

# swift

## the multiparty transport protocol

Victor Grishchenko

# 9 years of BitTorrent

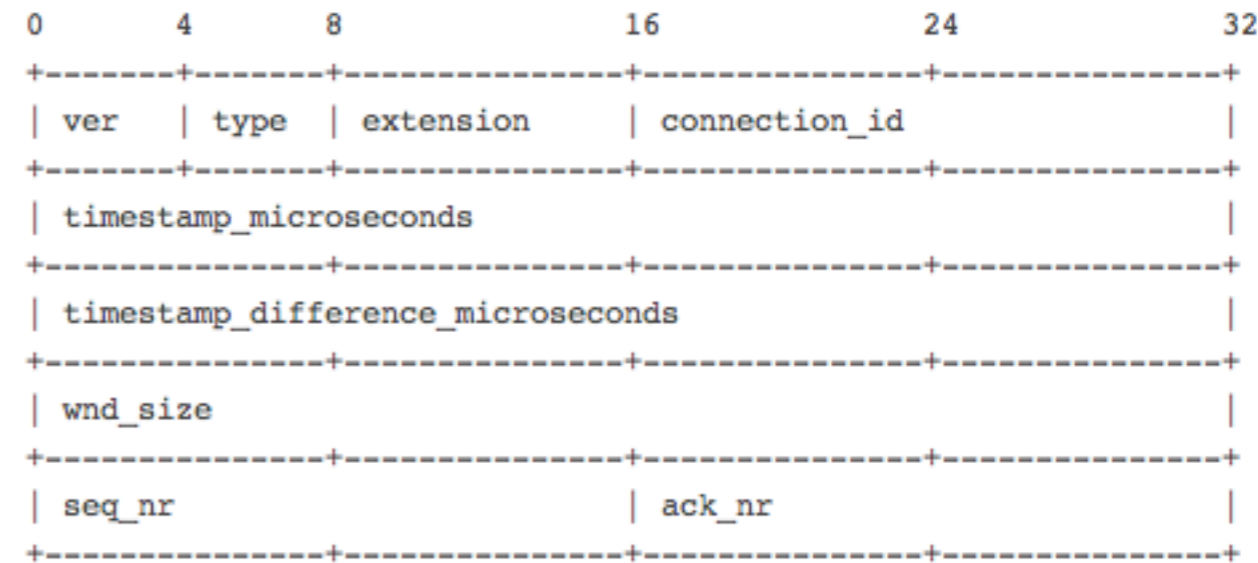
- “PEX won’t work”
- “don’t reinvent TCP/IP”
- tit4tat => rarest | st => out-of-order

# 9 years of BitTorrent

- “PEX ~~won't~~ work’s”
- “don't reinvent TCP/IP”
- tit4tat => rarest | st => out-of-order

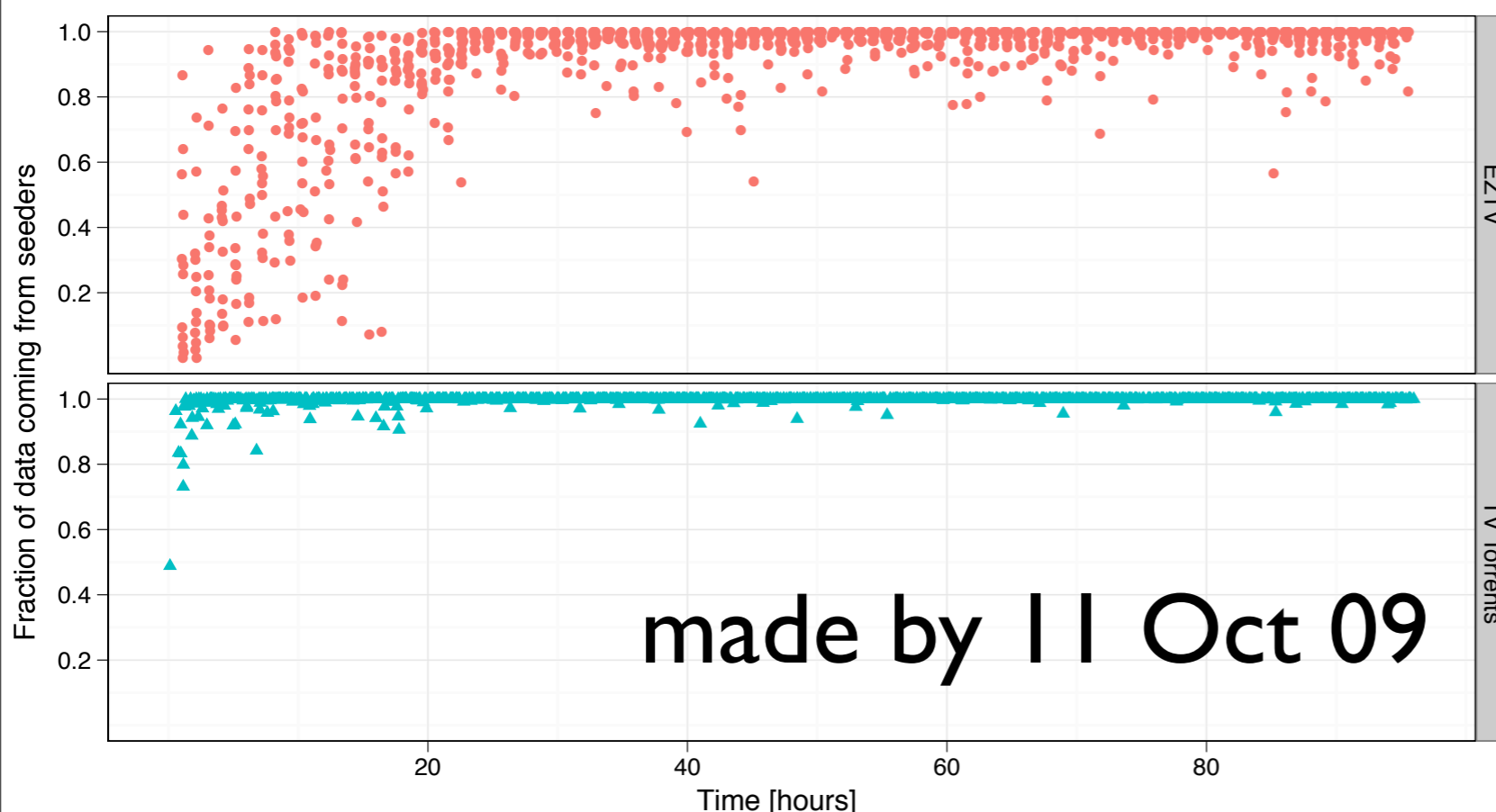
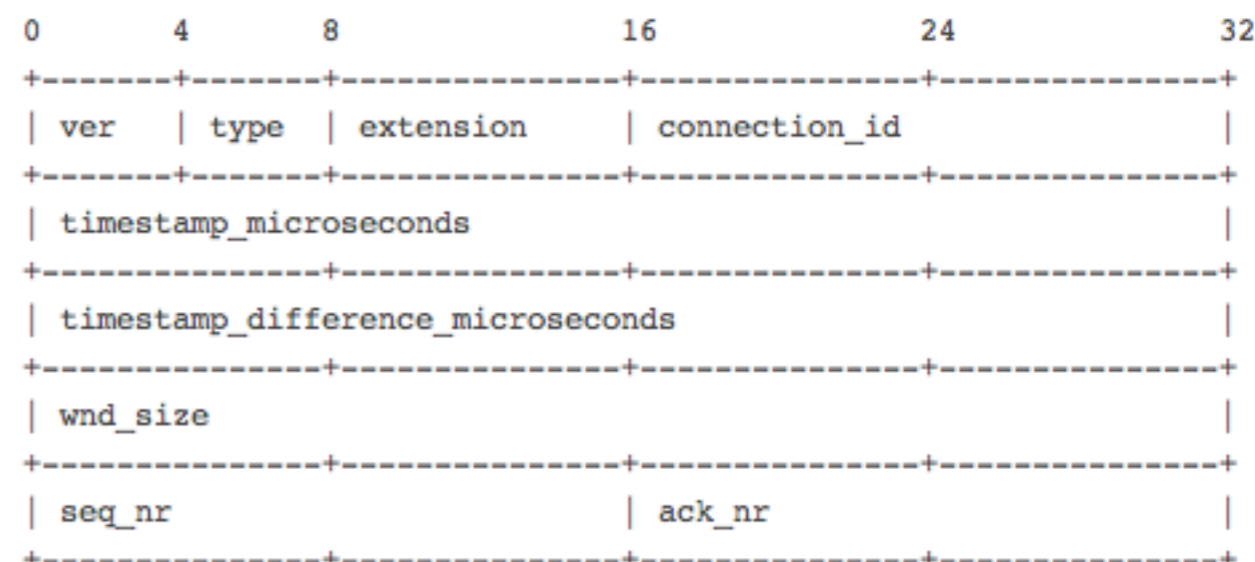
# 9 years of BitTorrent

- “PEX ~~won't~~ work”s
- “don't reinvent -”
- tit4tat => rarest | st => out-of-order



# 9 years of BitTorrent

- “PEX ~~won't~~ work”s
- “don't reinvent -”

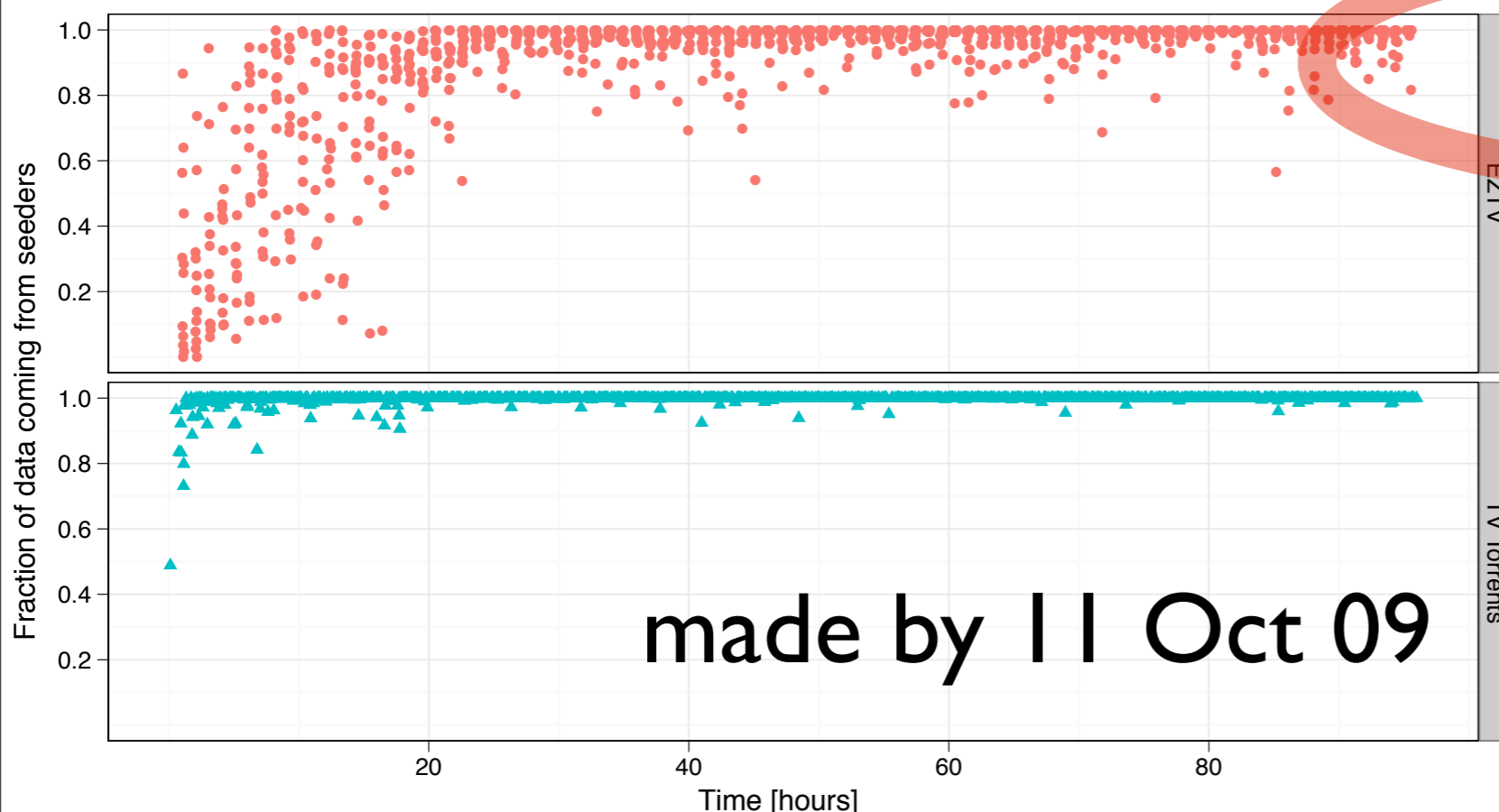
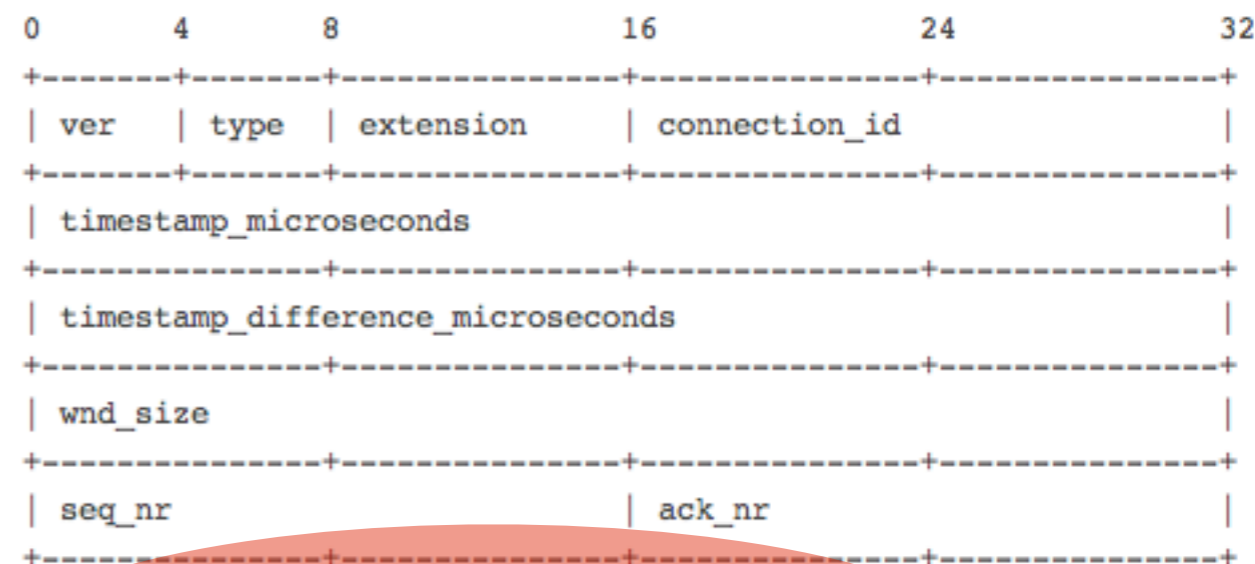


out-of-order

made by | | Oct 09

# 9 years of BitTorrent

- “PEX ~~won't~~ work”s
- “don't reinvent -”



out-of-order

made by I I Oct 09

# BitTorrent + TCP cake

FILES

TORRENT

PIECES

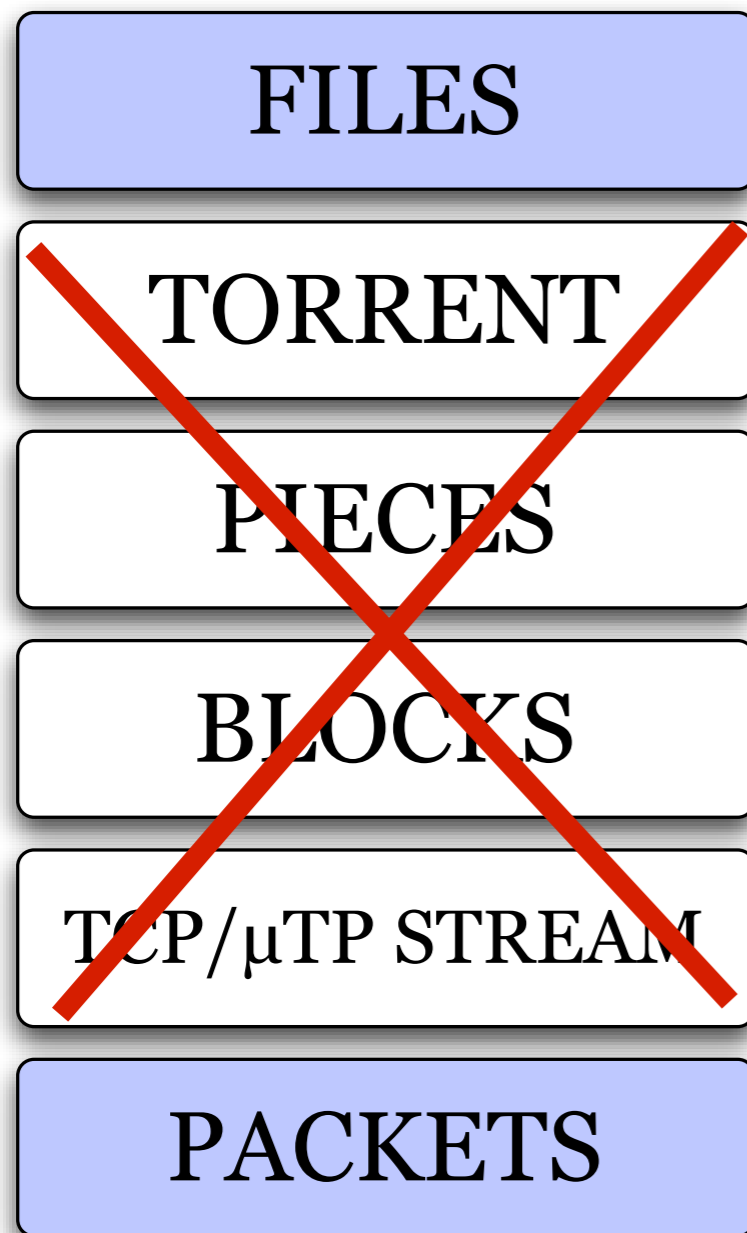
BLOCKS

TCP/ $\mu$ TP STREAM

PACKETS

- historical layering
- too generic but inflexible
- stream abstraction is heavy
- BT loop rides TCP loop

# BitTorrent + TCP cake



- historical layering
- too generic but inflexible
- stream abstraction is heavy
- BT loop rides TCP loop



# Objectives

- embeddable transport protocol
- data, VoD and live streaming
- NAT hole punching
- non-intrusive (esp. seeding)
- content-addressable

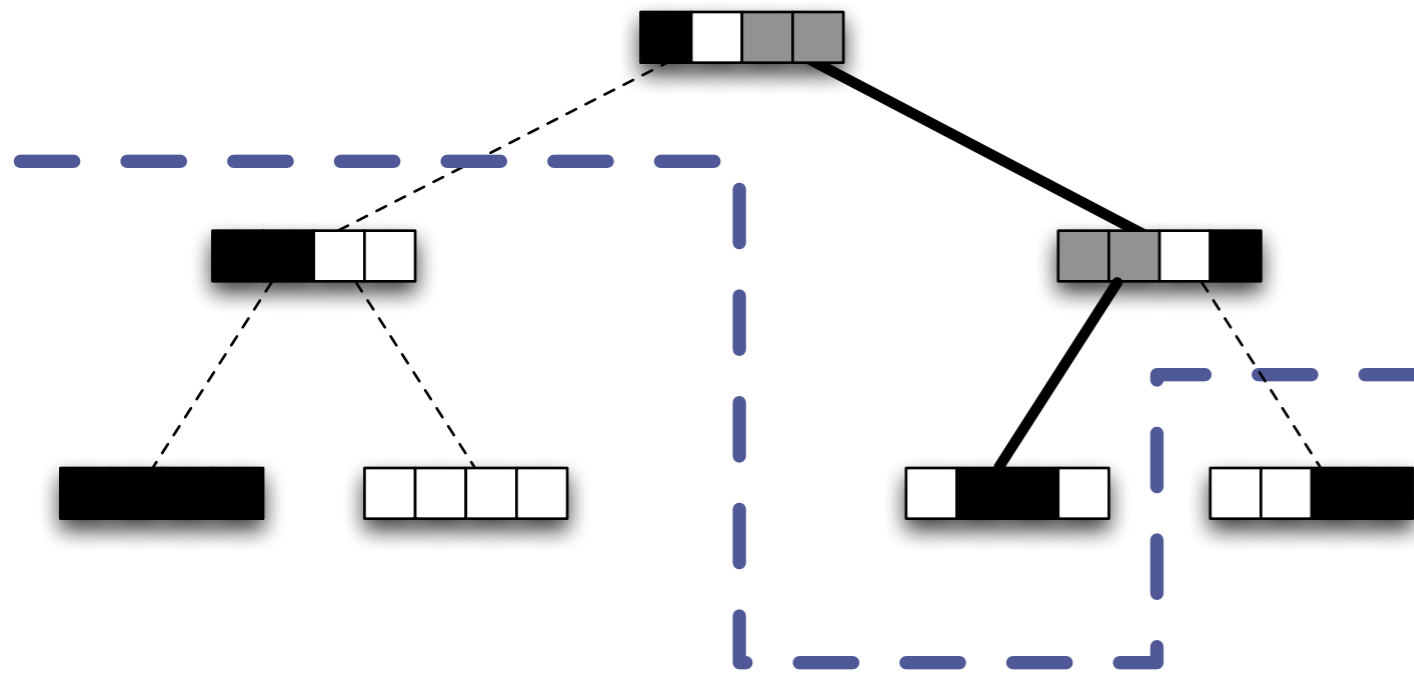
# *swift*: timeline

- Nov 08: first sketches
- Dec 08: work starts
- Jan 09: first code
- Apr 09: works “in a lab”
- Sep 09: LEDBAT implementation
- Oct/Nov: real-network debugging
- Dec 09: “covert” trial with BBC  
*dreams, dreams:*
- Mar 10: VoD runs in-the-wild
- Jun 10: live streaming “in a lab”

# Architecture

*high churn, high redundancy*

- messages (unreliable anyway)
- generic-scale units (bins)
- Merkle hash trees (packet-level)
- based on atomic (unreliable) datagrams
- LEDBAT

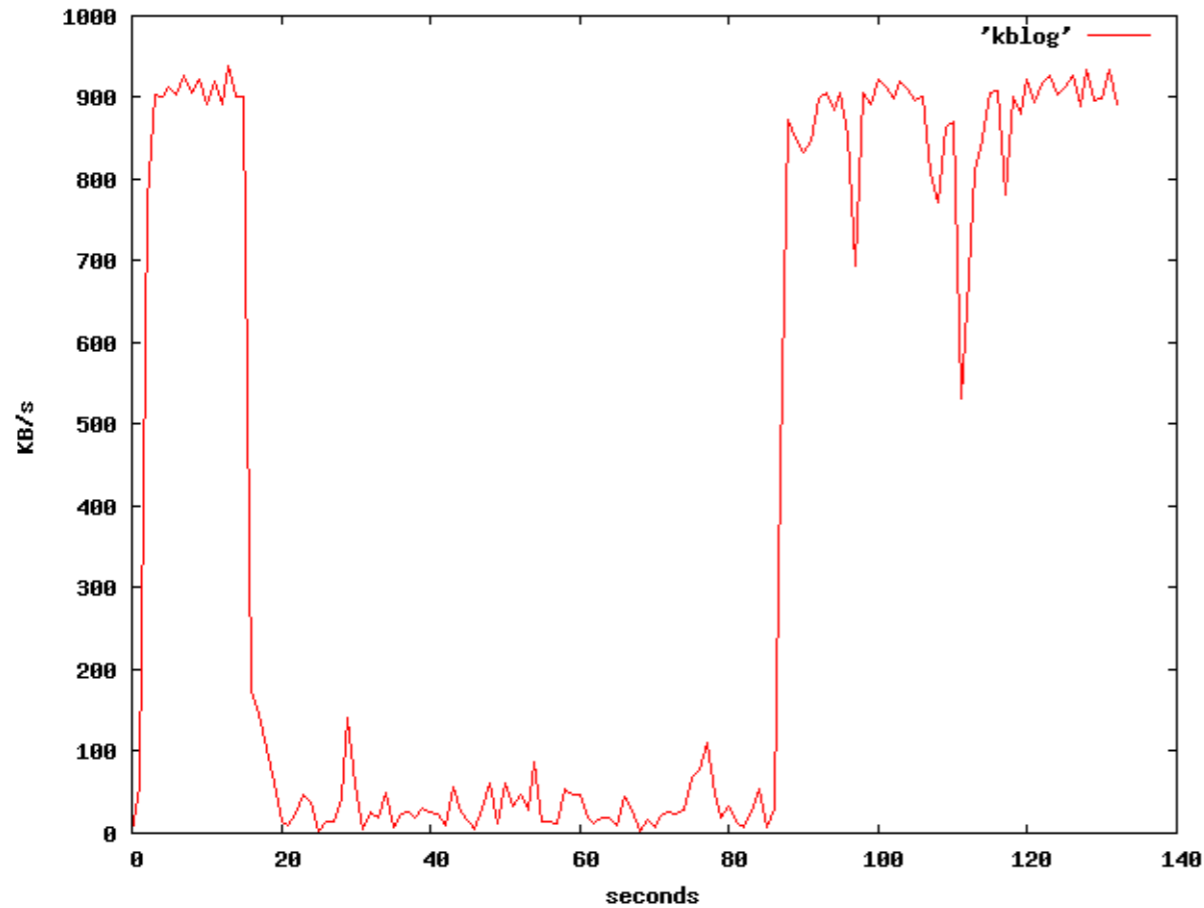


# Binmaps

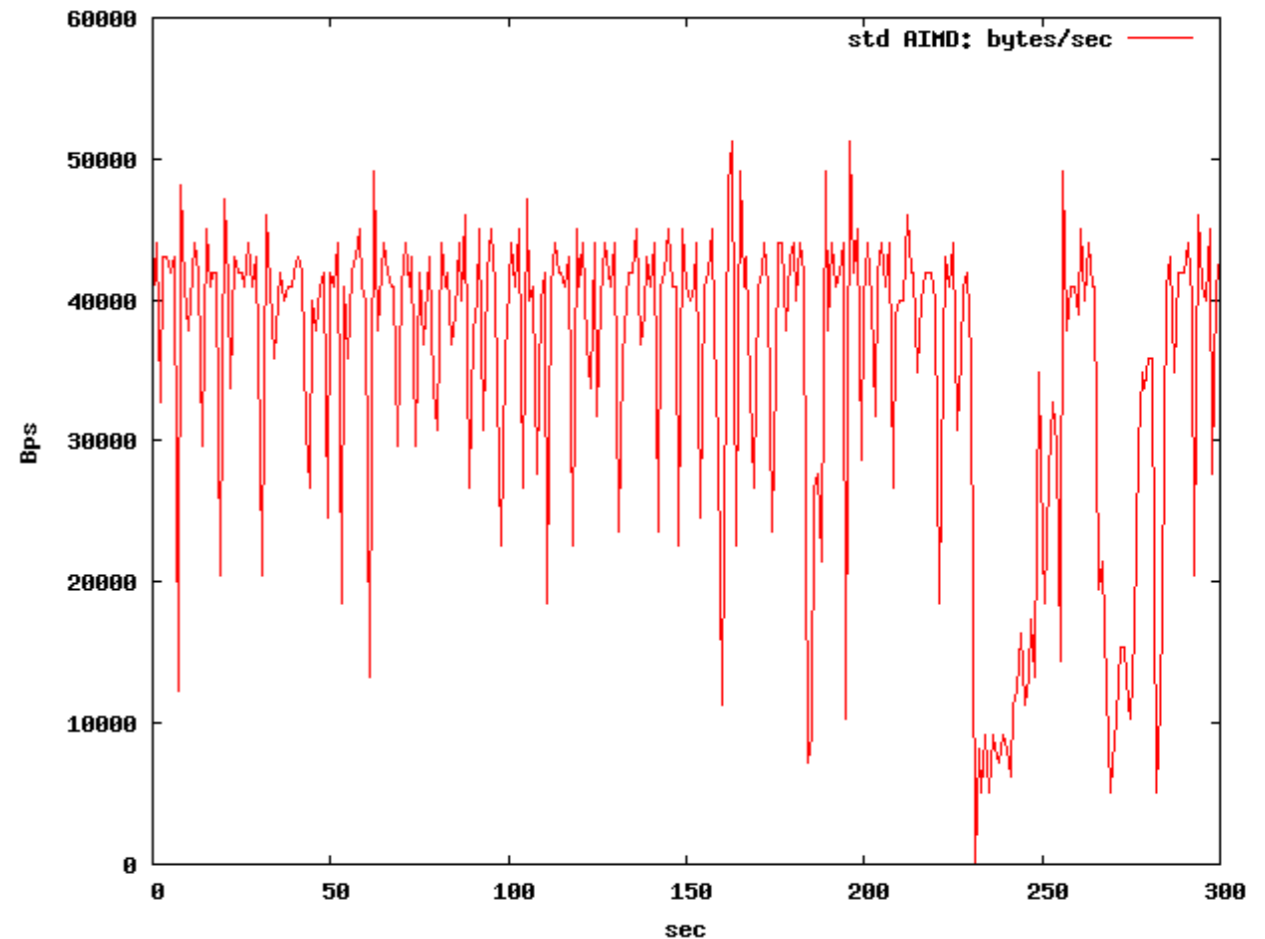
- $f: \text{bin} \rightarrow \{0, 1, \text{mixed}\}$
- bitmap + binary tree
- ext4: x2 worst, x0.2 avg, x0.02 best
- keep own and peer's state
- finer-grained, unified

# Congestion control

LEDBAT: yields to TCP



TCP-like AIMD



# Manifold testing

- Problem: “works for Holland, doesn’t work for Russia“
- Needs thorough testing
- Mix of real and emulated servers,  $N \sim 30$
- Combinations/coverage:  $O(N^2)$

from	to 200	to 201	to 202	to 203	to 204	to 205	to 206	to 208	to 209	to 210	to 212	to 213	to 214	to 215
@200	0	10M	4M	4M	3M	4M	2M	3M	3M	2M	2M	1M	1M	980K
@201	189K	1K	5M	6M	4M	5M	3M	3M	2M	1M	1M	952K	818K	544K
@202	117K	169K	942	83K	1M	1M	1M	1M	1M	1M	1M	929K	858K	2M
@203	114K	175K	3K	1K	80K	261K	1M	1019K	1M	2M	1M	1M	798K	668K
@204	104K	148K	63K	2K	1K	92K	363K	320K	949K	759K	1M	1M	909K	1M
@205	108K	156K	41K	9K	3K	1K	84K	624K	1M	1M	1M	1M	969K	802K
@206	83K	124K	65K	58K	115K	3K	1K	41K	550K	399K	378K	406K	1M	735K
@208	92K	136K	96K	37K	631K	22K	1K	1K	115K	905K	1M	1M	3M	809K
@209	93K	66K	62K	51K	57K	46K	768K	4K	1K	180K	314K	889K	684K	1M
@210	89K	40K	54K	66K	56K	49K	29K	32K	6K	0	68K	813K	749K	494K
@212	78K	37K	38K	53K	61K	54K	492K	58K	71K	2K	3K	701K	10K	807K
@213	59K	33K	33K	41K	43K	62K	14K	37K	33K	31K	33K	0	140K	348K
@214	46K	28K	30K	28K	32K	34K	49K	111K	24K	27K	750	5K	1K	167K
@215	35K	19K	81K	23K	52K	28K	26K	28K	68K	17K	28K	13K	6K	1K

# Contact details

- <http://github.com/gritzko/swift>
- <http://libswift.org>
- <http://tribler.org>
- <http://p2p-next.org>