
多段リンクにおける輻輳制御 アルゴリズムの公平性分析

立命館大学

原田香子 上山憲昭

目次

1. インターネットの輻輳制御
2. 輻輳制御アルゴリズム
3. 着目課題
4. 本研究の目的
5. 評価方法
6. 評価条件
7. バッファサイズ
8. 評価結果
9. まとめ

インターネットの輻輳制御

- 輻輳制御アルゴリズム (CCA: Congestion Control Algorithms)
 - ネットワーク混雑による遅延・損失を防ぐため送信レートを調整するアルゴリズム
- OS・アプリケーションは既定のCCAを利用
 - 例) Linux OS: CUBIC-TCP
- ネットワーク上には様々なCCAが共存
 - 代表例: TCP-Reno, CUBIC-TCP, BBRなど
- 複数のフローが同一リンクを経由する際, そのリンクの伝送帯域を奪い合う

輻輳制御アルゴリズム(1)

■ 損失ベースCCA

- パケット損失を輻輳の兆候とみなし, 送信量を減少
- TCP Reno: 初期の標準的なCCA
- CUBIC-TCP: 現在の標準的なCCA

■ 遅延ベースCCA

- 遅延の増加を輻輳の兆候とみなし, 送信量を減少
- BBR : Google開発のCCA。

輻輳制御アルゴリズム(2)

■ TCP-Reno:

- 低帯域・低遅延ネットワーク
- ウィンドウサイズは線形増加
- フロー間の公平性を実現

■ CUBIC-TCP:

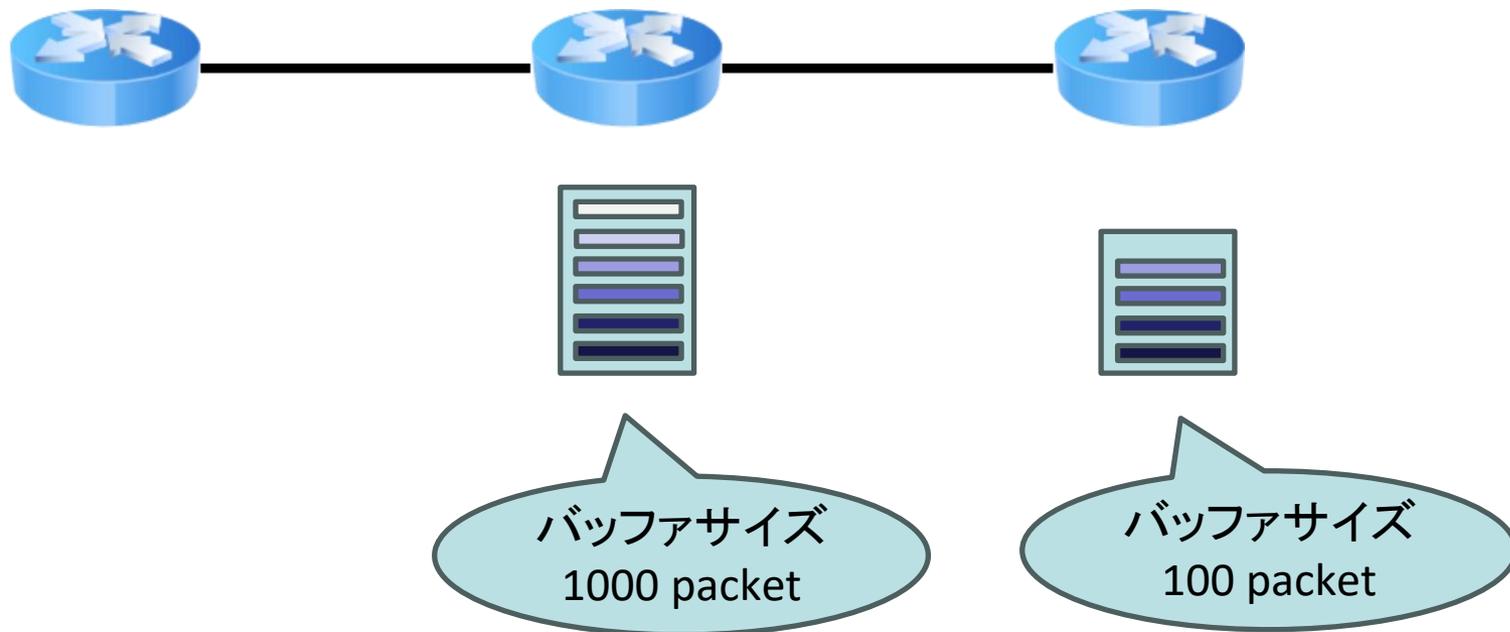
- 三次関数によるウィンドウサイズの増加
- 現在のスタンダードな輻輳制御法
- 広帯域・高遅延のネットワーク

■ BBR:

- 帯域幅と最小RTTを推定し、それに基づいて最適な送信レートを設定

出カルータのバッファサイズ

- 出カルータの最大データ量
- 1BDPの計算方法
 - $BDP = \text{帯域幅}[bps] \times \text{遅延}[\text{秒}]$



着目課題

■ 主な知見

- BBRがスループットを約80%占有し, 他のCCAのスループットが著しく低下(バッファサイズ小)
- CUBICがスループットを約80%占有し, BBRのスループットが著しく低下(バッファサイズ大)
- CCAの種類により帯域配分の不公平が生じる可能性

■ 単一ボトルネックリンクを考慮した評価*1, 2

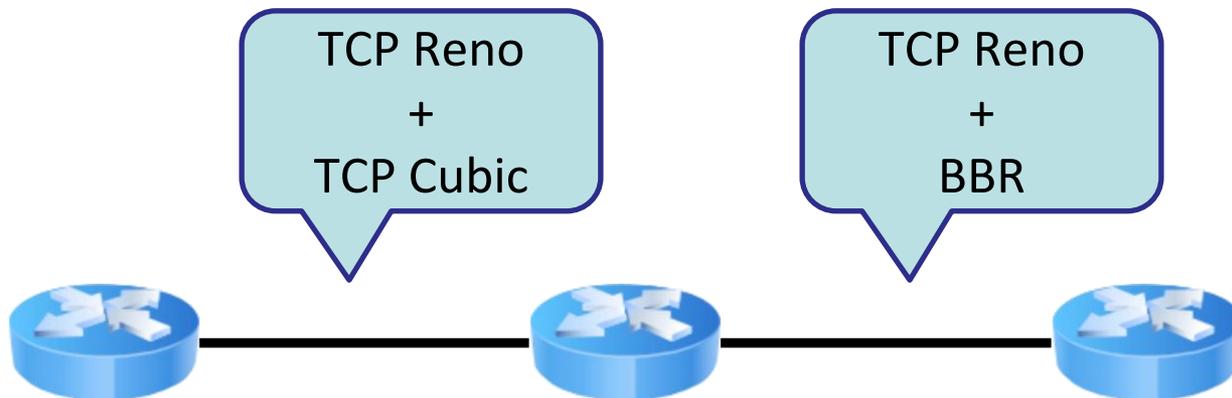
- 実際のインターネットは多数のリンクを経由
- 経路の途中で競合フローが変化
- 複数リンクを経由する場合, 公平性の結果が異なる可能性あり

*A. Philip, et al., "Prudentia: Findings of an Internet Fairness Watchdog", ACM SIGCOMM 2024

*R. Ware, M. K. Mukerjee, S. Seshan, and J. Sherry, "Modeling BBR 's Interactions with Loss-Based Congestion Control", the ACM Internet Measurement Conference, (IMC 2019)

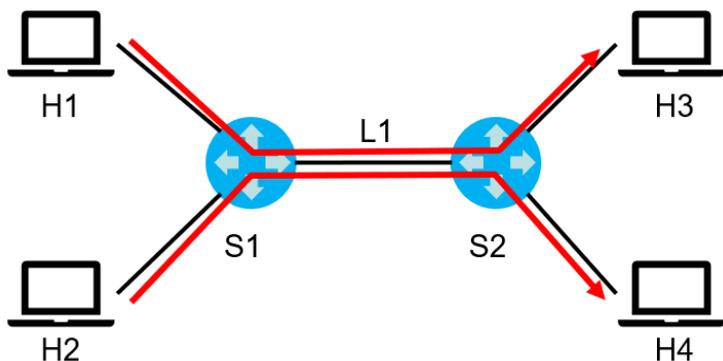
本発表の目的

- 複数リンクを経由するフロー間のCCAの公平性が、単一リンクを経由する場合の評価結果と異なるかを分析
- アプローチ
 - ダンベルトポロジと複数ホップトポロジの異なるトポロジを用いて評価

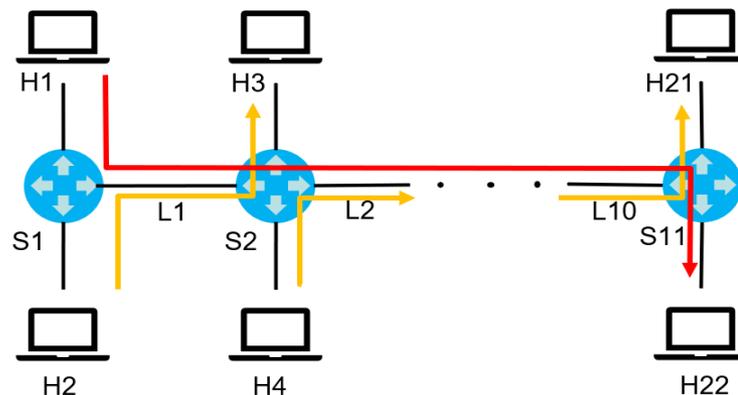


評価方法

- ネットワークエミュレータCOREを用いて評価
- 2つの異なるトポロジ
 - ダンベルトポロジ
 - 10ホップトポロジ
 - ホップ数の異なるフローが同時にリンクを共有して競合



ダンベルトポロジ

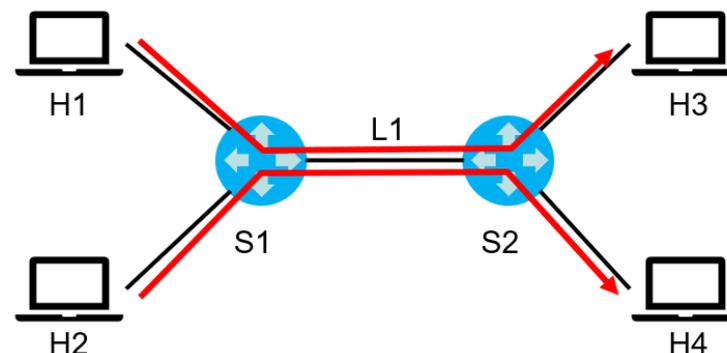


10ホップトポロジ

評価条件(ダンベルトポロジ)

■ 構成

- 2スイッチ: S1, S2
- 4ホスト: H1~H4
- 2フロー:
 - H1→H3, H2→H4



■ 競合リンク

- 帯域: 50Mbps
- 遅延: 2ms, 18ms, 100ms
- $B = \alpha \times BDP$ [packets] ($\alpha \in \{0.5, 1, 2, 4, 8, 16\}$)

■ 比較CCA

- TCP-Reno, CUBIC-TCP, BBR

■ フロー2本を同時生成し, 1分間スループットを観測

評価条件(10ホップトポロジ)

■ 構成

- 11スイッチ: S1~S11

- 22ホスト: H1~H22

- 11フロー:

- H1→H22(10競合リンク経由)

- H2→H3, H4→H5, ..., H20→H21(1競合リンク経由)

■ 競合リンク(L1, L2, ..., L10)

- 帯域: 50Mbps

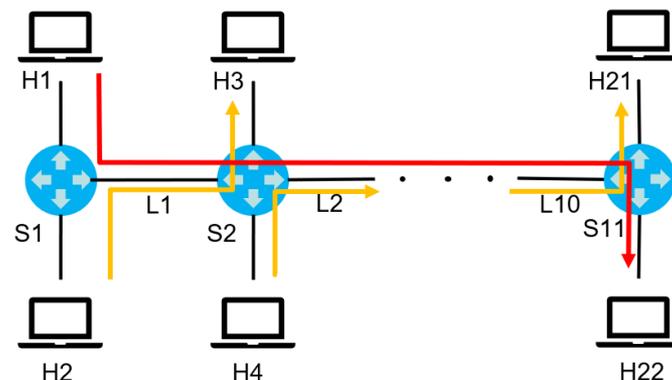
- 遅延: 2ms, 18ms, 100ms

- $B = \alpha \times BDP$ [packets] ($\alpha \in \{0.5, 1, 2, 4, 8, 16\}$)

■ 比較CCA

- TCP-Reno, CUBIC-TCP, BBR

- フロー11本を同時生成し, 1分間スループットを観測



ルータの出力バッファサイズ

■ 1BDPの計算方法

- $BDP = \text{帯域幅}[bps] \times \text{遅延}[\text{秒}]$

■ 例) ダンベルトポロジ, 帯域幅50Mbps, 遅延1ms, 2ms, 1パケット1500Bytesの場合

- $50 \text{ Mbps} = 50 \times 10^6 \text{ bps}$

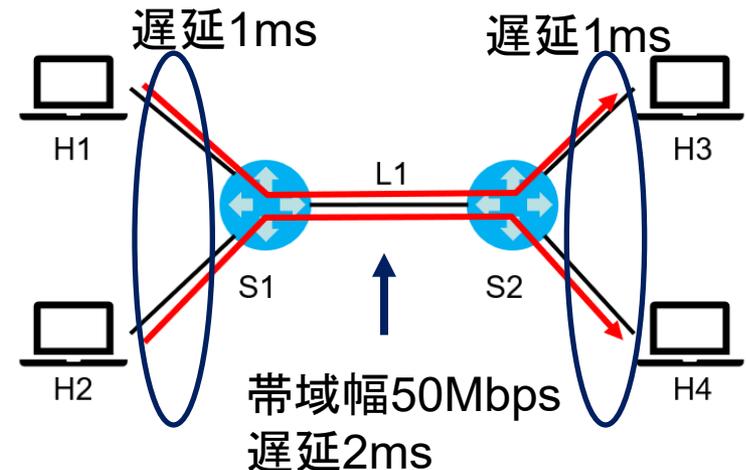
- $RTT = (1 + 2 + 1) \times 2 = 8 \text{ ms} = 8 \times 10^{-3} \text{ s}$

- $BDP = 50 \times 10^6 \times 8 \times 10^{-3}$

$$= 400,000 \text{ bits}$$

$$= 50,000 \text{ Bytes}$$

$$= \frac{50000}{1500} \approx 33.33 \text{ packet}$$



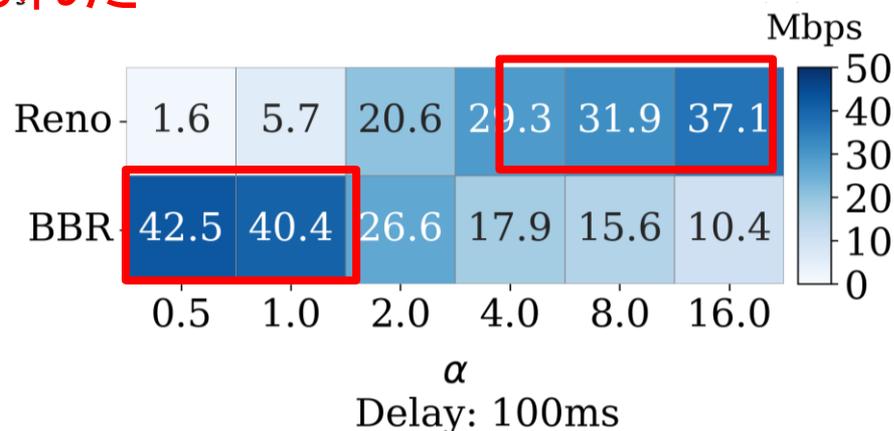
評価結果(ダンベルトポロジ)

- バッファサイズ α の値が小さい時:
 - BBRが帯域を占有
- バッファサイズ α の値が大きい時:
 - Renoが帯域を占有
- 遅延が大きい時, BBRのスループットは低下しRenoのスループットが上昇

ルータバッファサイズ =
 $\alpha \times$ 帯域遅延積BDP

➤ 先行研究と同様の知見が得られた

- バッファサイズ α の値が小さい時:
 - BBRが帯域を占有
- バッファサイズ α の値が大きい時:
 - Renoが帯域を占有
- 遅延が大きい時, BBRのスループットが上昇

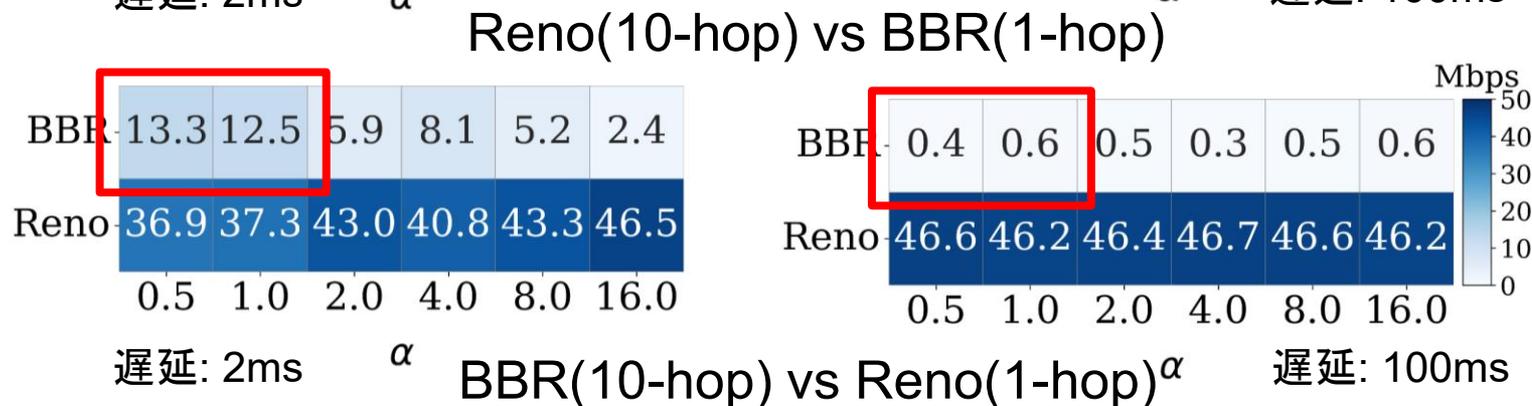
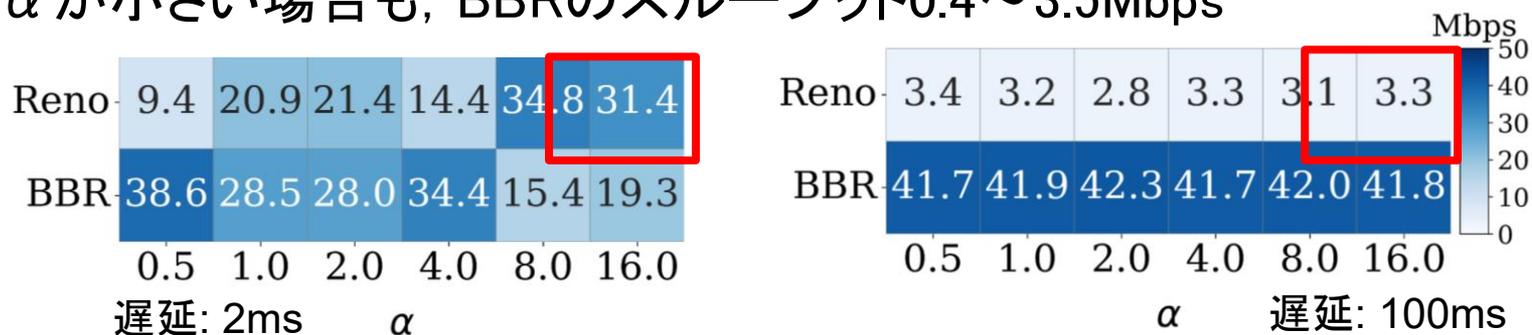


➤ 先行研究と同様の知見が得られた

Reno vs BBR

評価結果 (10ホップトポロジ)

- ホップ数の少ないフローが多くのスループットを取得する傾向
- 遅延が増大した時, 10ホップRenoのスループットは低下し, 1ホップBBRのスループットは上昇
- BBRが10ホップフローに割り当てられた時, スループットの低下率が激しい
 - α が小さい場合も, BBRのスループット0.4~3.5Mbps



まとめ

- 複数リンクを経由する場合CCAの公平性を評価
- エミュレーション評価結果
 - ダンベルトポロジ:
 - バッファサイズにより支配的CCAが変動
 - 先行研究と同様の傾向
 - 複数リンク経由トポロジ:
 - ホップ数小さいフローが帯域を占有
 - α が小さい場合もBBRのスループットが著しく低下
- 今後の展望
 - QUICで実装されたRenoやCubic, BBRでの公平性分析

バッファサイズと各CCAの挙動

■ バッファサイズが大きい

- 損失ベース: パケットロスが発生するまでルータのバッファを詰め込む
- BBR: バッファにパケットが溜まり始め、RTTが増大する→輻輳発生と判断し抑制

■ バッファサイズが小さい

- 損失ベース: バッファが小さいため、すぐにパケットロスが発生→抑制
- BBR: パケットロスを輻輳の兆候としないため通常運行