

# Multipath TCP Congestion Control

Costin Raiciu, Mark Handley and Damon Wischik  
draft-ietf-mptcp-congestion-06

# Status

- draft 05 approved by IESG for publication (with comments)
- draft 06 clarifies the extent to which the goals are satisfied, in response to comments by David Black
- Done!

# Multipath TCP Specification

Mark Handley

on behalf of Alan Ford, Costin Raiciu and Olivier  
Bonaventure

draft-ietf-mptcp-multiaddressed-04

# Status

- Close to final.
  - Reviewed by Andrew McGregor and Stein Gjessing

# Clarifications

- Section 2 rewritten to make overall behaviour clearer before diving into detail.

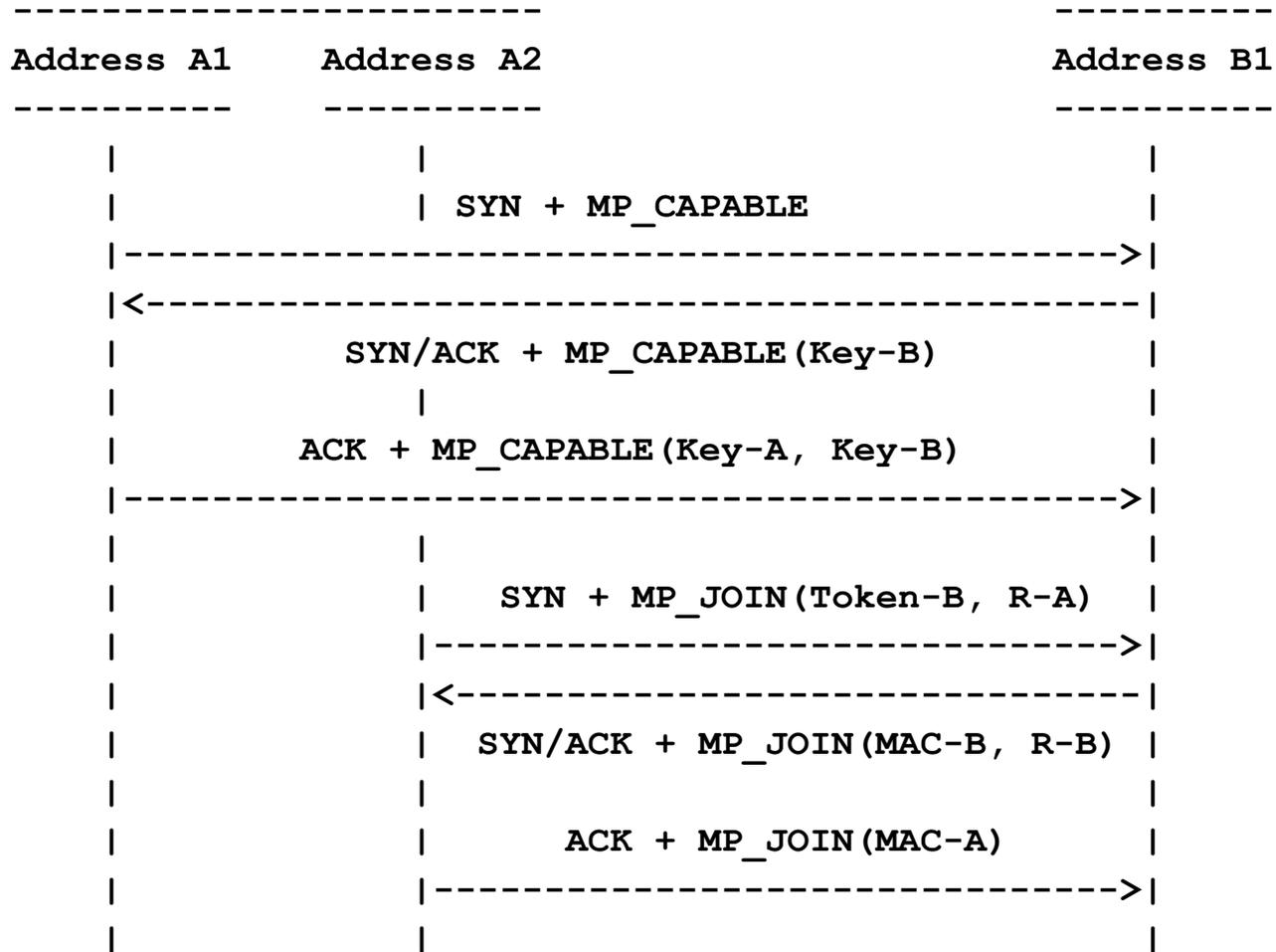
# Refresh of ADD\_ADDR

*“From the connection initiator's point of view, if an MP\_JOIN fails, it SHOULD NOT attempt to connect to the same IP address and port during the lifetime of the connection, unless the other host refreshes the information with another ADD\_ADDR option.”*

Previously needed to REMOVE\_ADDR first, now ADD\_ADDR is a refresh.

Host A

Host B



MAC-A = MAC (Key= (Key-A+Key-B) , Msg= (R-A+R-B) )

MAC-B = MAC (Key= (Key-B+Key-A) , Msg= (R-B+R-A) )

# Break before make

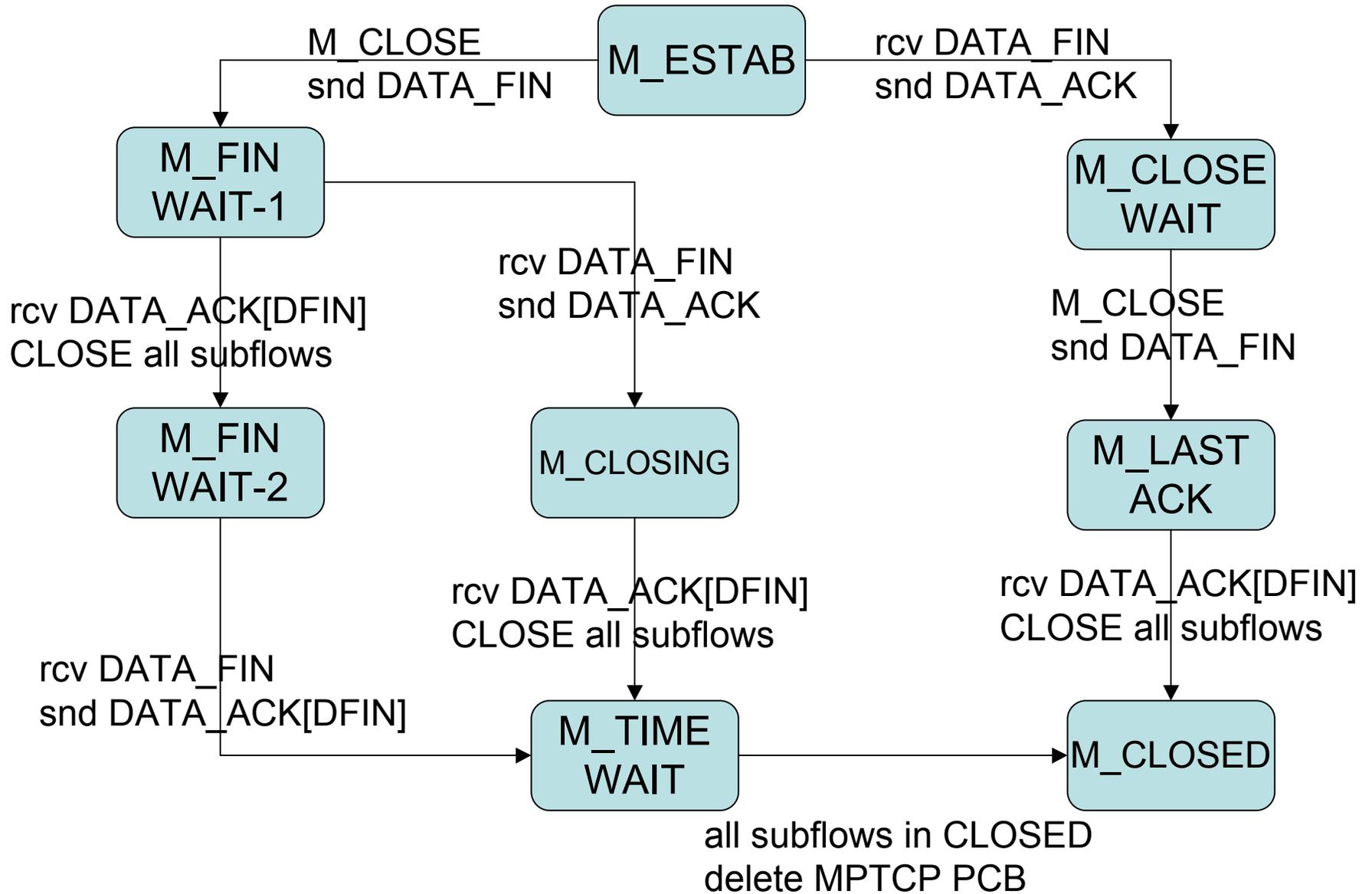
- Long discussion on the list
  - MP\_PRIO gained optional address ID.
  - Clarified relationship between when a connection finishes and when subflows die or finish (state machine).

# MP\_PRIO

- Subflows have a B bit indicating only use as backup.
- B=1 => don't use if any other B=0 subflow is working.
- What happens if the B=0 subflow dies in a break-before-make scenario?
  - Want to set B=1 on stalled subflow, not RST it, because it may recover later.
  - Need to be able to change MP\_PRIO on a different subflow.

# Closedown FSM

- Key issue is making sure all data is reliably received, even if you have to resend it on another subflow, and DATA\_FIN is received.
  - When precisely do you FIN the subflows?
  - Don't want to send last data on subflow A, FIN subflow B, then A fails before data is acked.



# Open Issues

- Mobility in fallback.
- Teardown of state when all subflows fail.

# Mobility in Fallback

*“When a connection is in fallback mode, only one subflow can send data at a time. Otherwise, the receiver would not know how to reorder the data. However, subflows can be opened and closed as necessary, as long as a single one is active at any point.”*

- This works with make-before-break.
  - If all data on a subflow is acked, can switch to another subflow.
- With break-before-make, don't know what was received.

# Teardown

- When do you finally give up if all subflows have timed out?

# Is MPTCP deployable?

*“Is it still possible to extend TCP?”*

Michio Honda, Yoshifumi Nishida, Costin Raiciu,  
Adam Greenhalgh, Mark Handley, Hideyuki Tokuda  
to appear in IMC 2011

<http://nrg.cs.ucl.ac.uk/mjh/tmp/mboxes.pdf>

# What actually happens to TCP in the wild?

- We studied 142 access networks in 24 countries.
- Ran tests to measure what actually happened to TCP.
  - Are new options actually permitted?
  - Does re-segmentation occur in the network?
  - Are sequence numbers modified?
  - Can you leave holes in the sequence space?
  - Do middleboxes proactively ack?

## Middleboxes and new TCP Options in SYN

Observed Behavior	34343	TCP Port 80	443
<i>Passed</i>	129 (96%)	122 (86%)	133(94%)
<i>Removed</i>	6 (4%)	20 (14%)	9 (6%)
<i>Changed</i>	0 (0%)	0 (0%)	0 (0%)
<i>Error</i>	0 (0%)	0 (0%)	0 (0%)
Total	135 (100%)	142 (100%)	142 (100%)

- Middleboxes that remove unknown options are not so rare, especially on port 80
- No path removed options from data but left them on SYNs.

# What actually happens to TCP in the wild?

- **Rewrote sequence numbers**: 10% of paths (18% on port 80)
  - Presumably to improve initial sequence number randomization
- **Resegmented data**: 3% of paths (13% on port 80)
- **Proxy Ack**: 3% of paths (7% on port 80)
  - Note: all of these paths also removed new options from the SYN

# What actually happens to TCP in the wild?

## Can we leave sequence space holes?

- Data segment after a hole: 5% of paths (10% on port 80) send dup ack from mbox or drop segment.
- Ack data not sent: 26% of paths (33% on port 80) do strange things if you send an ack for data not yet sent.