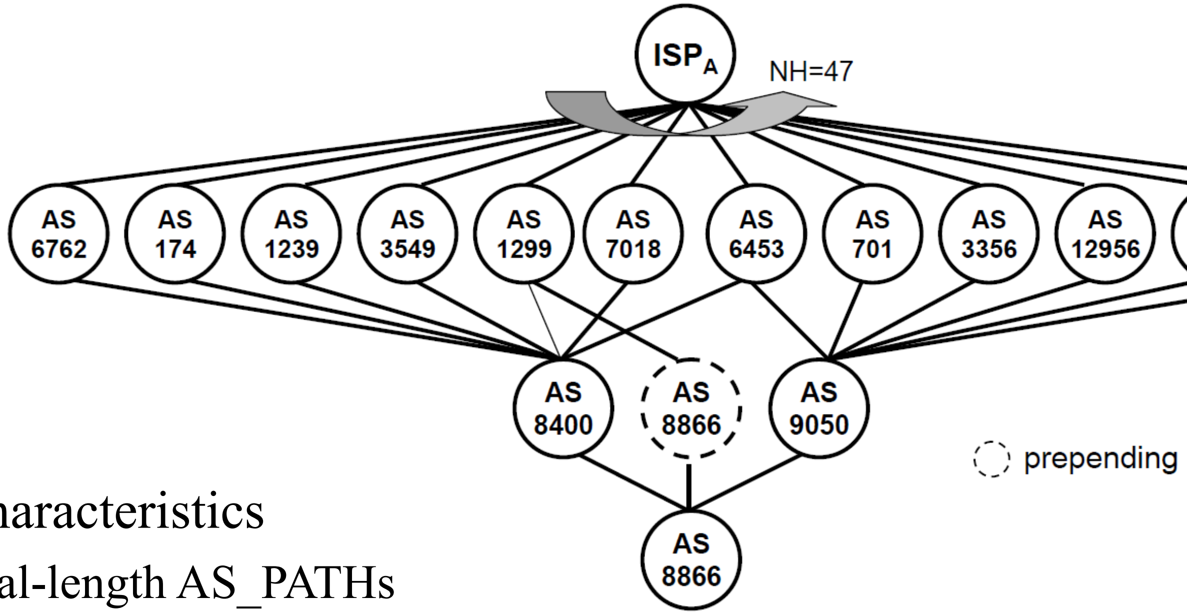
Moderate diversity

- Prefixes whose next-hop router diversity is between 6 and 12
  - Applies to more than half of all prefixes
  - Prefixes are reached through an AS that maintains multiple BGP peering sessions with ISP_A

- This case shows us that
  - # of peering routers has an impact on the next-hop diversity



High diversity



- Interesting characteristics
  - Many equal-length AS_PATHs
  - Length of AS_PATH to reach ISP_A is always > 1

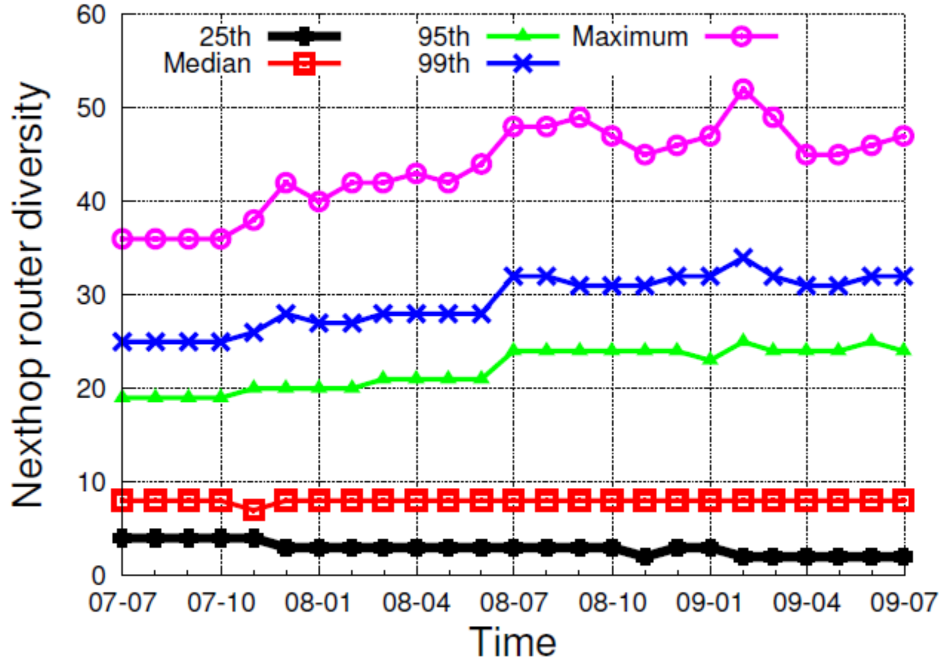
- Lack of geographical presence of ISP_A
  - For 89% of prefixes with high diversity, ISP_A do not have a presence at the prefix origination POP
  - Some prefixes can have a very high diversity regardless of the ISP's intention

Does diversity change in time?

- What is a general trend of next-hop diversity changes over time?

- Dataset
  - Sampled the routing table snapshot taken on 1st day of each month from July 20 07 to July 2009
  - Only consider common prefixes that exist in all RIBS
    - 220,432 prefixes in total

Next-hop diversity changes over 2 years



- Median values stay almost the same
  - The diversity of individual prefixes change in unpredictable manner, compensating the changes of other prefixes
  - As a result, no noticeable aggregate change in time
- 95th, 99th, Maximum values slightly increase
  - Number of backbone routers inside ISP_A have increased up to 19 additional routers during the 2 years

Summary

- Despite the promising efforts to increase path diversity, little understanding on path diversity in the *existing* system

- Our quantification and analysis on Tier-1 iBGP routing data show
  - Majority of prefixes can be reached through multiple next-hop routers
  - Some of the high diversity is unintended
  - ISP may be able to increase their diversity without any BGP modifications by a djusting path preference and number of peering routers
  - Overall diversity has not changed a lot in time

Any questions?

Thank you.