



# TCP-based Background Data Transfer using Inline Network Measurement Technique

津川 知朗 ( [t-tugawa@ist.osaka-u.ac.jp](mailto:t-tugawa@ist.osaka-u.ac.jp) )

大阪大学 大学院情報科学研究科

村田研究室

# 発表内容

- ◆ はじめに
- ◆ 従来のTCPによるバックグラウンドデータ転送
- ◆ ImTCP-bgアルゴリズム
- ◆ 性能評価
  - ◆ シミュレーションによる性能評価
  - ◆ 実ネットワーク上での性能評価
- ◆ まとめと今後の課題

# はじめに

## 研究背景: 優先度の異なるデータ転送の実現

- ⊕ インターネット上では様々なサービスが提供されている
  - ◆ データ転送に優先度を設けることによって, これらのサービスの品質が向上する場合が存在する
- ⊕ TCP Reno<sup>※1</sup>を用いて実現することは困難である

## 研究目的: TCPによるバックグラウンド転送の実現

- ⊕ 次の性質を満たすことが重要である
  - ◆ 他のトラフィックに影響を与えずにデータ転送を行う
  - ◆ 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利用して実現する

※1 インターネット上において、広く用いられているトランスポート層のプロトコル

# 従来のTCPによるバックグラウンド転送

**主眼:** 他のトラフィックに影響を与えない

- ⊕ e.g., TCP Nice [22], TCP-LP [23]
  - ◆ データパケットのRTTを輻輳の指標として用いる
  - ➡ TCP Renoよりも早期にネットワーク輻輳を検知できる
  - ➡ 適切な輻輳回避を行うことでバックグラウンド転送を実現できる

**問題点:** 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い

利用可能帯域を知るための効率的な方法を持たないため

[22] A. Venkataramani, R. Kokku, and M. Dahlin, "TCP Nice: A mechanism for background transfers," in Proceedings of OSDI 2002, December 2002.

[23] A. Kuzmanovic and E. W. Knightly, "TCP-LP: A distributed algorithm for low priority data transfer," in Proceedings of IEEE INFOCOM 2003, April 2003.

# Inline measurement TCP (ImTCP)

- ◆ インラインネットワーク計測手法のひとつ
  - ⊕ データ・ACKパケットのみを用いて計測を行う
- ◆ 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
- ◆ ImTCPの問題点
  - ⊕ 計測結果が不正確になる場合が存在する
    - ◆ 計測結果が実際の利用可能帯域よりも大きい場合
    - ◆ そのまま輻輳制御に利用すると他のトラフィックに影響を与える

[25] Cao Le Thanh Man, Go Hasegawa, and Masayuki Murata, "Available bandwidth measurement via TCP connection," in Proceedings of IFIP/IEEE MMNS 2004 E2EMON Workshop, October 2004.

# ImTCP-bgアルゴリズム

## ◆ ImTCPの計測結果を基に輻輳ウィンドウサイズを制御する

- ✦ 計測結果に対して平滑化を行う

$$\bar{A} \leftarrow (1-\gamma) \times \bar{A} + \gamma \times A_{cur}$$

$\gamma$ : 平滑化パラメータ ( $0 < \gamma < 1$ )  
 $A_{cur}$ : 最新の計測結果

- ✦ 送信側TCPの輻輳ウィンドウサイズの上限值を決定する

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

$RTT_{min}$ : RTTの最小値

- ✦ その他の輻輳制御方式はTCP Renoの方式を用いる

## ◆ RTTを用いた輻輳ウィンドウサイズの制御を併用する

- ✦ ImTCPの計測結果が信頼できない場合が存在するため

- ✦ RTTの現在値および最小値からネットワークの輻輳を検知する

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

$\delta$ : しきい値 ( $1 \leq \delta$ )  
 $RTT_{cur}$ ,  $RTT_{min}$ : RTTの現在値, 最小値

- ✦ RTTの値に応じて輻輳ウィンドウサイズを減少させる

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

# 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

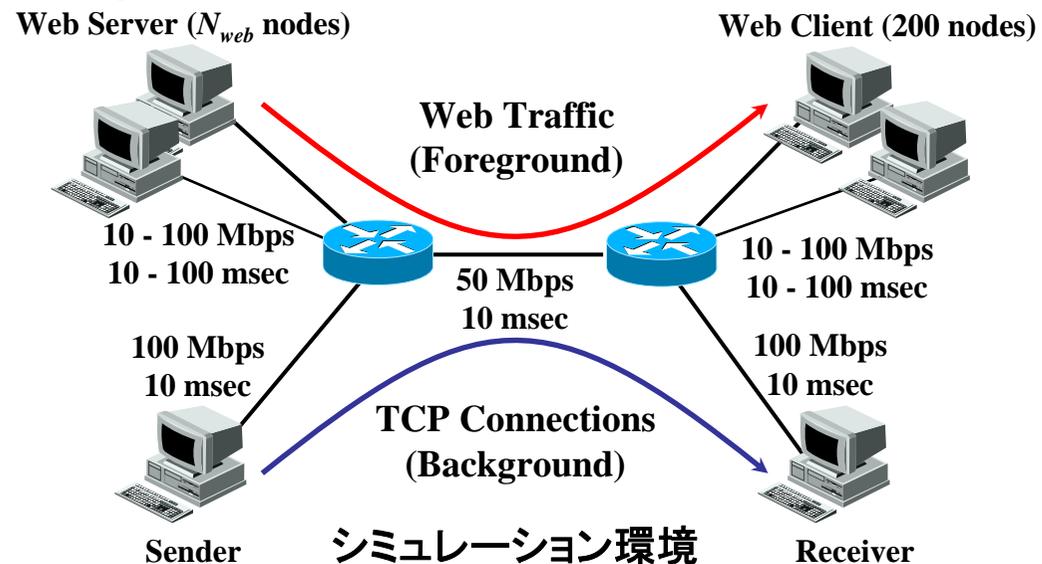
◆ TCP Reno

◆ TCP Nice [22]

◆ TCP-LP [23]

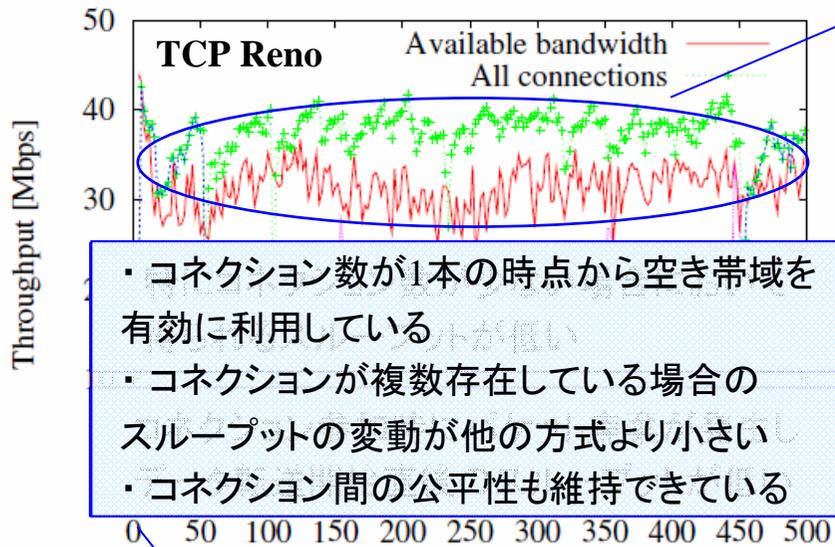
※ Webトラフィックは、平均60KByteのパレート分布に従ってWebページサイズを決定する

※ ルータのバッファはDropTail方式  
バッファサイズは1000パケット



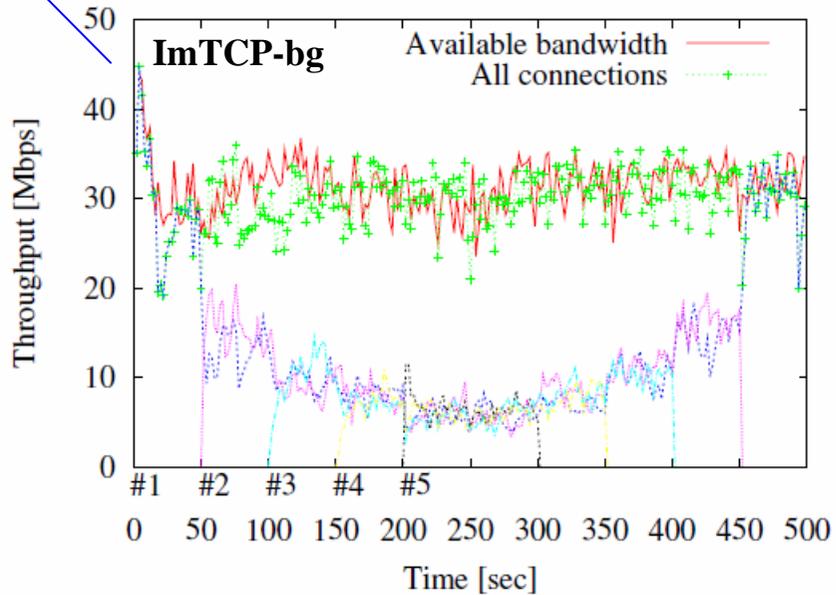
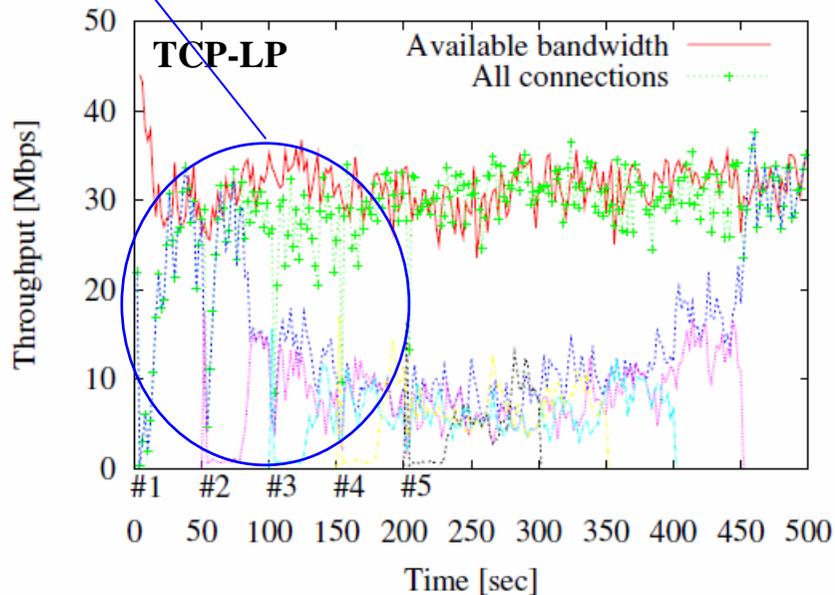
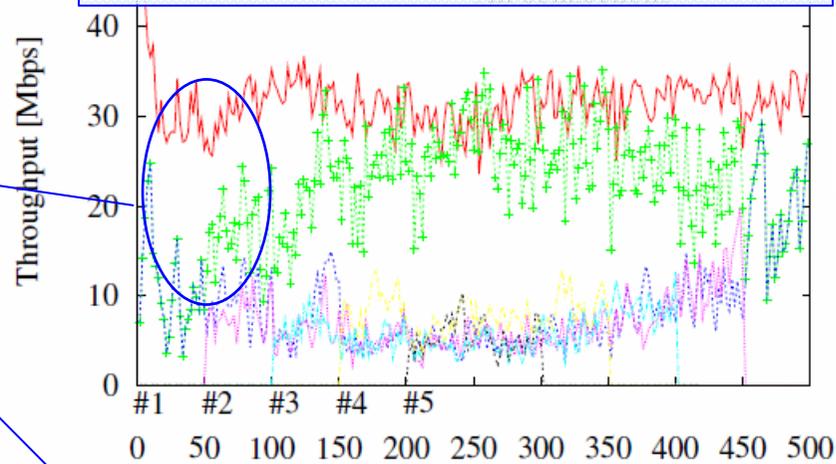
[22] A. Venkataramani, R. Kokku, and M. Dahlin, "TCP Nice: A mechanism for background transfers," in Proceedings of OSDI 2002, December 2002.

[23] A. Kuzmanovic and E. W. Knightly, "TCP-LP: A distributed algorithm for low priority data transfer," in Proceedings of IEEE INFOCOM 2003, April 2003.



- ・コネクション数が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存在している場合のスループットの変動が他の方式より小さい
- ・コネクション間の公平性も維持できている

全コネクションの合計スループットが  
利用可能帯域を大きく超えている



# 実ネットワーク上での性能評価

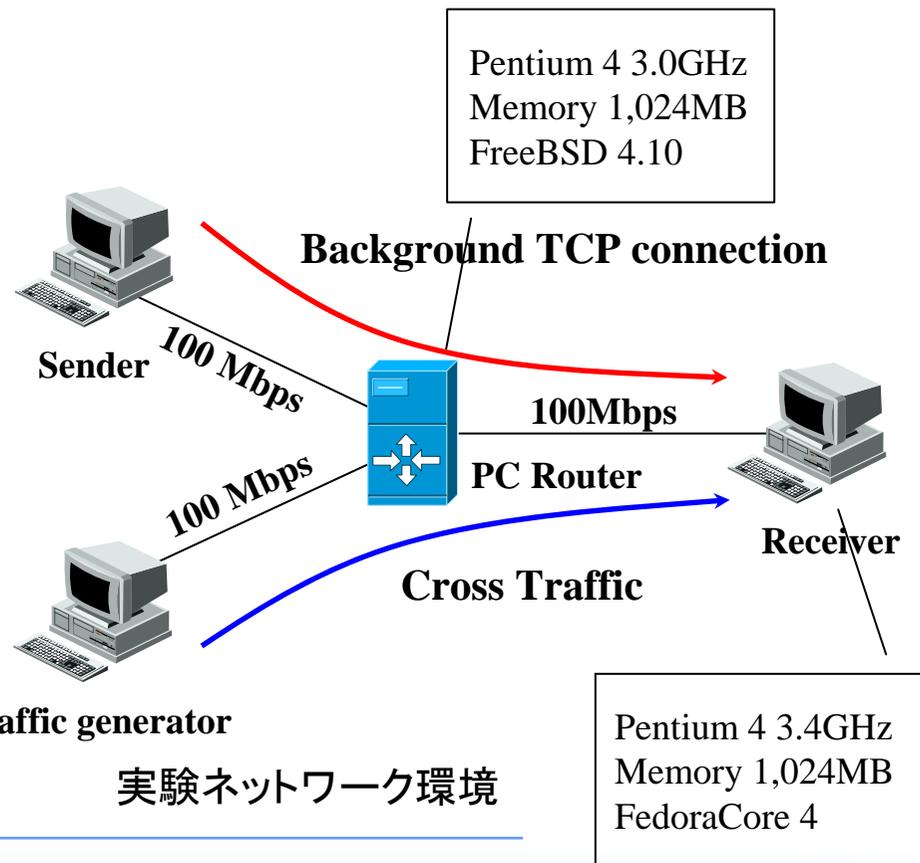
- ◆ 実験ネットワークを構築し実装実験を行う
- ◆ ImTCP-bgの比較対象

- ◆ TCP Reno

- ◆ TCP-LP

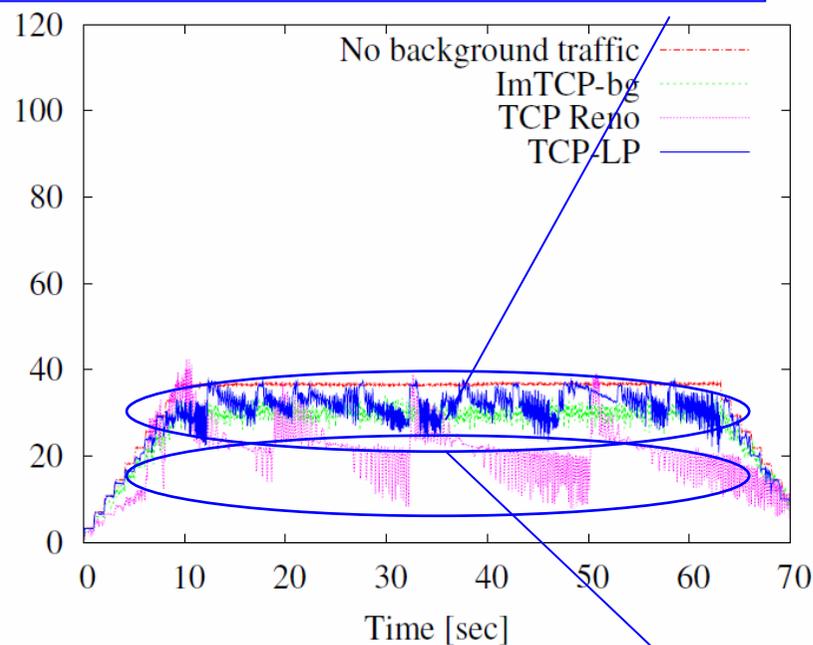
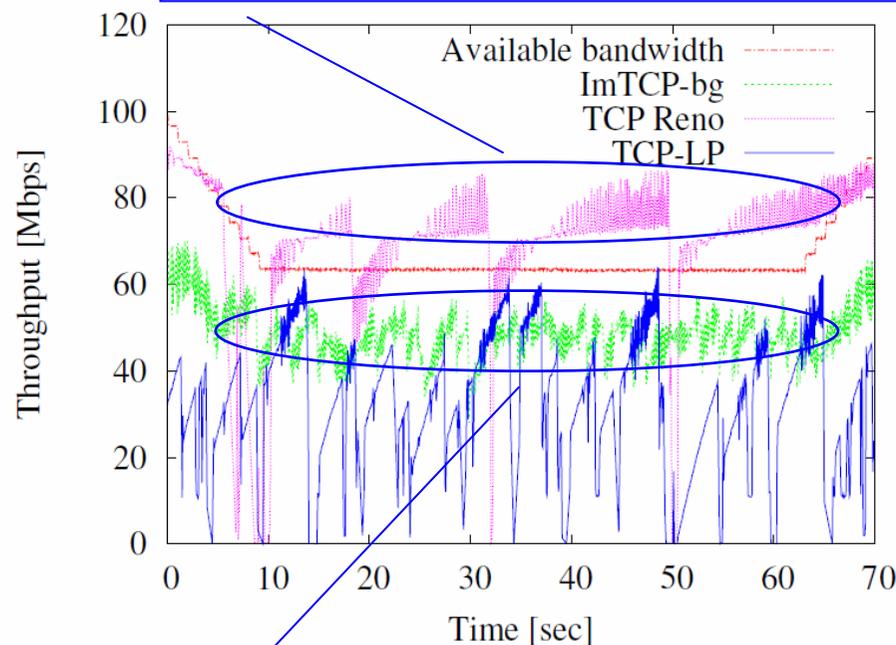
Pentium 4 3.0GHz  
Memory 1,024MB  
FreeBSD 4.10

Pentium 4 3.4GHz  
Memory 1,024MB  
FedoraCore 4



# 実験結果

バックグラウンドTCP接続のスループットが空き帯域を大きく超えている  
クロスラヒックのスループットも押し下げている



バックグラウンドTCP接続のスループット

ImTCP-bgを用いた場合に最も空き帯域に近いスループットを得ることができている

クロスラヒックのスループット

クロスラヒックのスループットをほとんど押し下げずにデータ転送を行うことができている

# まとめと今後の課題

## ◆ まとめ

- ✦ TCPによる新しいバックグラウンド転送方式の提案を行った
  - ◆ 他のトラフィックに影響を与えずにデータ転送を行う
  - ◆ ネットワーク上の空き帯域を有効に利用する
- ✦ 従来のバックグラウンド転送よりも優れていることを示した
  - ◆ シミュレーションによる性能評価
  - ◆ 実ネットワーク上での性能評価
    - ✦ ImTCP, ImTCP-bg の実装
    - ✦ 実験ネットワーク上での性能評価
    - ✦ インターネット上での実験

## ◆ 今後の課題

- ✦ 今回の実験と異なるネットワーク環境での性能評価