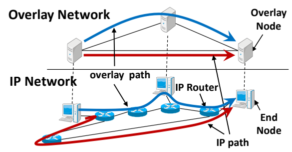
Performance evaluation of a method to reduce inter-ISP transit cost caused by overlay routing

Kazuhito MATSUDA †, Go HASEGAWA †, Satoshi KAMEI ††, Masayuki MURATA †

† Osaka University, Japan
†† NTT Service Integration Laboratories, Japan

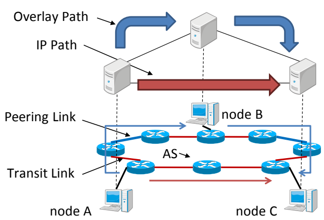
Background: Overlay routing

- A routing mechanism provided by overlay network
- Improves end-to-end network performance
  - Using user-perceived routing metrics such as end-to-end latency and available bandwidth
  - Performance gain is mainly based on the policy mismatch between IP routing and overlay routing
    - IP routing is based on router-level hop count, AS-level hop count and commercial contracts with neighboring ISPs
    - Overlay routing is based on user-perceived metrics



Problem definition: Increasing transit cost

- With overlay routing, there is a possibility of using additional transit links
  - This increases transit cost in the whole network



- Normal IP routing utilize two transit links
- When an overlay routing selects the path via node B, it traverses four transit links

Transit cost increases

Our goal

Propose a method to reduce inter-ISP transit cost caused by overlay routing

- Estimate the number of transit links with end-to-end network performance values which can be measured easily by overlay nodes
  - Because there are no public information of transit/peering relationships between ASes, we use multiple regression analysis of these values
- Select an overlay-level route with the estimated number of transit links as the routing metric
  - While reducing transit costs, the proposed method should maintain the performance improvement provided by overlay routing

Proposed method: Estimation of the number of transit links

- Evaluate correlations between "true" number of transit links and networks performance values which can be measured easily by overlay nodes
  - End-to-end latency, router-level hop counts, available bandwidth, etc.

⇨ Select parameters for a multiple regression analysis according to the degree of correlations

- Apply a multiple regression analysis to the selected parameters and derive the regression equation to estimate the number of transit links

Coefficients for each parameter

Network performance value between nodes *i* and *j*

$$C^e_{ij} = b_0 + b_1 x^1_{ij} + b_2 x^2_{ij} + \cdots + b_n x^n_{ij}$$

Estimated number of transit links between node *i* and *j*

Proposed method: Limited overlay routing

- Path selection methods of overlay routing
  1. The overlay routing selects the path which have the best end-to-end perf… …ber of transit links…
  2. The … …ber of trans… …erceived perf…

…nsit links …een node *i* and *j*

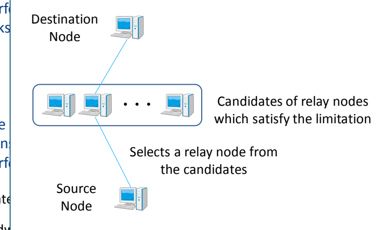
…ks on the relay … *i* and *j* via *k*

End-to-end late…

Available bandw… $P_{ikj} \geq P_{ij} \times (1 - \rho)$

…nce of the best relay path between node *i* and *j*

P_{ikj} : the performance of the relay path between node *i* and *j* via *k*



Proposed method: Limited overlay routing

- Path selection methods of overlay routing
  1. The overlay routing selects the path which have the best end-to-end performance under the limitation on the increase in the number of transit links on the relay paths than that of the direct path

$$C_{ikj} \leq C_{ij} + \alpha$$

C_{ij} : the number of transit links on the direct path between node *i* and *j*

C_{ikj} : the number of transit links on the relay path between node *i* and *j* via *k*

  2. The overlay routing selects the path which have smallest number of transit links under the limitation on the decrease in the user-perceived performance compared with the best relay path

End-to-end latency $P_{ikj} \leq P_{ij}^{opt} \times (1+\beta)$

Available bandwidth $P_{ikj} \geq P_{ij}^{opt} \times (1-\beta)$

P_{ij}^{opt} : the performance of the best relay path between node *i* and *j*

P_{ikj} : the performance of the relay path between node *i* and *j* via *k*

Evaluation environments

1. Generalized PlanetLab environment
   - Assume the overlay network constructed by PlanetLab nodes
   - To evaluate the performance of the proposed method in the general host distribution on the Internet, we use the "generalized PlanetLab environment" which has the node distribution according to the ratio of AS numbers assigned to each region

RIR (region name)	Number of ASes	Number of selected nodes from PlanetLab
ARIN (North America)	24,422	50
RIPE NCC (Europe)	21,065	43
APNIC (Asia)	5,782	12
LACNIC (South America)	2,815	6

Selects nodes randomly from each region

2. Japanese commercial network environment
   - Assume the overlay network constructed by nodes located at Japanese commercial ISPs
     - 18 nodes in 13 Japanese commercial ISPs

Dataset (1)

- Generalized PlanetLab environment
  - End-to-end latencies and available bandwidths between nodes
    - *Scalable Sensing Service* (S^3) [1]
      - Measurement results between PlanetLab nodes, which are summarized every four hours, are available
  - IP-level and AS-level paths between nodes
    - Conduct *traceroute* commands to obtain IP-level paths and convert these results into AS-level paths with AS number and IP address prefix database on *Route Views Project*
  - True number of transit links
    - We obtained the information of transit/peering relationships between ASes from CAIDA
    - We consider the number of transit links from these dataset as "true" number of transit links on the path

[1] Hewlett-Packard Laboratories, "Scalable Sensing Service," available at http://networking.hpl.hp.com/s-cube/.

Dataset (2)

- Japanese commercial network environment
  - End-to-end latencies between nodes
    - Conduct *ping* commands
  - IP-level and AS-level paths between nodes, true number of transit links
    - The same way as for the generalized PlanetLab environment

Evaluation results: The regression equations for two network environments

- Derive the regression equations for the generalized PlanetLab environment and the Japanese commercial network environment
  - Select "router-level hop count" and "end-to-end latency" as the parameters for the multiple regression equation

$$C_{ij}^{e} = b_y + b_r h_{ij} + b_d \delta_{ij}$$

C_{ij}^{e} : the estimated number of transit links between node *i* and *j*

h_{ij} : the router-level hop count between node *i* and *j*

δ_{ij} : the end-to-end latency between node *i* and *j*

	b_r	b_d	b_y
generalized PlanetLab environment	0.145 (7.56×10^{-4})	0.00120 (1.08×10^{-6})	0.846 (0.20)
Japanese commercial network environment	0.240	-0.000889	-1.48

average values of twenty node selections

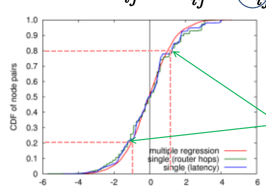
variance values of twenty node selections

Evaluation results: Estimation accuracy (generalized PlanetLab environment)

- Estimation accuracy of the regression equation with the difference between the estimated number and the true number of transit links

$$d_{ij} = C_{ij}^{e} - C_{ij}^{t}$$

C_{ij}^{t} : the true number of transit links between node *i* and *j*

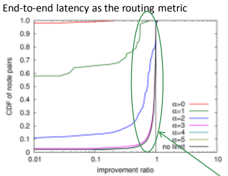


The multiple regression equation has high accuracy compared with other single regressions

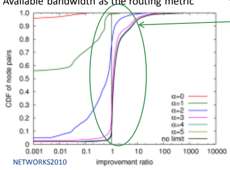
The regression equation can estimate where absolute estimation error is less than 1.0 in 60 % of node pairs

Evaluation results: User-perceived performance when limiting the number of transit links (generalized PlanetLab environment)

End-to-end latency as the routing metric



Available bandwidth as the routing metric



- The overlay routing selects the path which can achieve the best performance among all relay paths satisfying the limitation on the number of transit links
  - Equal to or less than direct path's transit links plus α

End-to-end latency $\text{Improvement} = P_{\text{direct}} / P_{\text{bestrelay}}$

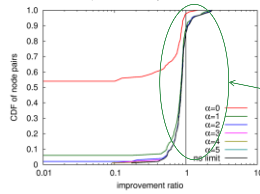
Available bandwidth $\text{Improvement} = P_{\text{bestrelay}} / P_{\text{direct}}$

When α is equal to or larger than three, the overlay routing can get almost the same improvement degree as that without the limitation

Our method can achieve the same improvement degree as that without the limitation, while limiting the increase in the estimated number of transit links equal to or less than three

Evaluation results: User-perceived performance when limiting the number of transit links (Japanese commercial network environment)

End-to-end latency as the routing metric



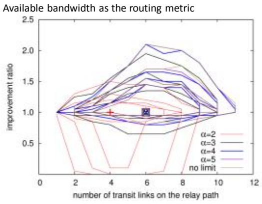
- Evaluation with the Japanese commercial network environment

When α is equal to or larger than one, the overlay routing can get the same improvement degree as that without the limitation

Our method can achieve the same improvement degree as that without the limitation, while limiting the increase in the estimated number of transit links equal to or less than one

Evaluation results: Trade-off between user-perceived performance and the reduction in the number of transit inks (generalized PlanetLab environment)

Available bandwidth as the routing metric



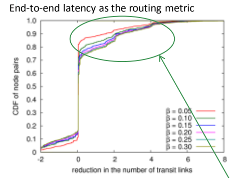
- Plot the distribution of node pairs on improvement ratio vs. number of transit links
  - The lines represent the density contours of 50, 100, 150 and 200 node pairs for each value of α

- The reduction in the number of transit links is largest when α is two, but the improvement ratio becomes less than one for a number of node pairs, it means many nodes can not get any performance improvement
- When α is greater than four, there is almost no reduction in the number of transit links

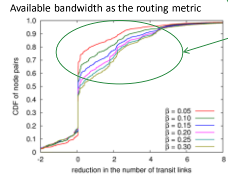
We can conclude that α = 3 is the best value from the viewpoint of the trade-off between user-perceived performance and reduction in the number of transit links

Evaluation results: Reduction of transit links when limiting the decrease on user-perceived performance (generalized PlanetLab environment)

End-to-end latency as the routing metric



Available bandwidth as the routing metric



- The overlay routing selects the path which has the smallest number of transit links satisfying the limitation on the user-perceived performance
  - Equal to or less than β compared with that of the best performance among all relay paths

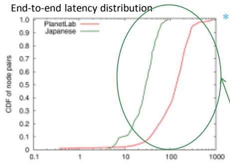
$$\text{Reduction} = C_{\text{bestrelay}} - C_{\text{smallestrelay}}$$

- When β = 0.05, we can reduce the number of transit links by at least one in 11% of node pairs in the case of end-to-end latency, 18% in the case of available bandwidth
- Although we can achieve a greater degree of reduction in the number of transit links, the decrease in the user-perceived performance becomes significant

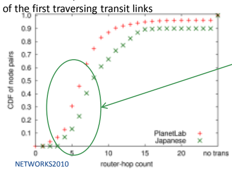
We can reduce a certain degree of the number of transit links, allowing only a 5% decrease in the user-perceived performance

Evaluation results: Network properties affecting the regression equations

End-to-end latency distribution



Router-level hop count of the first traversing transit links



- Discuss the network properties affecting the regression equation

	b_r	b_d	b_y
generalized PlanetLab environment	0.145	0.00120	0.846
Japanese commercial network environment	0.240	-0.000889	-1.48

The transit link estimation by end-to-end latency is difficult, because the latencies in the Japanese commercial network are significantly smaller than those in the generalized PlanetLab environment

Large portion of the node pairs in the generalized PlanetLab environment traverse the first transit links within the first few hops, while those in the Japanese commercial network environment dose not traverse within the first few hops
⇨ this is a reason why the intercept b_y in both environments is different

The proposed method could obtain the regression equation appropriately for both network environments

Conclusions & Future works

- Conclusions
  - Propose the method reducing inter-ISP transit cost in overlay routing
    - Utilize the estimated number of transit links on the relay paths calculated by the regression equation
  - Reveal the advantages of the proposed method in various network environments
    - We can control the number of transit links on the relay paths, while maintaining the performance provided by the overlay routing
    - Confirm that the proposed method can obtain the regression equation appropriately according to the network environments
- Future works
  - Consider a more practical billing mechanism of transit links
  - Utilize different mechanism such as P4P to reduce inter-ISP transit cost