

オーバレイルーティングに起因するネットワークただ乗り問題に関する一検討

長谷川剛^{1,3} 小林正好² 村田正幸³ 村瀬勉²

¹大阪大学サイバーメディアセンター

²NECシステムプラットフォーム研究所

³大阪大学大学院情報科学研究科

目次

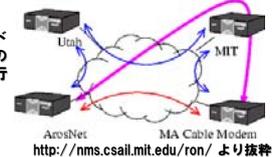
- 研究の背景と目的
- オーバレイルーティングによるネットワークただ乗り問題
- ただ乗りトラフィック量の定式化
- 数値例
- まとめと今後の課題

研究の背景 (1)

- オーバレイネットワーク技術によるネットワークサービスの多様化
 - CDN, Grid, ALM (Application-Level Multicast), VPN, Anonymous file sharing, SAN
 - サービスの早期展開が可能
- 下位層ネットワークに対する高い品質要求
 - アプリケーション (オーバレイノード) 自身が品質を監視・計測し、アプリケーション制御を行う
 - 遅延時間・遅延ジッタなどに基づくストリーミング制御
 - 過去の転送スループットに基づくサーバ選択や並列TCPコネクション数設定
 - ノード間の利用可能帯域や遅延などに基づくレプリカミラー配置

研究の背景 (2)

- トラフィックのルーティングそのものを目的としたオーバレイネットワークの登場：オーバレイルーティング
 - 特定アプリケーションを前提としない
 - ユーザの感じる性能を改善
- 例：Resilient Overlay Network (RON)
 - ノード間の直接経路と、他ノードを経由した経路を比較し、品質の良い経路を用いてデータ転送を行う



[3] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris. "Resilient overlay networks," in *Proceedings of 18th ACM Symposium on Operating Systems Principles*, Oct. 2001.

研究の背景 (3)

- オーバレイルーティングの効果
 - エンド間遅延時間の短縮
 - 40%のノードペア間で可能 (米国、[*])
 - 20%のノードペア間で可能 (日本、[**])
 - 特に、最大遅延時間の改善効果が大きい
 - 障害からの回復
 - BGPでは回復に数分~数時間かかる場合がある
 - 収束する場合、障害地点からのホップ数に依存
 - なかなか収束しない場合もある
 - フルメッシュリンクを張るオーバレイルーティングの場合 (例：RON)、固定時間 (例：30秒以内) で回復可能

[*] 中尾彰宏, "Network Embedded Test-beds for New Generation Network Research," オーバレイネットワークシンポジウム講演, 2006年12月

[**] M. Uchida, S. Kamel and R. Kawahara, "Performance Evaluation of QoS-aware Routing in Overlay Network," in *Proceedings of ICIN 2006*, January 2006.

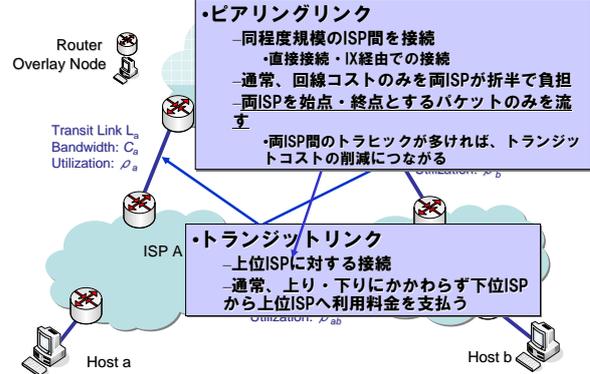
研究の背景 (4)

- オーバレイルーティング
 - ユーザ性能を指標として行われる
- IPルーティング
 - ユーザ性能を考慮して行われていない
 - ホップ数、遅延時間がベース
 - リンク利用率、エンド間パスの空き帯域などは使われない
 - ISP間の接続・BGPルーティングはネットワーク性能よりも、政治的・経済的な理由で決定される
- オーバレイルーティングとIPルーティングの違いにより、ISPの思惑 (コスト構造) を無視したトラフィックが発生する恐れ

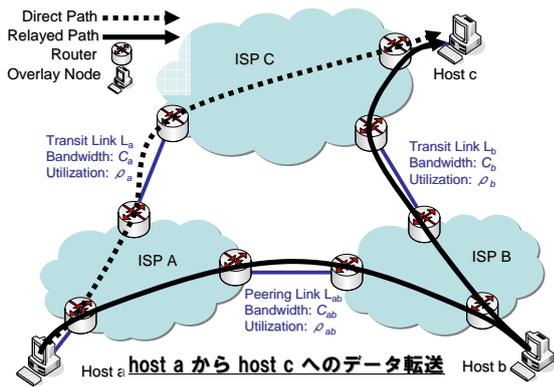
研究の目的

- オーバレイルーティングによるネットワークただ乗り問題に着目
 - 問題定義
 - ただ乗りトラフィック量の定式化
 - 遅延、利用可能帯域を指標として利用
- PlanetLab上の計測データを用いた数値例
 - オーバレイルーティングの効果
 - ただ乗りトラフィック量の推定
- ただ乗り問題がISPにとって無視できない問題になり得ることを指摘

ネットワークモデル (1)



ネットワークモデル (2)

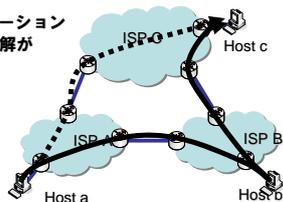


オーバレイトラフィックのコスト負担

- Direct pathを使う場合
 - ISP A 内の host a が、ISP A のトランジットリンクを使ってトラフィックを運ぶ
 - host a から課金可能
- Relayed pathを使う場合
 - ISP A 内の host a が、ISP A-B 間のピアリングリンクと、ISP B のトランジットリンクを使ってトラフィックを運ぶ
 - ピアリングリンク: ISP A, B で折半
 - ISP B のトランジットリンク: ISP B が負担
 - ネットワークただ乗り問題
 - ISP A 内のホストが発生させたトラフィックのトランジットコストを、直接的には関係のない ISP B が負担
 - ピアリングリンク上のトラフィック量に偏りが発生すると、コスト負担を要求されることもある

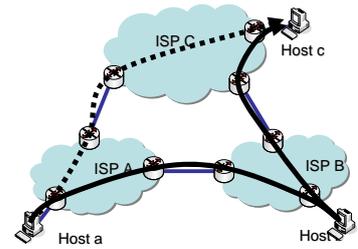
考えられる解決方法とその問題点 (1)

- オーバレイネットワークに参加している host b から課金する
 - ユーザは自分のPCが中継に使われていることを自覚していない
- アプリケーションそのものを遮断・制限
 - Winny であれば可能
 - Skype は?
 - 「有用」なアプリケーションの場合、ユーザの理解が得られない



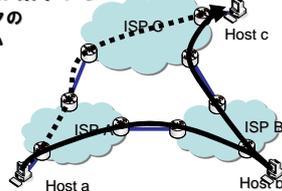
考えられる解決方法とその問題点 (2)

- Relayed pathを使うトラフィックを検出し、課金/遮断する
 - 通常のトラフィック (ISP A → ISP B) と区別できない
 - どちらも src: host a, dst: host b のパケット



考えられる解決方法とその問題点 (3)

- ISP Bが設備投資を行い、トランジットリンク帯域を増強
 - オーバーレイトラフィックのうち、中継経路 (Relayed Path) が使われる割合が増える
 - ISP A内のユーザが恩恵を受ける
 - ISP B内のユーザには思ったような効果が得られない
- 品質の良いネットワークを構築した結果、そこへオーバーレイトラフィックが集中する
 - そのコストをトラフィックの発生元から回収できない



13

Relayed pathを使うトラフィック量 (1)

- 利用可能帯域の比で経路を選択する場合
 - ISP Bがトランジットリンク帯域を増強することによって、流れ込んでくるオーバーレイトラフィック量

$$\Delta x_r = \frac{A_a(C'_b - C_b)}{(A_a + C'_a(1 - \rho'_a))(A_a + A_b)} x \quad \lim_{C'_b \rightarrow \infty} \Delta x_r = \frac{A_a}{A_a + A_b} x$$

- 帯域増強が進むと、オーバーレイトラフィックのほとんど全てが中継経路へ流れる
- 利用可能帯域の大きい経路を選択する場合

$$\Delta x_r = \frac{1}{2}(C'_b - C_b)$$

- 増強した帯域の50%を、オーバーレイトラフィックの中継に使われる

14

Relayed pathを使うトラフィック量 (2)

- 遅延時間 (RTT) の小さい経路を選択する場合
 - 各経路のRTT

$$RTT_{d,r} = \tau_{d,r} + \frac{1}{\mu_{d,r}(1 - \rho_{d,r})}$$

- 経路を流れるトラフィック量が増えると、RTTが増大する
- RTTが等しくなる時の、それぞれの経路の帯域利用率の差

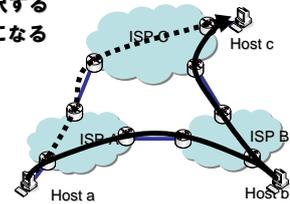
$$\Delta \rho = \frac{\Delta \tau \cdot \mu(1 - \rho_d)^2}{1 - \Delta \tau \cdot \mu(1 - \rho_d)}$$

- ネットワーク全体の負荷が高いと、中継経路がより多く使われる

15

Relayed pathを使うトラフィック量 (3)

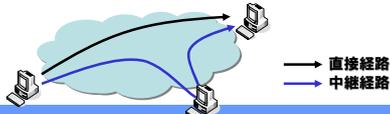
- ボトルネックが両経路が通過する場所 (例: host cに近い場所) にある場合
 - オーバーレイトラフィック量に関係なく、使われる経路は固定的になる
 - RTTで経路を選択する場合、常に全てのトラフィックが中継経路を流れる場合もある
 - 利用可能帯域を選択する場合は、50%ずつになる



16

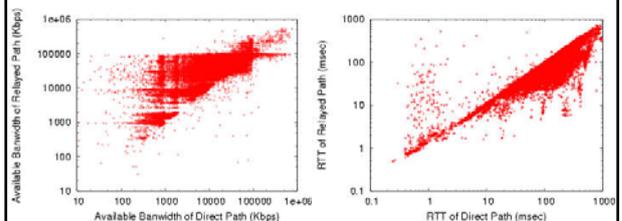
PlanetLabのデータを用いた数値例 (1)

- PlanetLab
 - <http://www.planet-lab.org/>
 - 世界規模の実験・サービス展開のためのオーバーレイネットワーク
 - 約700ノード
- Scalable Sensing Serviceプロジェクト
 - <http://networking.hpl.hp.com/s-cube/>
 - PlanetLabノード間のフルメッシュ計測情報を公開
 - 遅延時間、利用可能帯域、物理帯域、パケット廃棄率など
 - 4時間ごとの計測結果がダウンロード可能
- 各ノードペアに関して、下記を比較
 - 直接経路の遅延・利用可能帯域
 - どこか他のノードを経由した場合の遅延・利用可能帯域



17

PlanetLabのデータを用いた数値例 (2)



- 横軸：直接経路の性能、縦軸：最良の中継経路の性能
 - 利用可能帯域の場合で96.2%、RTTの場合で96.3%の経路で、良い中継経路が存在
 - 全ての中継経路の平均と比較すると、それぞれ25.4%/0.05%
- 利用可能帯域を指標にする場合、オーバレイルーティングがより有効

18

PlanetLabのデータを用いた数値例 (3)

- オーバレイネットワーク全体で、中継経路で運ばれるトラフィック量の割合

	利用可能帯域の比で経路を選択	利用可能帯域の大きい経路を選択
最良の中継経路	72.8%	96.2%
良い中継経路の平均	58.5%	96.2%
全中継経路の平均	49.2%	22.6%

- 中継経路=ただ乗り経路、ではない
 - トランジットリンクを1つでも経由される場合
 - ・ ただ乗りになる可能性がある
 - 流入・流出ともピアリングリンクの場合
 - ・ ただ乗りにはならないが、トラフィックを運ぶコストが発生

19

まとめと今後の課題

- オーバレイルーティングによるネットワークただ乗り問題
 - 問題提起
 - ISPにとって無視できない問題になり得る
- 今後の課題
 - より詳細な調査
 - ・ オーバレイトラフィック量が増えた場合の影響
 - ・ 物理帯域・トポロジーを考慮した検討
 - ・ PlanetLab以外のデータを用いた検討
 - ただ乗りトラフィックの検出手法

20