Performance issues with MPTCP

R. Khalili, LCA2, EPFL, Switzerland

R.khalili, N. Gast, M. Popovic, J.-Y Le Boudec, "Performance Issues with MPTCP", submitted draft (draft-khalili-mptcp-performance-issues-00).

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Measurement-based study

focus on congestion control part of MPTCP (RFC 6356)

- outline:
  1. examples of performance issues
  2. can this problem be fixed?

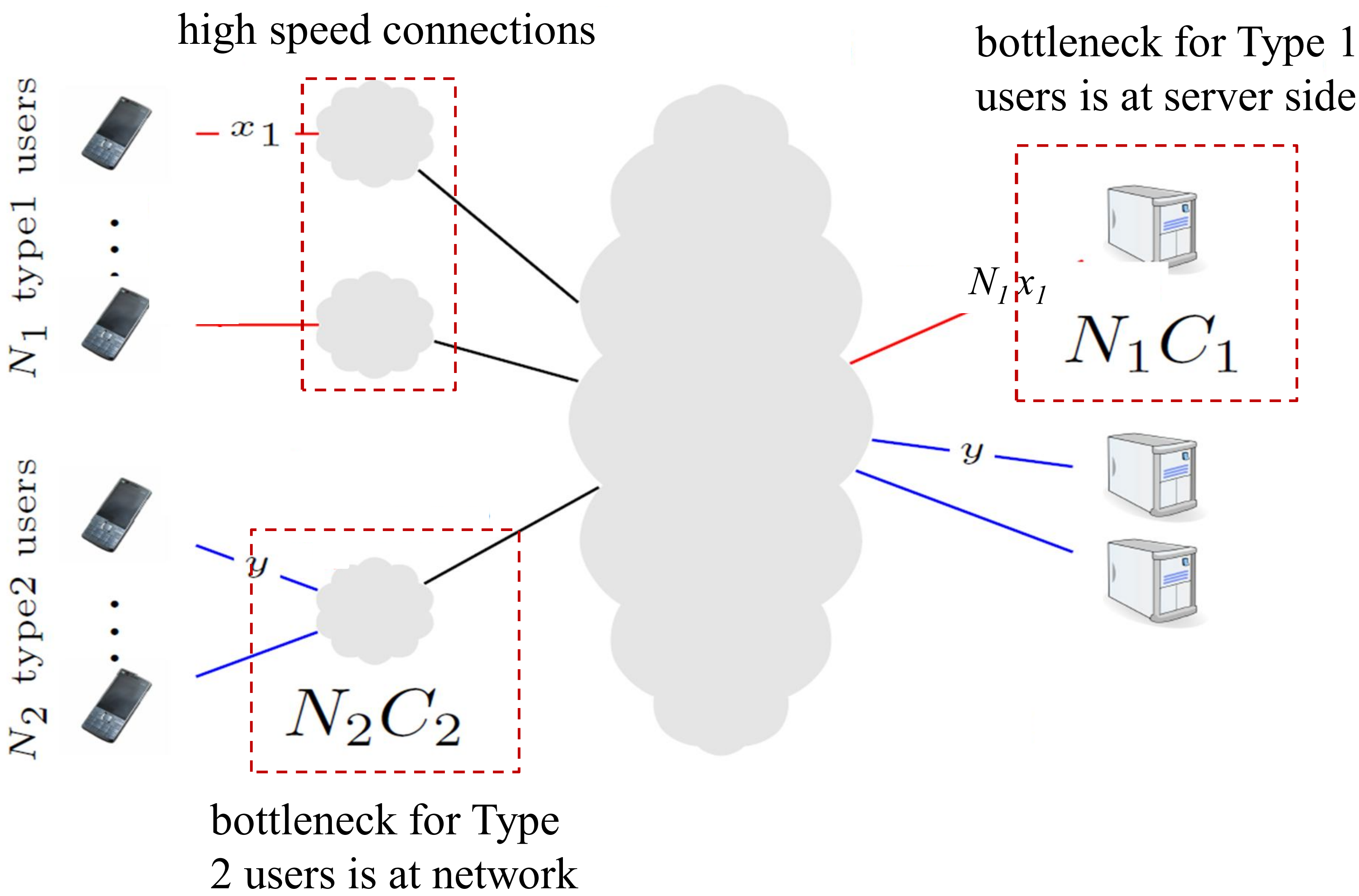
LIA (RFC 6356): "Linked Increases" Algorithm

- adhoc design based on 3 goals
  1. improve throughput: total throughput ≥ 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 stated in RFC 6356, LIA does not fully satisfy goal 3

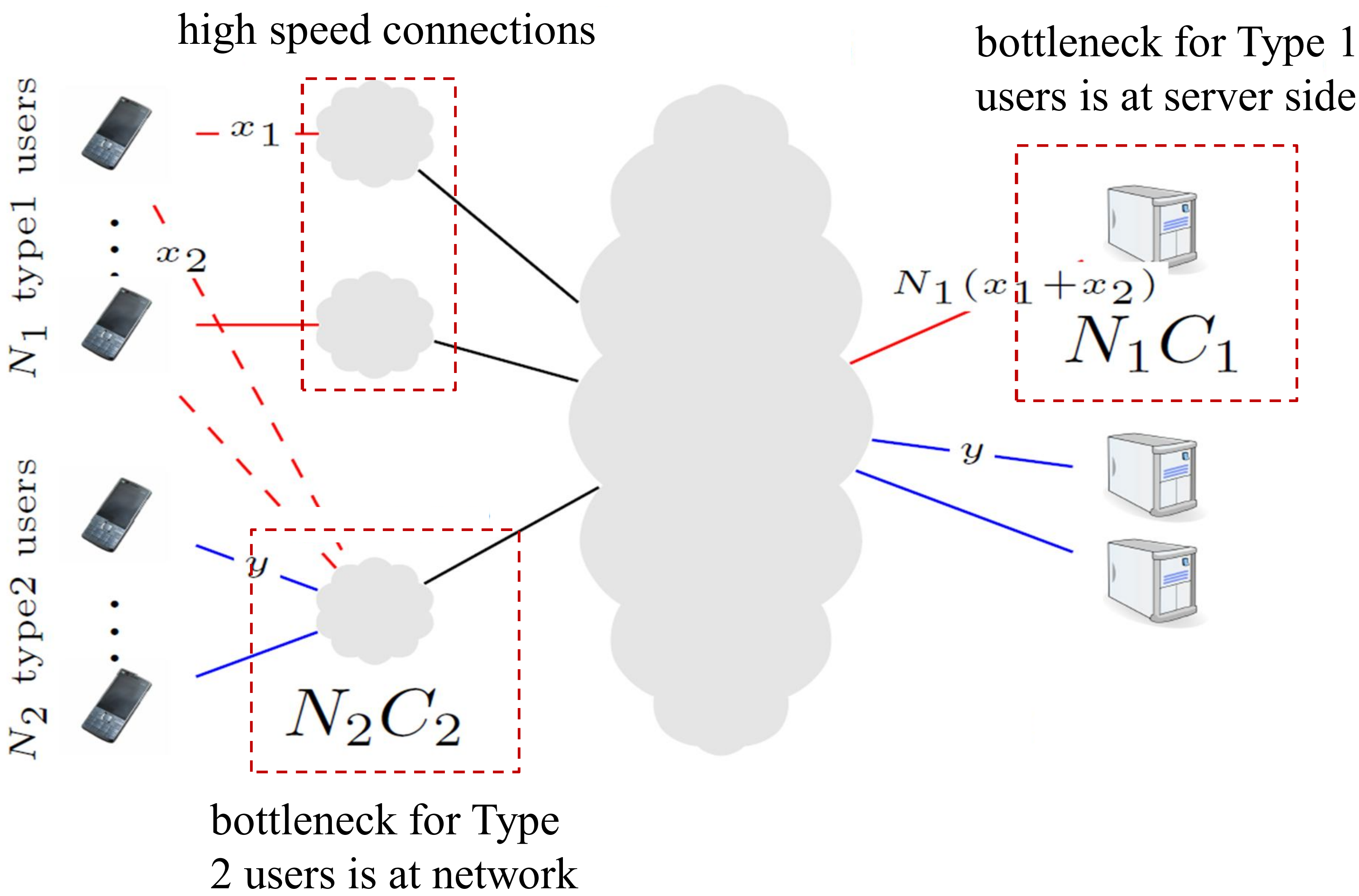
upgrading some TCP users to MPTCP can reduce the throughput of others without any benefit to the upgraded users

MPTCP CAN PENALIZE USERS

Scenario A: MPTCP can penalize TCP users



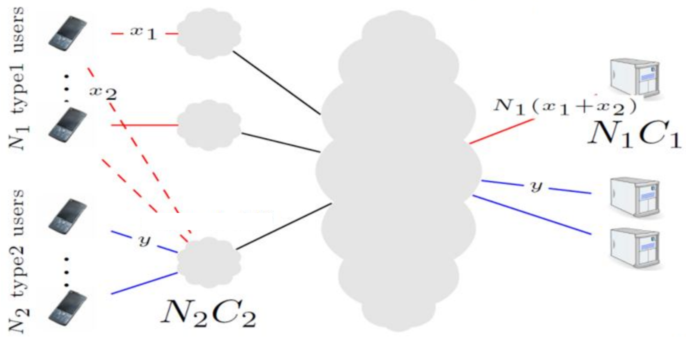
Scenario A: MPTCP can penalize TCP users



Throughput of type 2 users reduced without any benefit for type 1 users

C_1=C_2=1 Mbps

		Type1 users are single path (measurement)	Type1 users are multipath		
			MPTCP (meas.)		
N1=10	type1	0.98	0.96		
N2=10	type2	0.98	0.70		
N1=30	type1	0.98	0.98		
N2=10	type2	0.98	0.44		



We compare MPTCP with two theoretical baselines

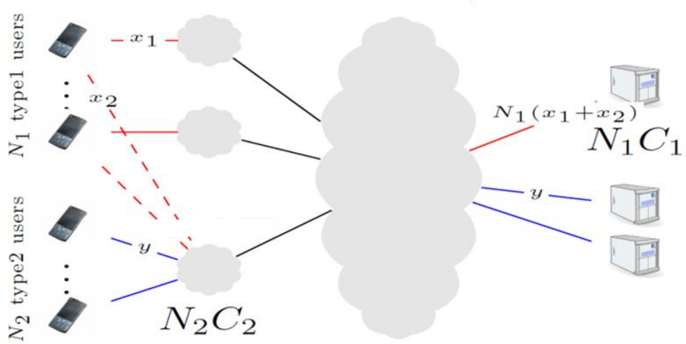
1. optimal algorithm (without probing cost): theoretical optimal load balancing [Kelly,Voice 2005]

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

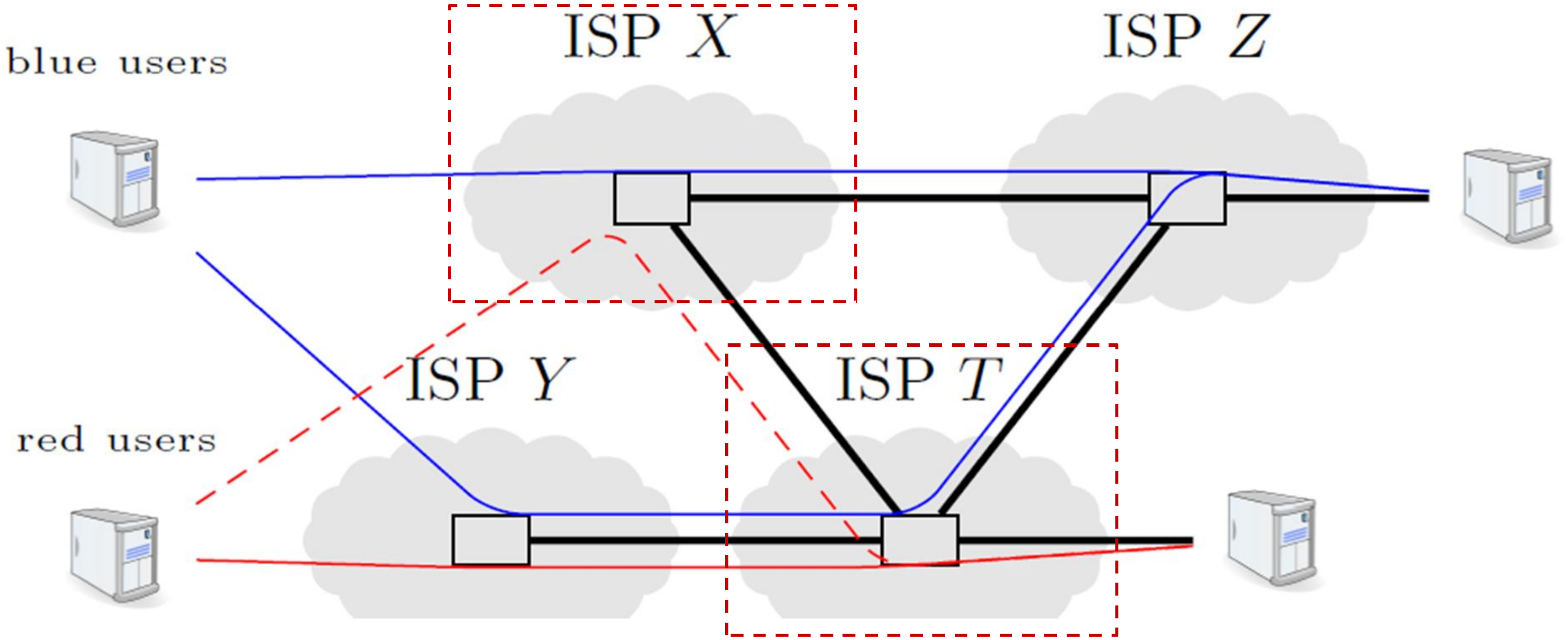
$C_1=C_2=1$ Mbps

		Type1 users are single path (measurement)	Type1 users are multipath: MPTCP (meas.)	Type1 users are multipath: optimal algorithm, with p. cost (theory)	Type1 users are multipath: optimal algorithm, w/out p. cost (theory)
N1=10	type1	0.98	0.96	1	1
N2=10	type2	0.98	0.70	0.94	1
N1=30	type1	0.98	0.98	1	1
N2=10	type2	0.98	0.44	0.8	1



Scenario B: MPTCP can penalize other MPTCP users

bottleneck with capacity C_x

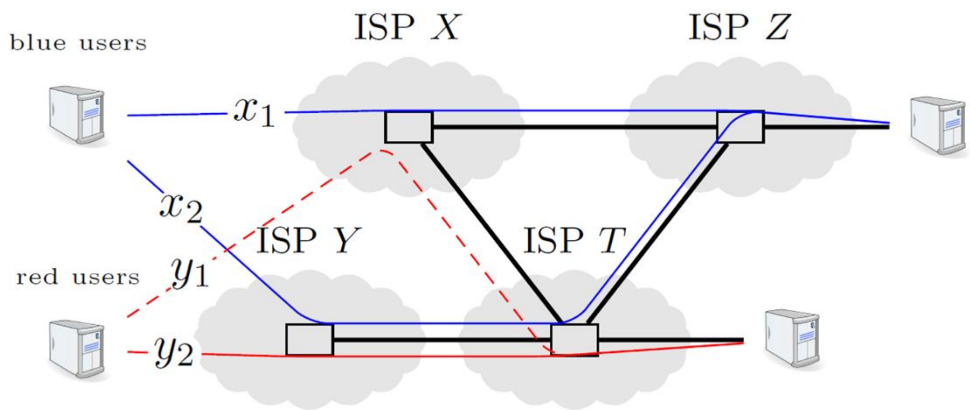


bottleneck with capacity C_T

By upgrading red users to MPTCP, the throughput of everybody decreases

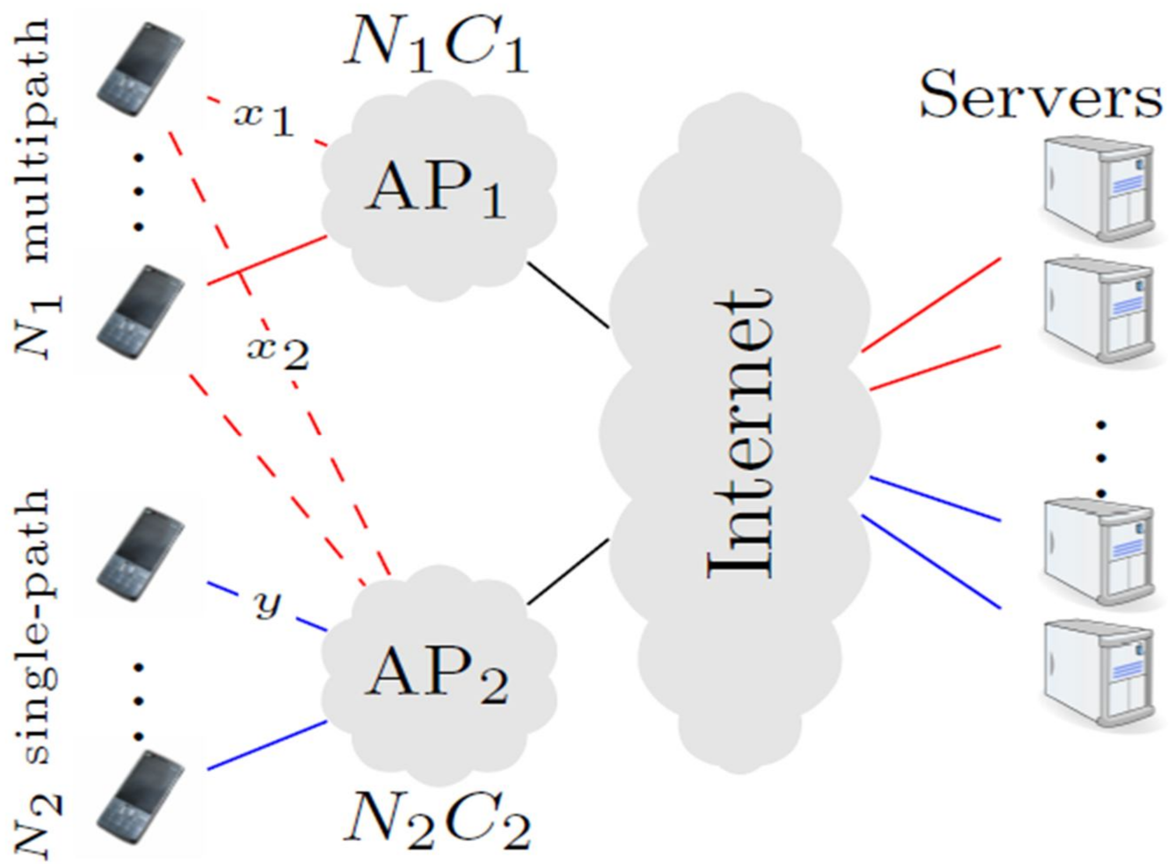
15 blue and 15 red users, C_X=27 and C_T=36 Mbps

	Red users are single-path: Blue users use			Red users are multipath: Blue and Red users use		
	MPTCP (meas.)	optimal algorithm with p. cost (theory)	optimal algorithm w/out p. cost (theory)	MPTCP (meas.)	optimal algorithm with p. cost (theory)	optimal algorithm w/out p. cost (theory)
Red users	1.5	2.1	2.1	1.4	2.04	2.1
Blue users	2.5	2.1	2.1	2.0	2.04	2.1



MPTCP CAN BE EXCESSIVELY AGGRESSIVE TOWARDS TCP USERS

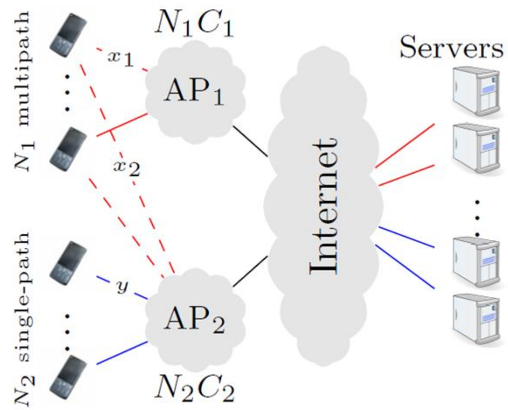
Scenario C: single-path and multipath users share resources



MPTCP is excessively aggressive toward TCP users

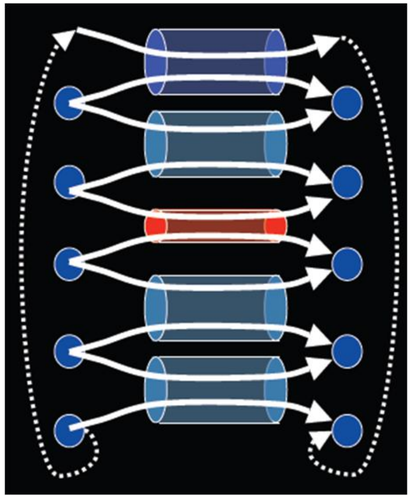
$C_1 = C_2 = 1$ Mbps

		multipath users use MPTCP (measurement)	multipath users use optimal algorithm: with p. cost (theory)	multipath users use optimal algorithm: w/out p. cost (theory)
N1=10	multipath users	1.3	1.04	1
N2=10	single-path users	0.68	0.94	1
N1=30	multipath users	1.19	1.04	1
N2=10	single-path users	0.38	0.8	1

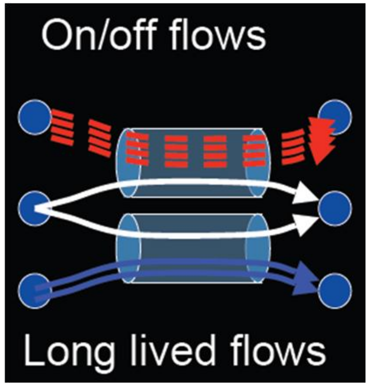


CAN THE SUBOPTIMALITY OF MPTCP WITH LIA BE FIXED?

LIA's design forces tradeoff between responsiveness and load balancing



provide load balancing



be responsive

optimal load balacing but not responsive — $\varepsilon=0$

LIA's implementation (RFC 6356) — $\varepsilon=1$

responsive but bad load balancing — $\varepsilon=2$

ε is a design parameter

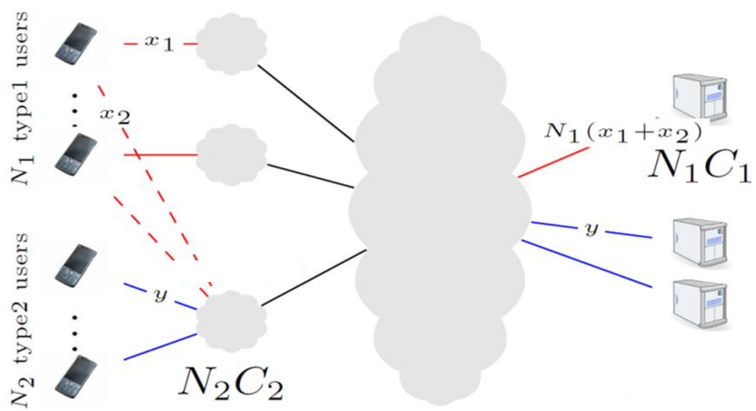
OLIA*: simultaneously provides responsiveness and load balancing

- an adjustment of optimal algorithm [Kelly,Voice 2005]
- adapts windows increases as a function of
  - # of transmitted bytes since last loss to make it responsiveness and non-flappy
  - RTTs of paths to compensate for different RTTs
- implemented in the MPTCP release supported on the Linux kernel 3.0.0

* R. Khalili, N. Gast, M. Popovic, U. Upadhyay, and J.-Y. Le Boudec. "Non Pareto-optimality of mptcp: Performance issues and a possible solution", EPFL Technical report.

Scenario A: OLIA performs close to optimal algorithm with probing cost

		Type1 users are single path	Type1 users are multipath		
			MPTCP w. OLIA [with LIA]	optimal algorithm	
				with p. cost	w/out p. cost
		(measurement)	(measurement)	(theory)	(theory)
N1=10	type1	0.98	0.98 [0.96]	1	1
N2=10	type2	0.98	0.86 [0.70]	0.94	1
N1=30	type1	0.98	0.98 [0.98]	1	1
N2=10	type2	0.98	0.75 [0.44]	0.8	1



Summary

- MPTCP with LIA suffers from important performance problems
- these problems can be mitigated in practice
- suggestion: congestion control part of MPTCP should be revisited by the IETF committees

BACKUP SLIDES

OLIA: Opportunistic "Linked Increases" Algorithm

For a user u, on each path r in R_u:

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

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

optimal congestion balancing → (first term)

responsiveness → (second term)

$$\alpha_r = \begin{cases} \dfrac{1/|\mathcal{R}_u|}{|\mathcal{B}\setminus\mathcal{M}|} & \text{if } r \in \mathcal{B}\setminus\mathcal{M} \neq \emptyset \\ -\dfrac{1/|\mathcal{R}_u|}{|\mathcal{M}|} & \text{if } r \in \mathcal{M} \text{ and } \mathcal{B}\setminus\mathcal{M} \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

"best" paths

$$\mathcal{B}(t) = \left\{ j(t) \,|\, j(t) = \arg\max_{r\in\mathcal{R}_u} \frac{\ell_r(t)}{\text{rtt}_r(t)^2} \right\}$$

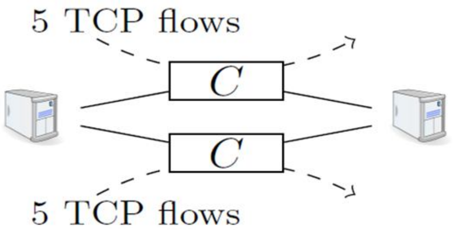
$\ell_r(t)$ = inter-loss #Bytes transmitted

$$\mathcal{M}(t) = \left\{ i(t) \,|\, i(t) = \arg\max_{r\in\mathcal{R}_u} w_r(t) \right\}$$

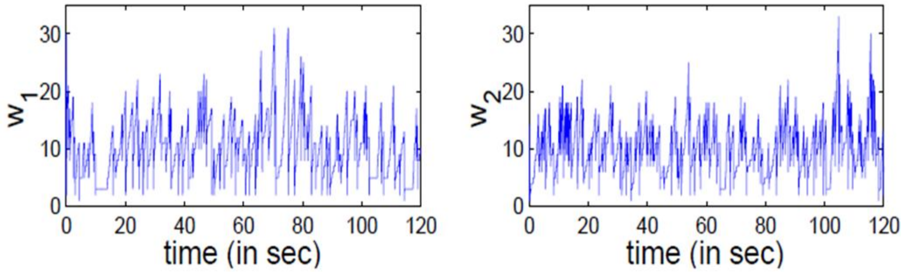
paths with maximum window size

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

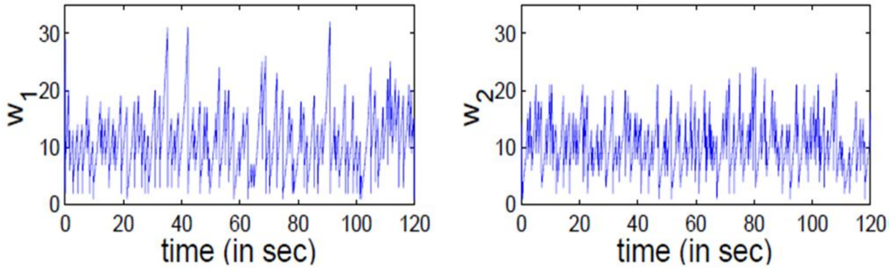
An illustrative example of OLIA's behavior



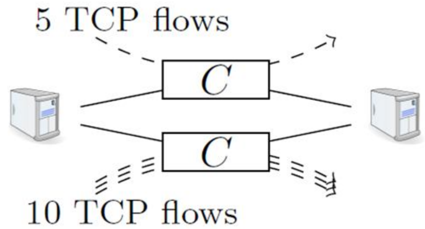
Symetric scenario



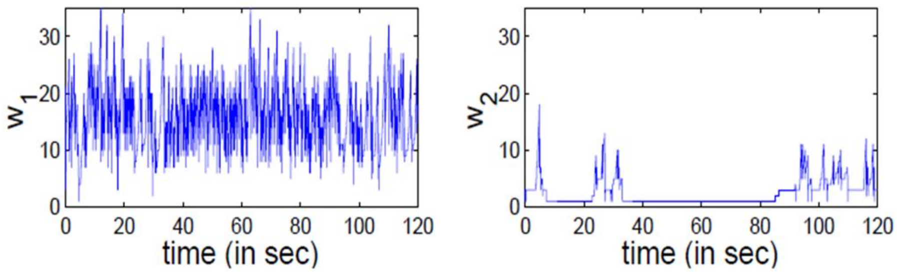
MPTCP with OLIA



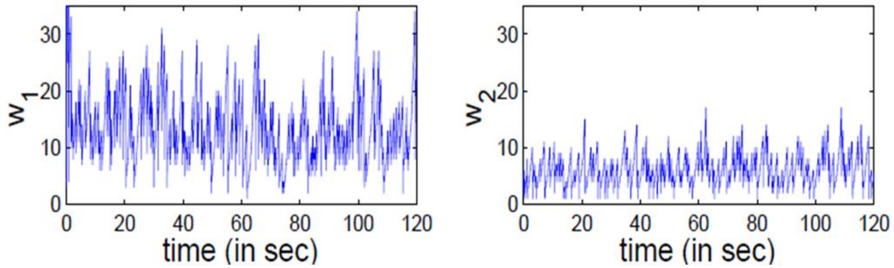
MPTCP with LIA



Asymetric scenario



MPTCP with OLIA



MPTCP with LIA

paths have similar quality, OLIA uses both (non-flappy)

second path is congested, OLIA uses only the first one