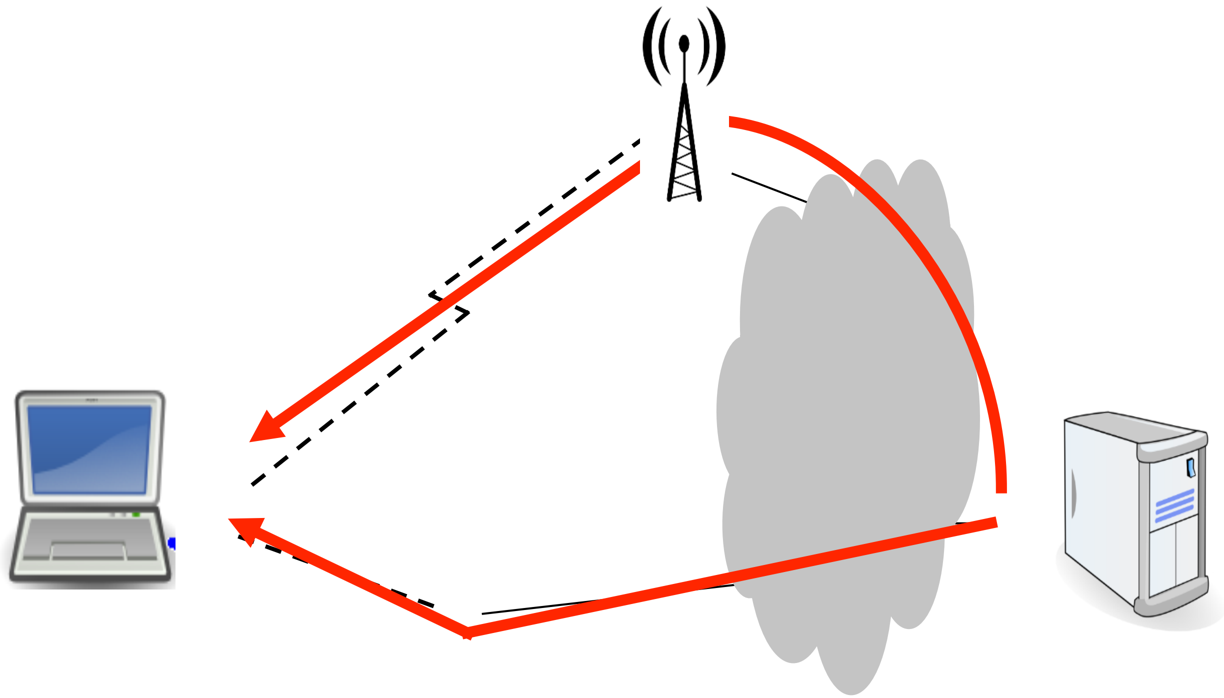
# Congestion Control of Multipath TCP: Problems and Solutions

Ramin Khalili, T-Labs/TU-Berlin, Germany

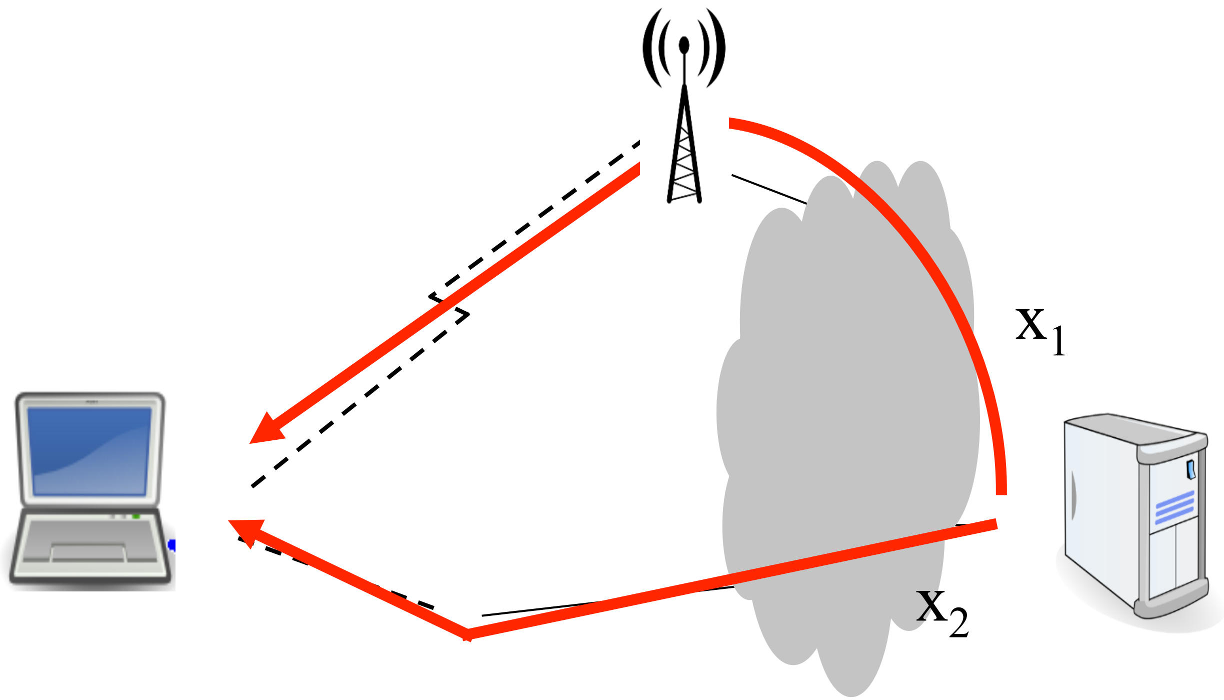
draft-khalili-mptcp-performance-issues-03
draft-khalili-mptcp-congestion-control-01

# Multipath TCP (MPTCP)



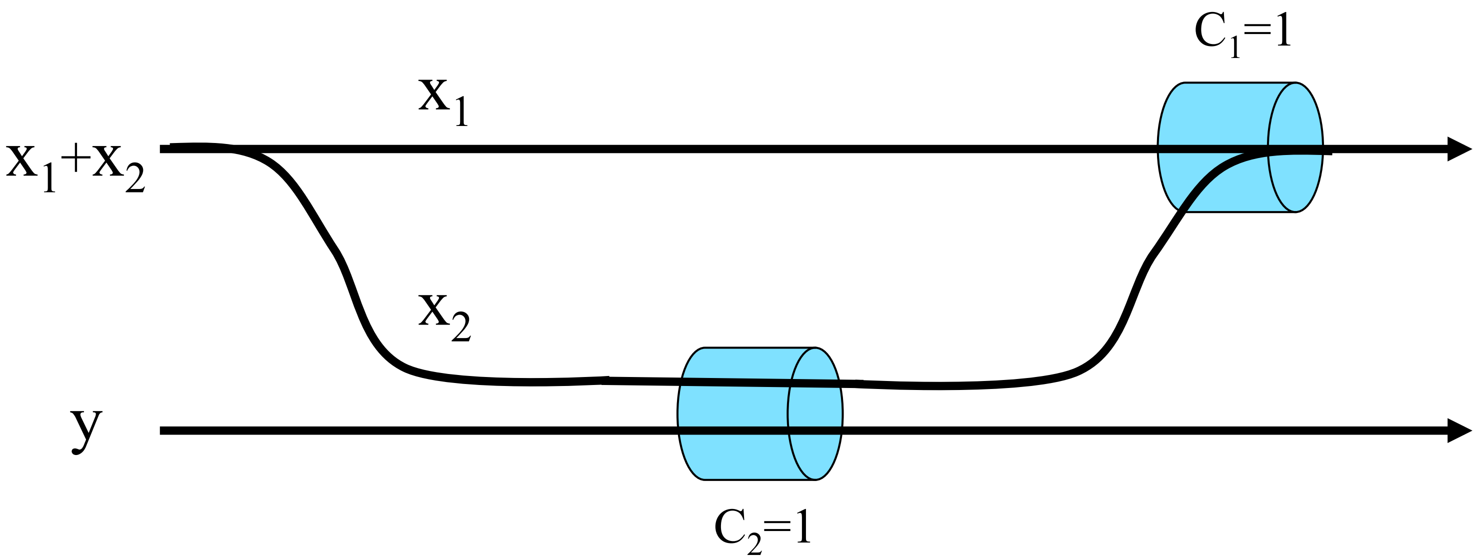
- allows a user to split its traffic across multiple paths
- improve reliability and throughput

# Congestion control




- to provide load balancing in the network
- what are rates $x_1$ and $x_2$

# Load balancing is not only about fairness




- with uncoupled congestion control

  $x_1+x_2 = 1$ ($x_1 = 0.6$ & $x_2 = 0.4$) and $y = 0.6$

- with an optimal algorithm

  $x_1+x_2 = 1$ ($x_1 = 1$ & $x_2 = 0$) and $y = 1$

# Kelly & Voice 2005: optimal load balancing, theoretical results

- optimal in static networks with all paths have similar RTT
- in practice, however,
  - not responsive: fails to detect free capacity in dynamic setting
  - flappy: when multiple good paths available, randomly flip traffic between them

# LIA [RFC 6356]: "Linked Increases" Algorithm

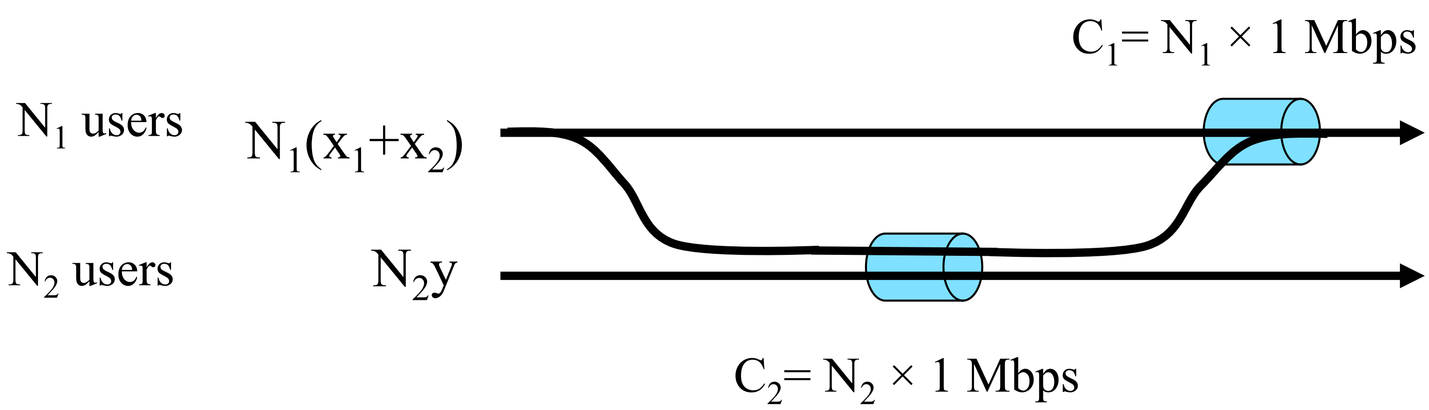
- adhoc design based on 3 goals
  1. improve throughput: total throughput $\geq$ TCP over best path
  2. do not harm: not more aggressive than a TCP over a path
  3. balance congestion while meeting the first two goals
- as also said in the RFC 6356, LIA does not fully satisfy goal 3

# LIA FAILS TO PROVIDE AN EFFICIENT LOAD BALANCING

R. Khalili, N. Gast, M. Popovic, J.-Y. Le Boudec, "Performance Issues with MPTCP", draft-khalili-mptcp-performance-issues-03

# MPTCP with LIA is suboptimal

| | | MPTCP with LIA (measurement) | | |
|---|---|---|---|---|
| $N_1=10$ | $x_1+x_2$ | 0.96 | | |
| $N_2=10$ | $y$ | 0.7 | | |
| $N_1=30$ | $x_1+x_2$ | $\boxed{0.96}$ | | |
| $N_2=10$ | $y$ | $\boxed{0.4}$ | | |



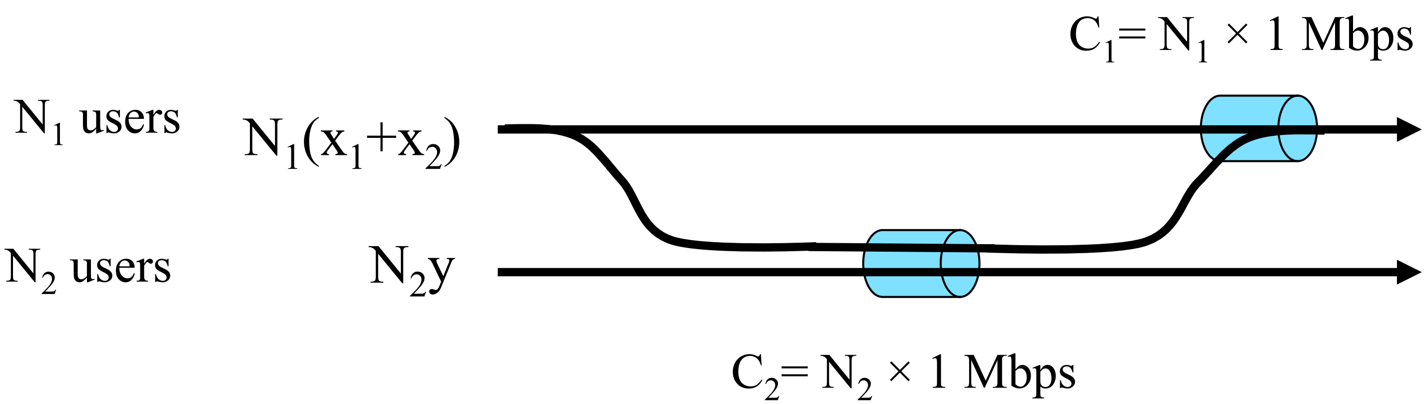
We compare MPTCP with two theoretical baselines

1. optimal algorithm (without probing cost):
   theoretical optimal load balancing [Kelly,Voice 05]

2. optimal algorithm with probing cost:
   theoretical optimal load balancing including minimal probing traffic
   - using a windows-based algorithm, a min probing traffic of 1 MSS/RTT is sent over each path

# Part of problem is in nature of things, but MPTCP seems to be far from optimal

| | | MPTCP with LIA (measurement) | optimal with probing cost (theory) | optimal w/out probing cost (theory) |
|---|---|---|---|---|
| $N_1$=10 | $x_1+x_2$ | 0.96 | 1 | 1 |
| $N_2$=10 | y | 0.7 | 0.94 | 1 |
| $N_1$=30 | $x_1+x_2$ | 0.96 | 1 | 1 |
| $N_2$=10 | y | 0.4 | 0.8 | 1 |



# CAN THE SUBOPTIMALITY OF MPTCP WITH LIA BE FIXED IN PRACTICE?

R. Khalili, N. Gast, M. Popovic, J.-Y. Le Boudec, "Opportunistic Linked-Increases Congestion Control Algorithm for MPTCP ", draft-khalili-mptcp-congestion-control-01

# LIA forces a tradeoff between responsiveness and load balancing

- to provide responsiveness, LIA departs from optimal load balancing
- Question: is it possible to come with a new design that provides both simultaneously?

# OLIA: an algorithm inspired by utility maximization framework

- simultaneously provides responsiveness and congestion balancing
- an adjustment of optimal algorithm [Kelly,Voice 05]
  - by adapting windows increases as a function of quality of paths, we make it responsive and non-flappy
- part of the Louvain MPTCP implementation

# Set of collected paths (collected_paths)

- $l_r$: smoothed estimation of number of bytes transmitted between last two losses
- best_paths: set of paths with max $(l_r * l_r)/rtt_r$
  - paths that are presumubly the bests for the MPTCP connection (based on TCP loss-throughput formula)
- max_w_paths: set of path with max windows
- collected_paths: set of paths in best_paths but not in max_w_paths

# OLIA: "Opportunistic Linked-Increases Algorithm"

For each path r:

- **increase part:** for each ACK on r, increase $w_r$ by

$$\frac{w_r/\mathrm{rtt}_r^2}{\left(\sum_{p\in\mathcal{R}_u} w_p/\mathrm{rtt}_p\right)^2} + \frac{\alpha_r}{w_r}$$

α is possitive for collected_paths

optimal congestion balancing: adaptation of [kelly, voice 05]

responsiveness; reacts to changes in current windows

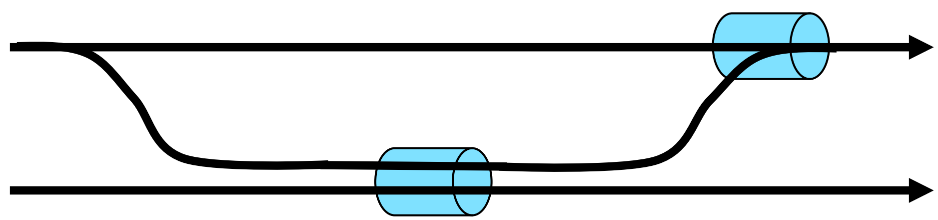
- **decrease part:** each loss on r, decreases $w_r$ by $w_r/2$

# Theoretical results: OLIA solves problems with LIA

- using a fluid model of OLIA
- Theorem: OLIA satisfies design goals of LIA (RFC 6356)
- Theorem: OLIA is Pareto optimal
- Theorem: when all paths of a user have similar RTTs, OLIA provides optimal load balancing similarly to [kelly, voice 05]

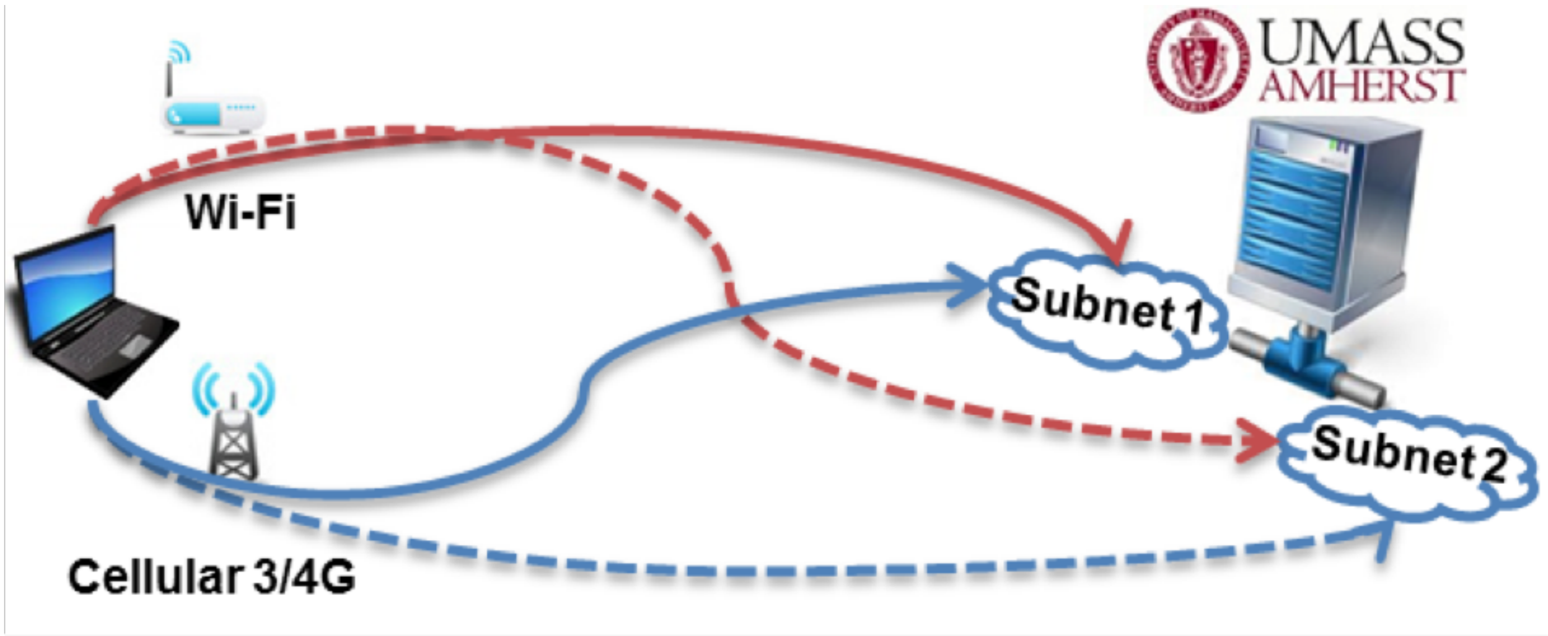
# OLIA performs close to optimal algorithm with probing cost

| | | MPTCP with LIA (measurement) | MPTCP with OLIA (measurement) | optimal with probing cost (theory) |
|---|---|---|---|---|
| $N_1$=10 | $x_1+x_2$ | 0.96 | 0.96 | 1 |
| $N_2$=10 | y | 0.7 | 0.86 | 0.94 |
| $N_1$=30 | $x_1+x_2$ | 0.96 | 0.96 | 1 |
| $N_2$=10 | y | 0.4 | 0.75 | 0.8 |

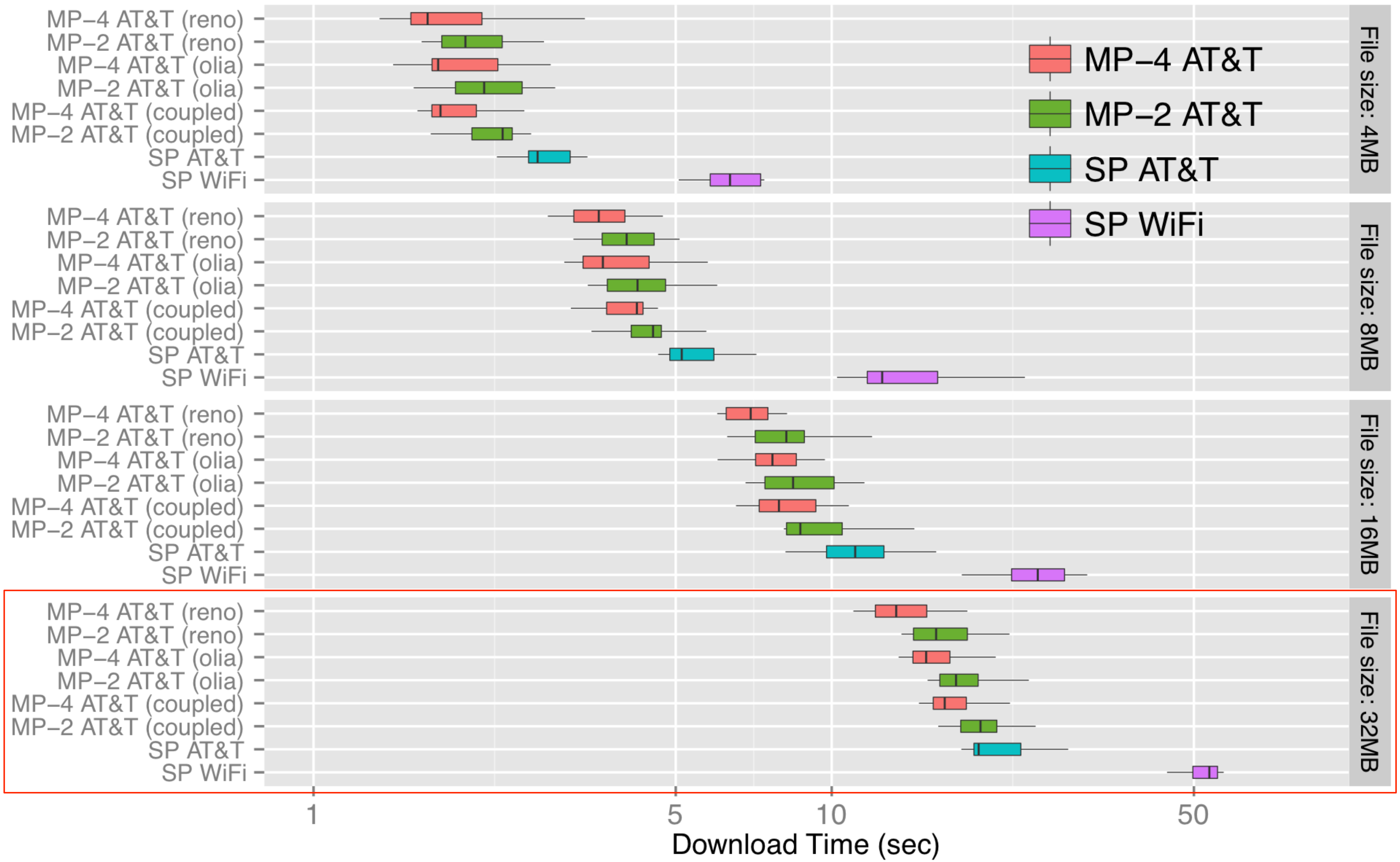


# Is OLIA responsive and non-flappy?

- multiple examples in [CoNEXT 12]
- nothing is better than measurements in real world



# Download time [ACM IMC 13]



# Summary

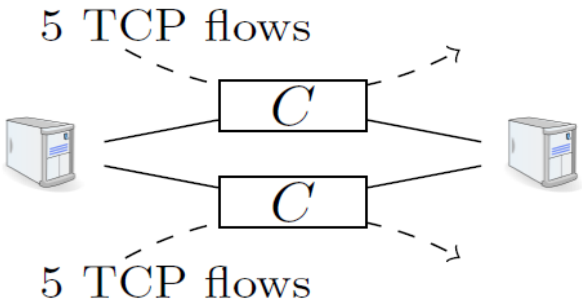
- MPTCP with LIA suffers from important performance problems
- these problems can be mitigated in practice
- OLIA outperforms LIA in all scenarios we studied
- Question: shouldn't we set OLIA as the default congestion control of MPTCP?

# References

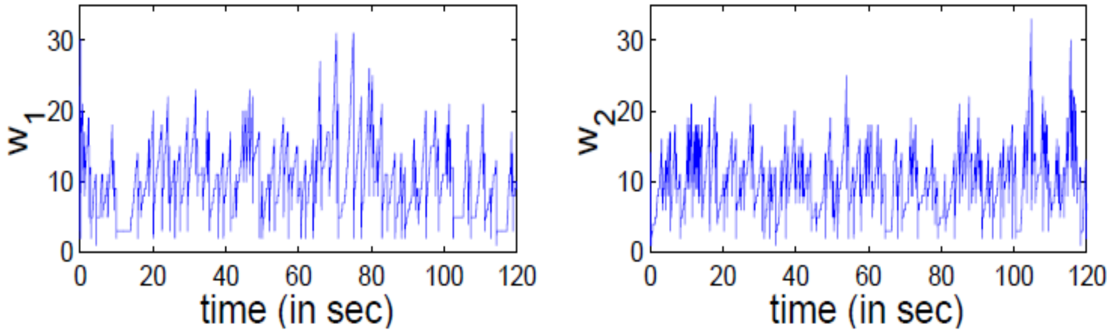

- [RFC 6356]: C. Raiciu, M. Handly, and D. Wischik. “Coupled congestion control for multipath transport protocols”. 2011
- [Kelly, Voice 05]: F. Kelly and T. Voice. “Stability of end-to-end algorithms for joint routing and rate control”. ACM SIGCOMM CCR, 35, 2005.
- [CoNEXT 12]: R. Khalili, N. Gast, M. Popovic, U. Upadhyay, and J.-Y. Le Boudec. “Non pareto-optimality of mptcp: Performance issues and a possible solution”. ACM CoNEXT 2012 (best paper).
- [IMC 13]: Y.-C. Chih, Y.-S. Lim, R. J. Gibbens, E. Nahum, R. Khalili, and D. Towsley. " A Measurement-based Study of MultiPath TCP Performance over Wireless Networks”, accepted at ACM IMC 2013.

# BACK UP SLIDES

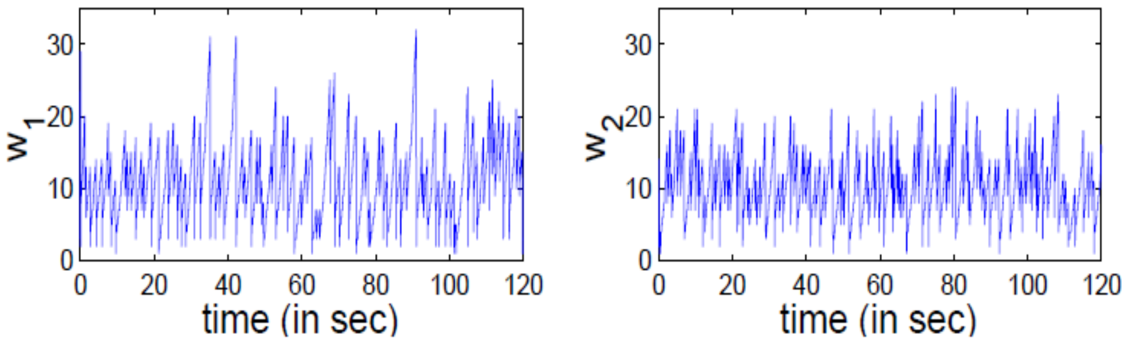
# An illustrative example of OLIA's behavior symmetric scenario




both paths are equally good



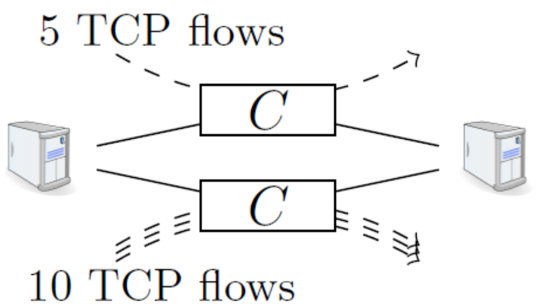

MPTCP with OLIA



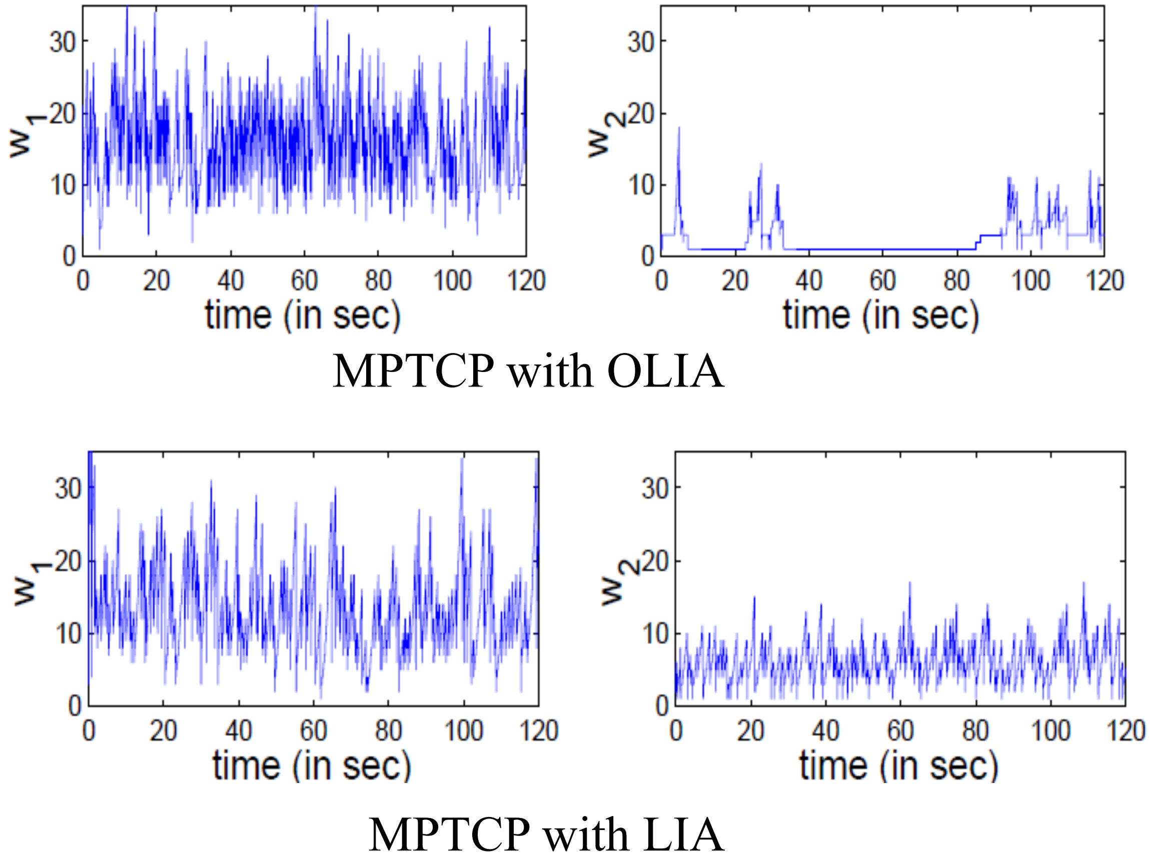

MPTCP with LIA

OLIA uses both paths; it is non-flappy and responsive

# An illustrative example of OLIA's behavior asymmetric scenario

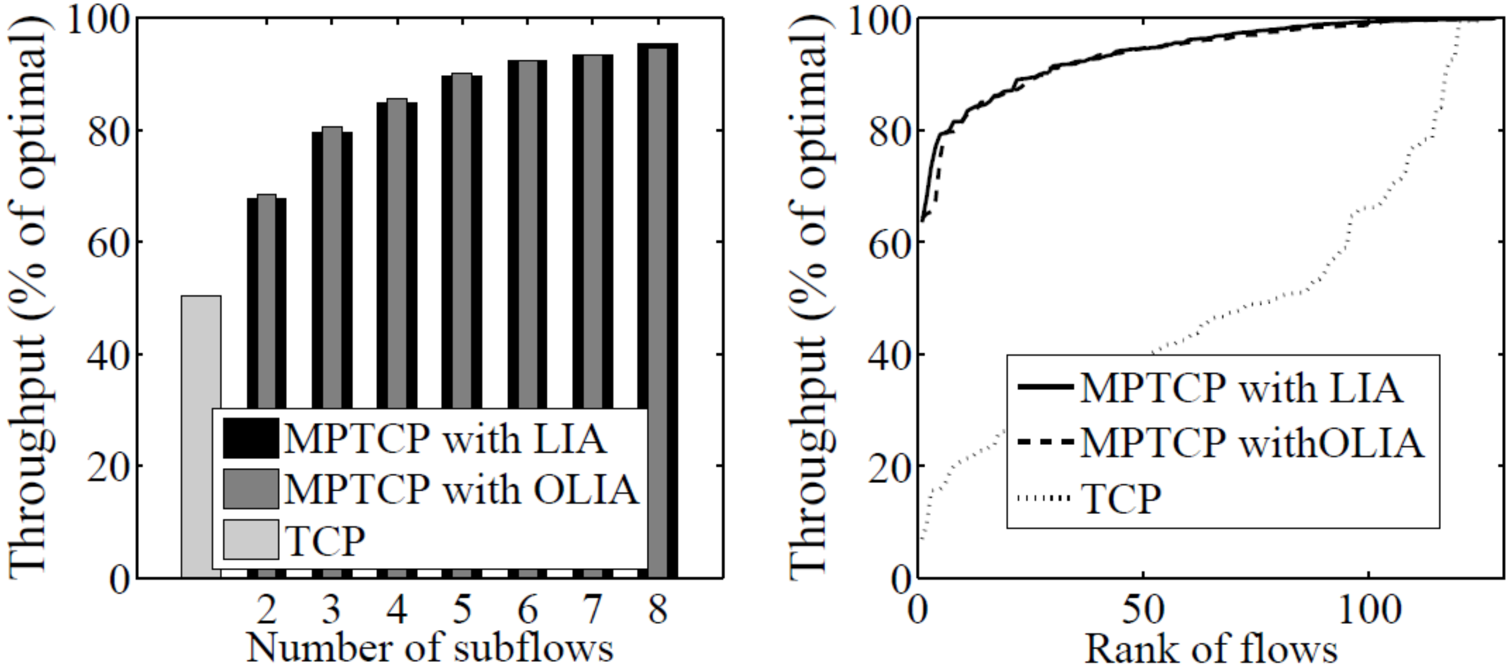



second path is congested




OLIA uses only the first one; it balances the congestion

# Static fat-tree topology: OLIA explores path diversity and show no flappiness

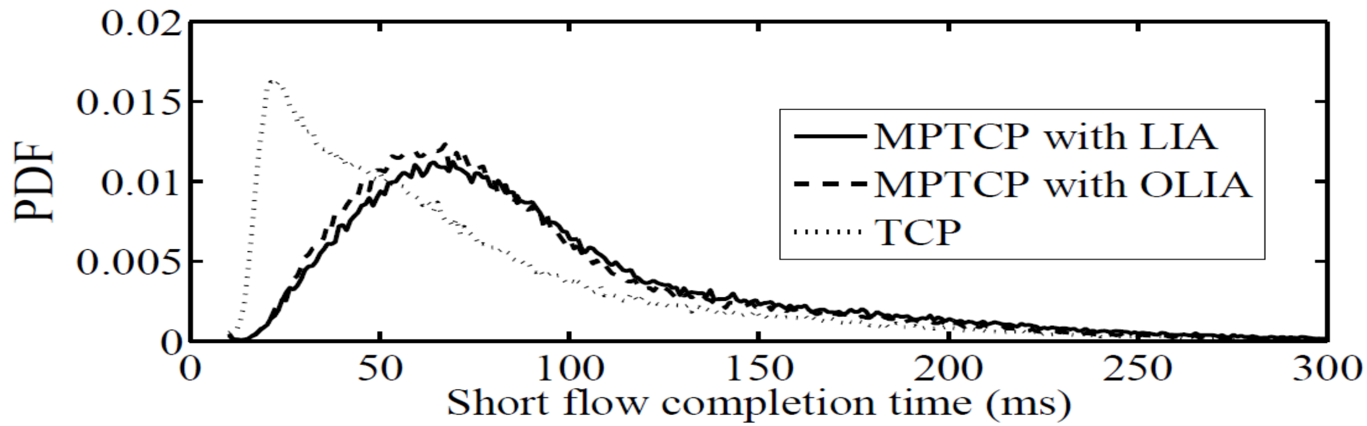



(a) Aggregated throughput. (b) Throughput of users.

a data center with fat-tree topology (similarly to what studied at [MPTCP-Sigcomm 2011])

# Highly dynamic setting with short flows

| algorithm | Short flow finish time (mean/stdev) | Network core utilization |
|---|---|---|
| MPTCP - LIA | $98 \pm 57$ ms | 63.2% |
| MPTCP - OLIA | $90 \pm 42$ ms | 63% |
| Regular TCP | $73 \pm 57$ ms | 39.3% |




4:1 oversubscribed fat-tree; 1/3 of flows are long flows and 2/3 are short flows (similarly to [MPTCP-Sigcomm 2011])