

インライン計測に基づく TCPによるバックグラウンド転送方式

大阪大学 大学院情報科学研究科
情報ネットワーク学専攻
博士前期課程 津川 知朗

発表内容

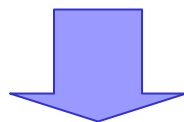
- ④ 研究背景
 - 優先度の異なるデータ転送
- ④ 研究目的
 - TCPによるバックグラウンド転送方式の実現
- ④ 従来のTCPによるバックグラウンド転送方式
- ④ 提案方式 ImTCP-bg
- ④ シミュレーションによる性能評価
- ④ まとめと今後の課題

研究背景

- ④ 優先度の異なるデータ転送
 - サービスの品質が向上する場合がある
 - ④ e.g. Contents Delivery Network (CDN)
 - ④ プリフェッチ等のデータ転送に低い優先度を設定する事でユーザからの要求に対するデータ転送が優先される
- ④ TCP Renoを用いて実現することは非常に難しい
 - 輻輳制御機構の基本的な性質に原因が存在する
 - ④ パケット廃棄が発生するまで輻輳ウィンドウサイズを増加させ続けるためフォアグラウンドトラヒックにも影響を与える

優先度の異なるデータ転送の実現

- a IPネットワークに新しい機構を組み込む方法
 - o 全てのルータが対応していなければならない
 - u ネットワーク規模に対するスケーラビリティの面で困難である
- a アプリケーションが優先度を与える方法
 - o エンドホストのみで実現可能
 - o アプリケーションでの制御は非常に難しい
 - u データ転送速度はTCPの制御によって決定されるため



TCPで実現する事がもっとも望ましい

研究目的

- ④ TCPによるバックグラウンド転送方式の実現
ImTCP-bg (ImTCP background mode)
 - 既存のTCPよりも優先度の低いデータ転送
 - 次の性質を持つことが重要である
 1. フォアグラウンドトラフィックに影響を与えない
 2. ネットワーク上の空き帯域を有効に利用する
 - インライン計測の計測結果を用いて実現する

従来のTCPによるバックグラウンド転送方式

- ④ フォアグラウンドトラフィックに影響を与えないことに主眼
 - e.g. TCP Nice [8], TCP-LP [9]
 - ④ データパケットのRTTや片道遅延時間を輻輳の指標として用いる
 - ④ TCP Renoよりも優先度の低いデータ転送を実現している
- ④ ネットワークの空き帯域を十分に利用できない
 - 輻輳ウィンドウサイズの減少量が一定かつ大きく設定されてある
 - ④ フォアグラウンドトラフィックへ影響を与えないため
 - 利用可能な帯域を知る為の効率的な方法がないため

[8] A. Venkataramani, R. Kokku, and M. Dahlin, “TCP Nice: A mechanism for background transfers,” in Proceedings of the 5th Symposium on Operating Systems Design and Implementation, Dec. 2002.

[9] A. Kuzmanovic and E. W. Knightly, “TCP-LP: A distributed algorithm for low priority data transfer,” in Proceedings of IEEE INFOCOM 2003, Apr. 2003.

Inline Measurement TCP (ImTCP)

- ④ インラインネットワーク計測手法のひとつ
 - データ・ACKパケットを利用して利用可能帯域の計測を行う
 - ④ 計測アルゴリズムによりデータパケットの送信間隔を調節する
 - ④ ACKパケットの受信間隔から利用可能帯域を導出する
- ④ ImTCPの特徴
 - 少ない数のパケットで計測を行うことができる
 - 短い周期 (1 – 4 RTT) で継続的に計測を行うことができる
 - ④ 利用可能帯域の変動に素早く追随することができる
 - 送信側のTCPの修正のみで計測を行うことができる

[14] M. L. T. Cao, G. Hasegawa, and M. Murata, “Available bandwidth measurement via TCP connection,” in Proceedings of IFIP/IEEE MMNS 2004 E2EMON Workshop, Oct. 2004.

ImTCPの問題点

- ④ 計測失敗や計測結果が不正確になる場合が存在
 - 計測できない場合
 - ④ 計測に必要なデータパケットを確保できない場合
 - ④ 輻輳ウィンドウサイズが小さいため
 - 計測結果が不正確になる場合
 - ④ 複数のImTCPコネクションが同時に計測を行う場合
 - ④ 計測結果が実際の利用可能帯域の値よりも大きくなる
- ④ そのまま輻輳ウィンドウサイズの制御に用いた場合
 - フォアグラウンドトラヒックへ影響を与える

ImTCP-bg の目標

- ④ 計測結果が正確かどうかをRTTを観測して判断
 - 輻輳ウィンドウサイズが大きくなるにつれてRTTも増大するため
- ④ ImTCPの計測結果が正確な場合
 - 計測結果を用いて輻輳ウィンドウサイズの制御を行う
- ④ ImTCPの計測結果が不正確な場合
 - RTTの大きさに応じて輻輳ウィンドウサイズを減少させる

ImTCP-bg アルゴリズム (1)

- 次式により計測結果が正確であるかどうかの判断を行う

$$\frac{RTT_{cur}}{RTT_{min}} > \delta$$

: しきい値 (1)

RTT_{cur} , RTT_{min} : RTTの現在値 , 最小値

- 計測結果が正確な場合

- ImTCPの計測結果を輻輳ウィンドウサイズ制御に用いる

- 計測結果に対して平滑化を行う

$$\bar{A} \leftarrow (1 - \gamma) \times \bar{A} + \gamma \times A_{cur} \quad : \text{平滑化パラメータ (0 } \gamma < 1)$$

A_{cur} : 最新の利用可能帯域

- 送信側のTCPの輻輳ウィンドウサイズの上限值を決定する

$$maxcwnd = \bar{A} \times RTT_{min}$$

- その他の輻輳制御はTCP Renoの方式を用いる

ImTCP-bg アルゴリズム (2)

a 計測結果が不正確な場合

- RTTの大きさに応じて輻輳ウィンドウサイズを減少させる
 - ↳ 現在のRTTおよびRTTの最小値を用いて減少量を決定する

$$cwnd \leftarrow cwnd \times \frac{RTT_{min}}{RTT_{cur}}$$

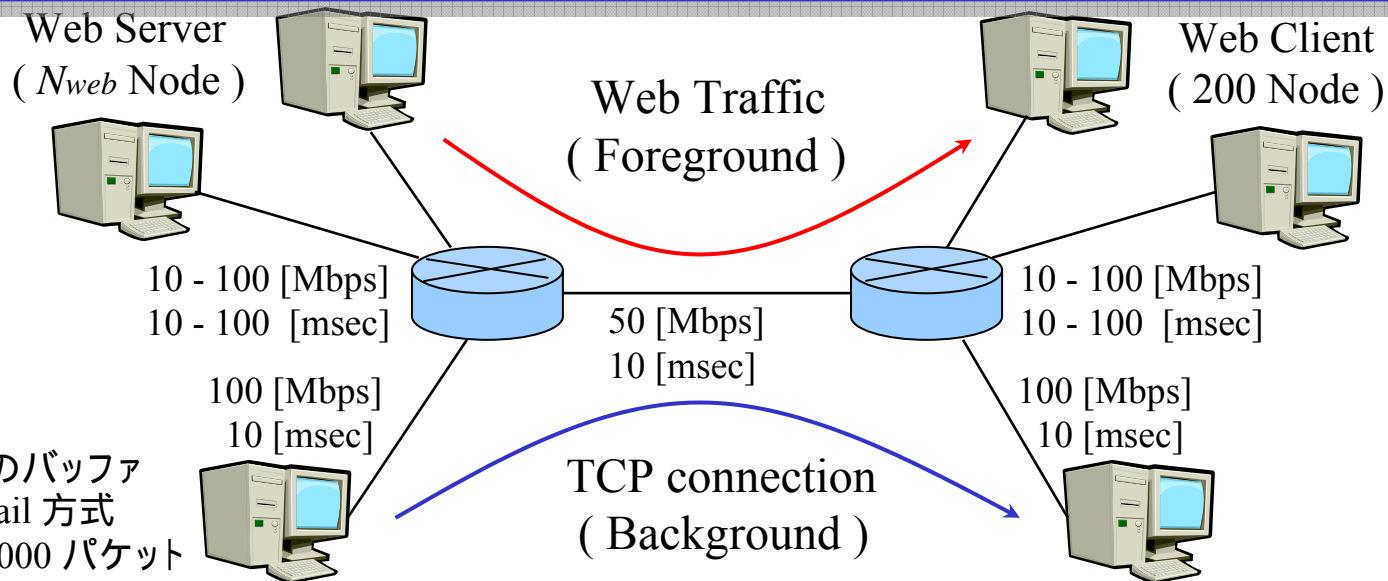
- 輻輳ウィンドウサイズの上限値は変化させない

シミュレーションによる性能評価

- ns-2を用いたシミュレーション
- ImTCP-bg の比較対象
 - TCP Reno
 - TCP Nice
 - TCP-LP

シミュレーション環境

平均60KByteのパレート分布に従ってWebページサイズを決定



評価項目

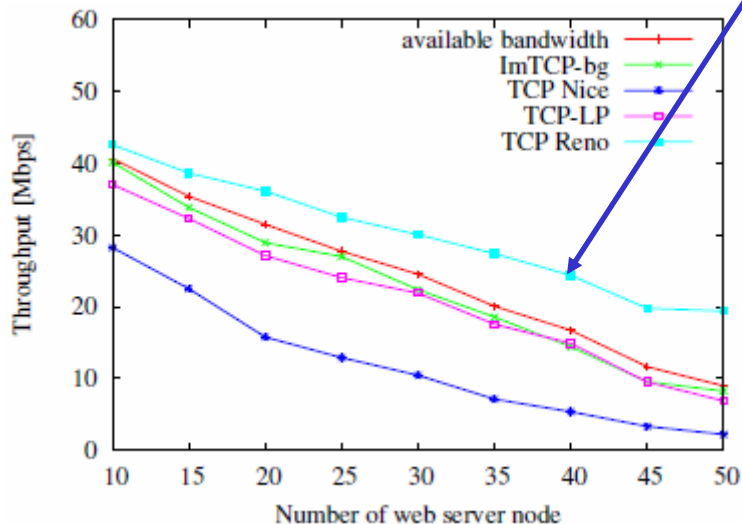
- フォアグラウンドトラフィックへの影響度
- バックグラウンドトラフィックのスループット

1本のバックグラウンドトラフィックの評価

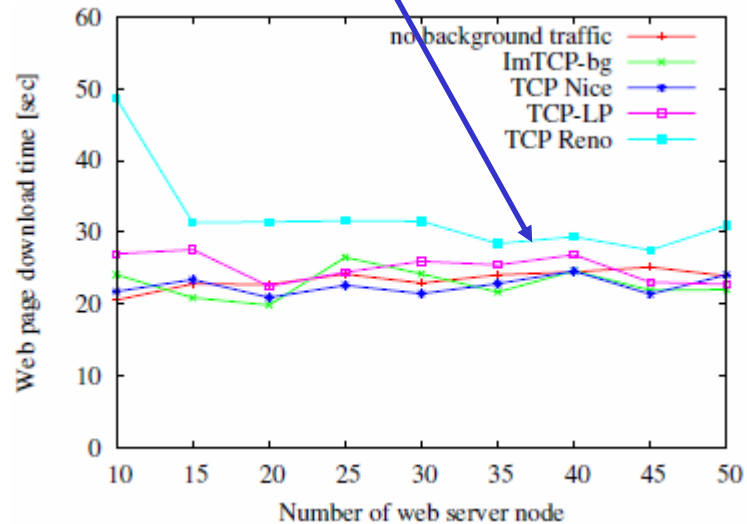
TCP Renoを用いた場合,最も高いスループットが得られるが
Webページの平均ダウンロード時間が増大している



フォアグラウンドトラフィックに大きな影響を与えている



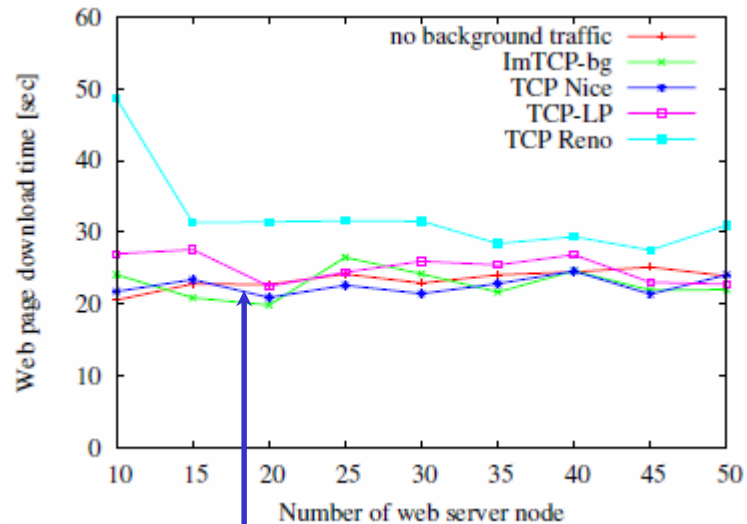
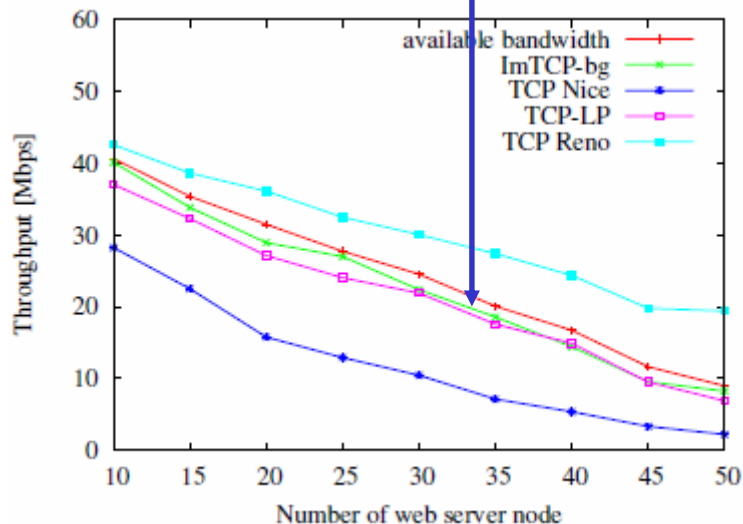
Webサーバノード数に対するバックグラウンド
トラフィックの平均スループットの変化



Webサーバノード数に対するWebページの
平均ダウンロード時間の変化

1本のバックグラウンド転送の評価

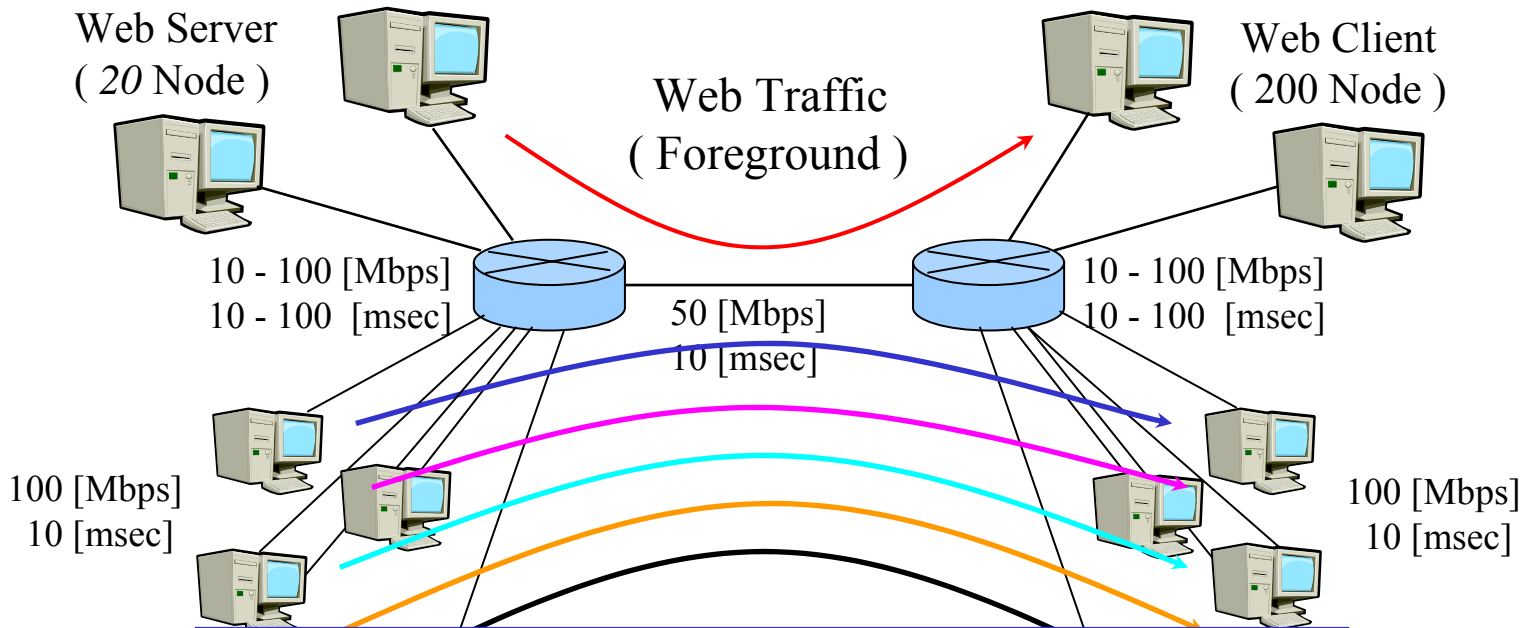
ImTCP-bgを用いた場合, 3つのバックグラウンド転送方式の中で最も真の利用可能帯域に近いスループットを得ている



Webサーバノード数に対するバックグラウンドトラフィックの平均スループット

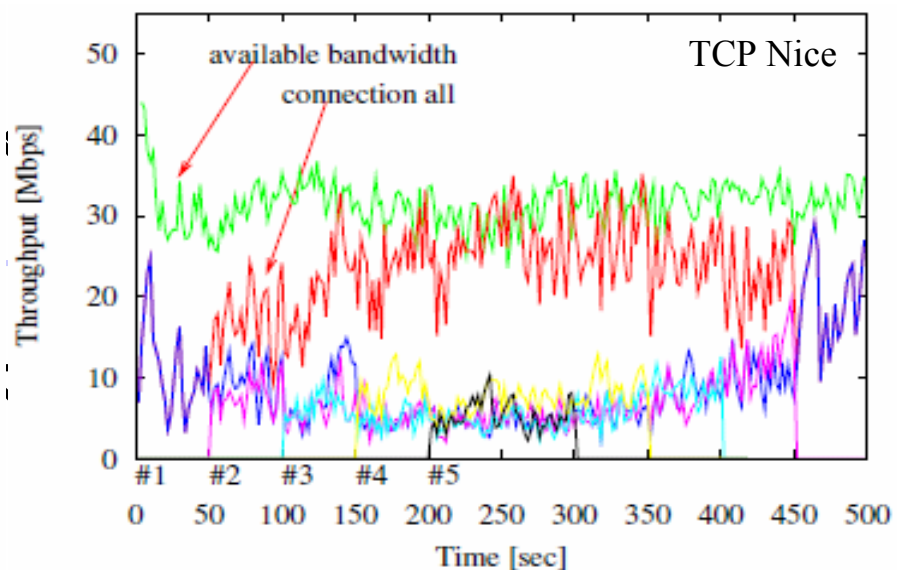
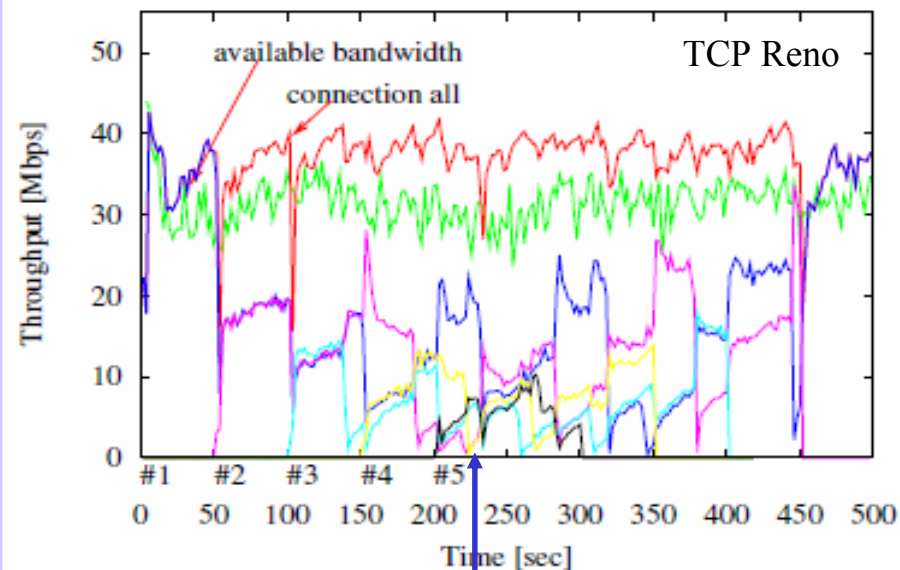
TCP Nice, TCP-LP, ImTCP-bgを用いた場合Webページの平均ダウンロード時間がバックグラウンドトラフィックが存在しない場合とほぼ同等である

シミュレーション環境

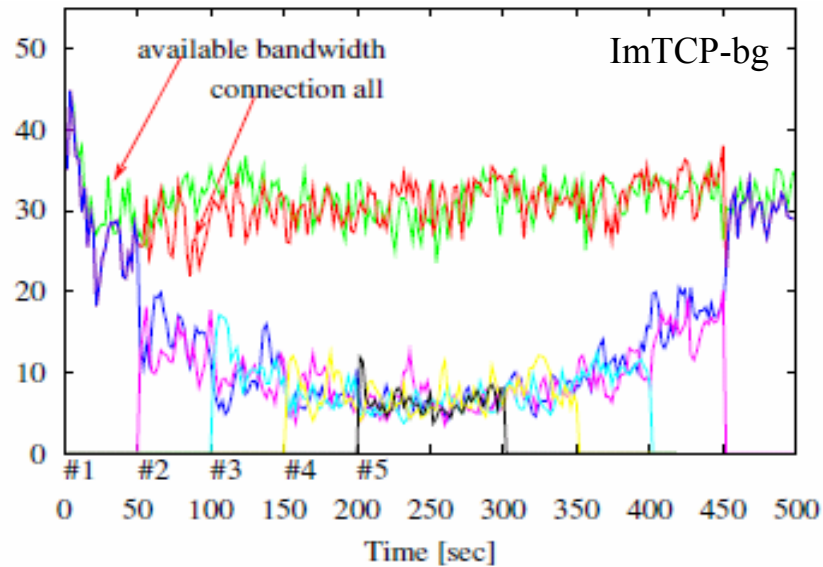
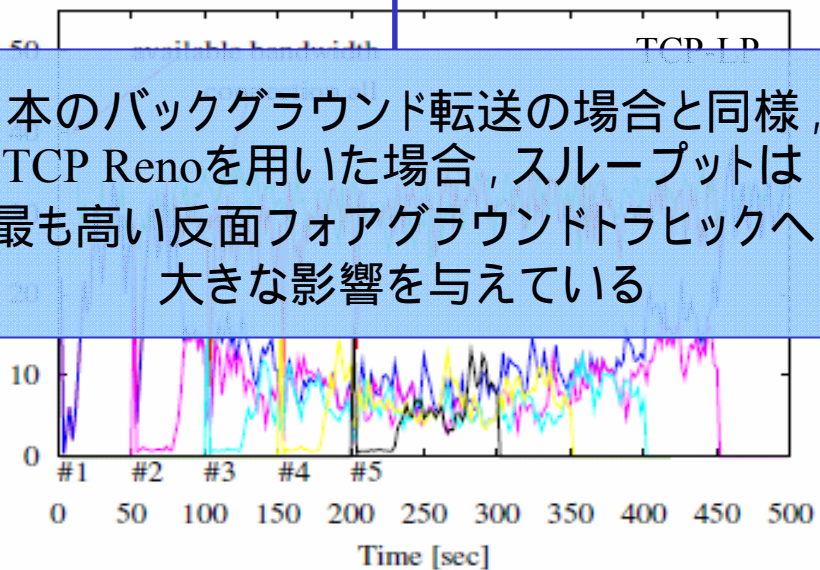


評価項目

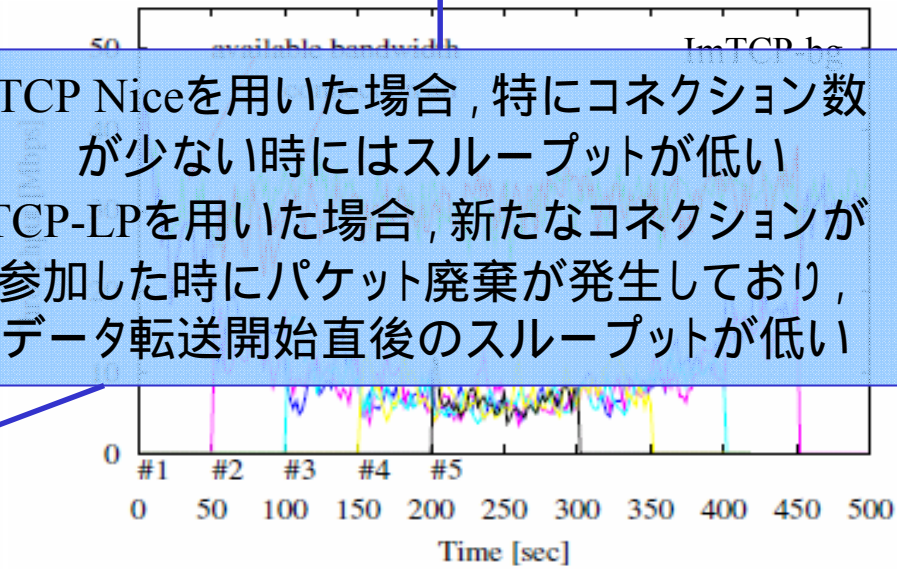
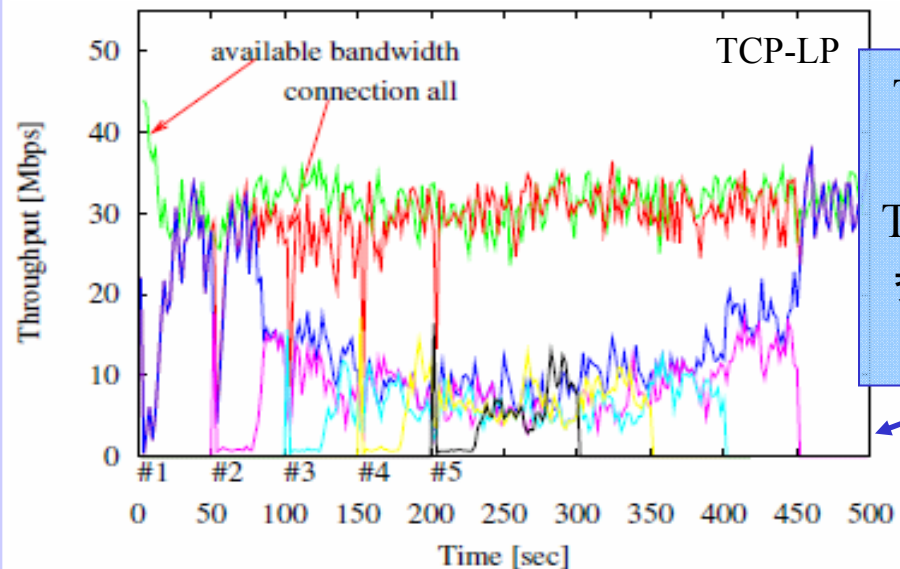
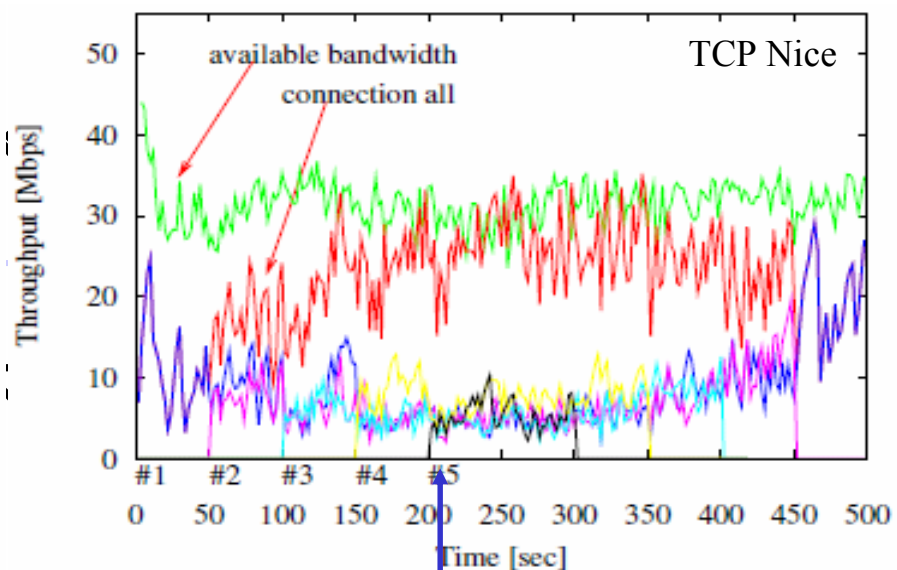
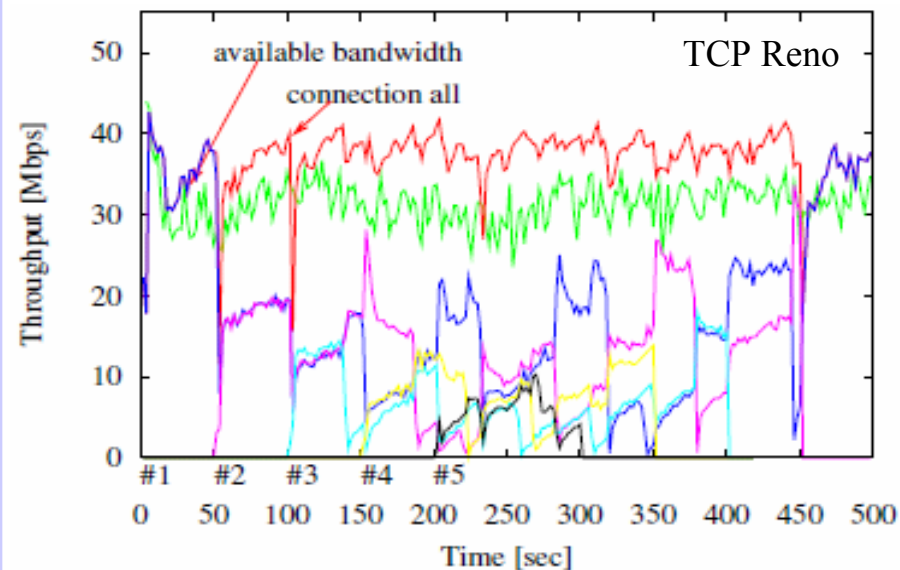
- フォアグラウンドトラフィックへの影響度
- バックグラウンドトラフィックのスループット
- コネクション間の公平性



1本のバックグラウンド転送の場合と同様、TCP Renoを用いた場合、スループットは最も高い反面フォアグラウンドトラフィックへ大きな影響を与えている



経過時間に対するバックグラウンドトラフィックのスループットの変化

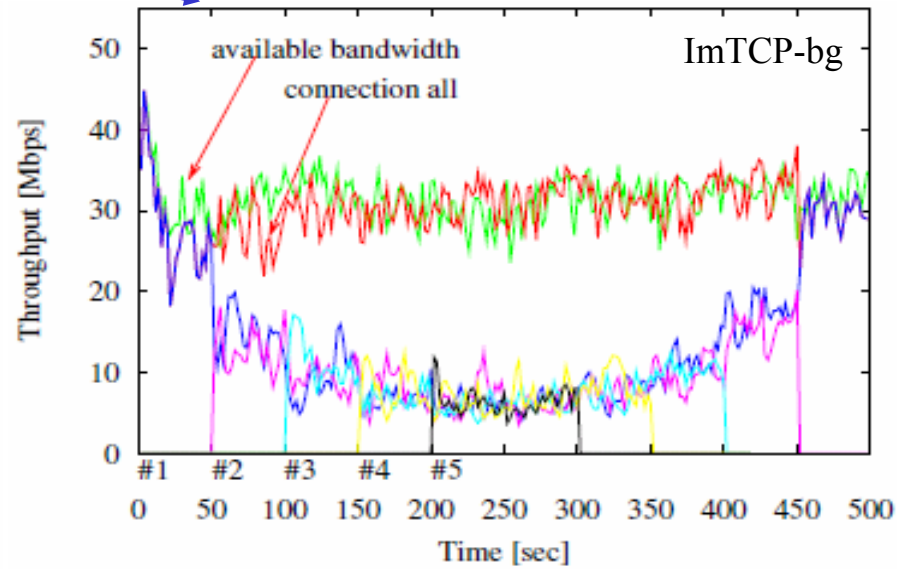
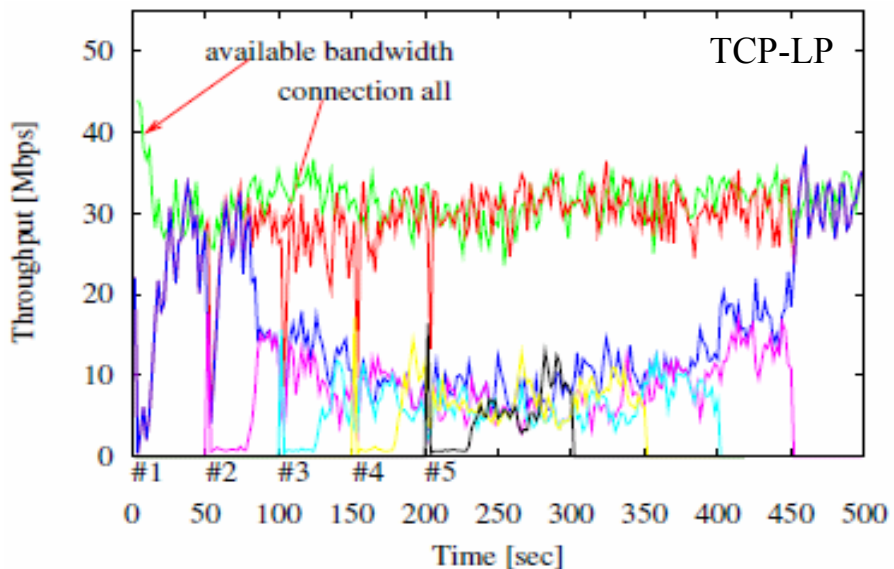
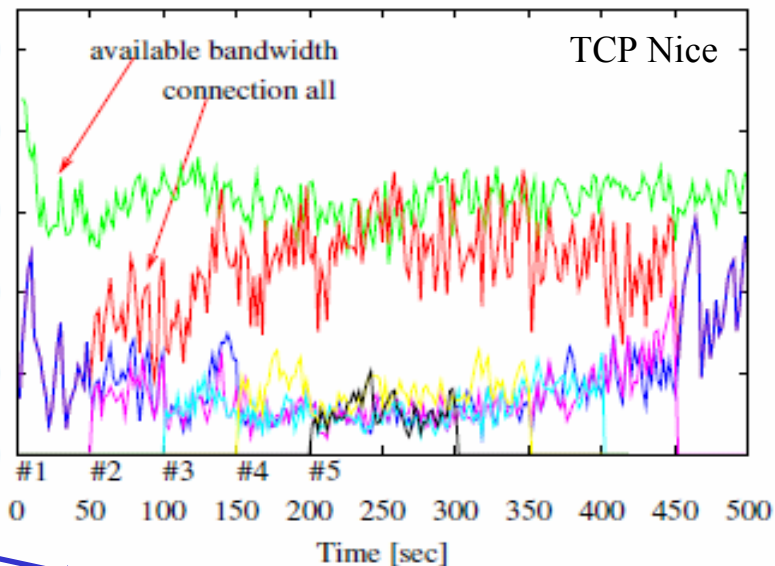
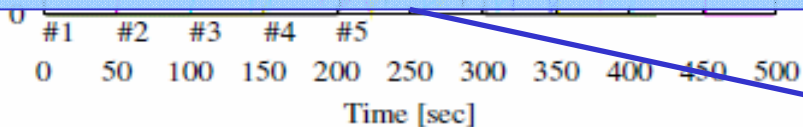


TCP Niceを用いた場合, 特にコネクション数が少ない時にはスループットが低い
 TCP-LPを用いた場合, 新たなコネクションが参加した時にパケット廃棄が発生しており, データ転送開始直後のスループットが低い

経過時間に対するバックグラウンドトラフィックのスループットの変化

ImTCP-bgを用いた場合、コネクションが1本の時点から空き帯域を有効に利用している

コネクションが複数本存在している場合、スループットの変動が他の方式に比べて小さい
TCP NiceやTCP-LPと同様、コネクション間の公平性を維持することが出来ている



経過時間に対するバックグラウンドトラフィックのスループットの変化

まとめと今後の課題

- ④ TCPによる新しいバックグラウンド転送方式の提案
 - インライン計測を用いて利用可能帯域を取得する
 - 計測結果を用いて輻輳ウィンドウサイズを制御する
- ④ シミュレーションにより目的が達成されていることを示した
 - フォアグラウンドトラフィックへ影響を与えない
 - ネットワーク上の空き帯域のみを有効に利用する
- ④ 今後の課題
 - 適切なパラメータ設定に関する議論
 - ⑤ 計測結果が不正確になった場合に輻輳を検知するためのしきい値
 - 実ネットワーク環境での性能評価