

# MPTCP Hackathon 2016

curl

Arnaud Dethise  
Jacob Eliat-Eliat

# Curl

- File transfer tool (similar to wget)
- Also has a nice application-level API
- Supports several common file transfer protocols (HTTP, FTP, IMAP, SCP...)

# MPTCP awareness

- MPTCP opens multiple flows as soon as the connexion is established
- Problem : for small transfers, the new subflows are not established in time to be useful
- Change : initiate the communication with a single subflow, and add more if needed



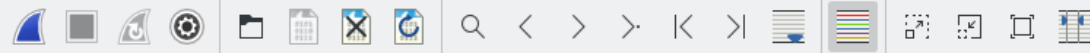
Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Len	Info
1	0.000000	b2:66:75:3c:01:52		ARP	44	Who has 10.0.0.1? Tell 10.0.0.2
2	0.000007	82:69:01:be:f0:18		ARP	44	10.0.0.1 is at 82:69:01:be:f0:18
3	3.438872	10.0.0.1	10.0.0.2	MPTCP	88	41990→80 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM=1 TSval=2278935 TSecr=0 WS=512
4	3.451730	10.0.0.2	10.0.0.1	MPTCP	88	80→41990 [SYN, ACK] Seq=0 Ack=1 Win=28560 Len=0 MSS=1460 SACK_PERM=1 TSval=2278938 TSecr=2278935 WS=512
5	3.451755	10.0.0.1	10.0.0.2	MPTCP	96	41990→80 [ACK] Seq=1 Ack=1 Win=29696 Len=0 TSval=2278940 TSecr=2278938
6	3.451757	10.0.0.1	10.0.0.2	MPTCP	84	[TCP Dup ACK 5#1] 41990→80 [ACK] Seq=1 Ack=1 Win=29696 Len=0 TSval=2278940 TSecr=2278938
7	3.452026	10.0.0.1	10.0.0.2	HTTP	164	GET / HTTP/1.1 [Packet size limited during capture]
8	3.462255	10.0.0.2	10.0.0.1	MPTCP	84	[TCP Window Update] 80→41990 [ACK] Seq=1 Ack=1 Win=28672 Len=0 TSval=2278941 TSecr=2278940
9	3.462643	10.0.0.2	10.0.0.1	MPTCP	76	80→41990 [ACK] Seq=1 Ack=77 Win=28672 Len=0 TSval=2278941 TSecr=2278940
10	3.462673	ae:9b:40:c8:23:a1		ARP	44	Who has 10.0.0.2? Tell 10.0.0.3
11	3.462684	10.0.0.1	10.0.0.4	MPTCP	88	59200→80 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM=1 TSval=2278943 TSecr=0 WS=512
12	3.462692	ae:9b:40:c8:23:a1		ARP	44	Who has 10.0.0.4? Tell 10.0.0.3
13	3.462958	10.0.0.2	10.0.0.1	HTTP	1516	HTTP/1.1 200 OK [Packet size limited during capture]
14	3.462967	10.0.0.1	10.0.0.2	MPTCP	76	41990→80 [ACK] Seq=77 Ack=1429 Win=32256 Len=0 TSval=2278943 TSecr=2278941
15	3.462958	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
16	3.462970	10.0.0.1	10.0.0.2	MPTCP	76	41990→80 [ACK] Seq=77 Ack=2857 Win=35328 Len=0 TSval=2278943 TSecr=2278941
17	3.464107	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
18	3.464115	10.0.0.1	10.0.0.2	MPTCP	76	41990→80 [ACK] Seq=77 Ack=4285 Win=38400 Len=0 TSval=2278943 TSecr=2278941
19	3.465343	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
20	3.465353	10.0.0.1	10.0.0.2	MPTCP	76	41990→80 [ACK] Seq=77 Ack=5713 Win=40960 Len=0 TSval=2278943 TSecr=2278941
21	3.466557	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
22	3.466565	10.0.0.1	10.0.0.2	MPTCP	76	41990→80 [ACK] Seq=77 Ack=7141 Win=44032 Len=0 TSval=2278944 TSecr=2278941
23	3.467792	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
24	3.467801	10.0.0.1	10.0.0.2	MPTCP	76	41990→80 [ACK] Seq=77 Ack=8569 Win=47104 Len=0 TSval=2278944 TSecr=2278941
25	3.469125	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
26	3.469132	10.0.0.1	10.0.0.2	MPTCP	76	41990→80 [ACK] Seq=77 Ack=9997 Win=49664 Len=0 TSval=2278944 TSecr=2278941

- Frame 43: 92 bytes on wire (736 bits), 92 bytes captured (736 bits)  
 - Linux cooked capture  
 - Internet Protocol Version 4, Src: 10.0.0.4, Dst: 10.0.0.1  
 - Transmission Control Protocol, Src Port: 80, Dst Port: 59200, Seq: 0, Ack: 1, Len: 0

```

0000  00 00 00 01 00 06 b2 66 75 3c 01 52 00 00 08 00    .....f u<.R...
0010  45 00 00 4c fb 64 40 00 40 06 2b 43 0a 00 00 04    E..Ld@. @.+C...
0020  0a 00 00 01 00 50 e7 40 e5 03 21 d4 b6 56 77 a6    .....P@ !.!Vw.
0030  e0 12 6f 90 76 c0 00 00 02 04 05 b4 04 02 08 0a    ..o.V... ..
0040  00 22 c6 20 00 22 c6 1f 01 03 03 09 1e 10 10 03    ". " " " " " "
0050  6a aa 62 48 58 2e 0a 8b 44 13 c2 cb                j.bHX... D...
    
```



tcp.stream eq 1 Expression...

No.	Time	Source	Destination	Protocol	Len	Info
11	3.482684	10.0.0.1	10.0.0.4	MPTCP	88	59200→80 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM=1 TSval=2278943 TSecr=0 WS=512
44	3.479338	10.0.0.1	10.0.0.4	MPTCP	92	59200→80 [ACK] Seq=1 Ack=1 Win=101888 Len=0 TSval=2278947 TSecr=2278944
89	3.739136	10.0.0.1	10.0.0.4	MPTCP	76	59200→80 [ACK] Seq=1 Ack=586 Win=162304 Len=0 TSval=2279012 TSecr=2279011
92	3.751635	10.0.0.1	10.0.0.4	MPTCP	76	59200→80 [FIN, ACK] Seq=1 Ack=586 Win=162304 Len=0 TSval=2279015 TSecr=2279011
96	3.761889	10.0.0.1	10.0.0.4	MPTCP	76	59200→80 [ACK] Seq=2 Ack=587 Win=162304 Len=0 TSval=2279017 TSecr=2279016
43	3.479324	10.0.0.4	10.0.0.1	MPTCP	92	80→59200 [SYN, ACK] Seq=0 Ack=1 Win=28560 Len=0 MSS=1460 SACK_PERM=1 TSval=2278944 TSecr=2278943 WS=512
81	3.501985	10.0.0.4	10.0.0.1	MPTCP	76	[TCP Window Update] 80→59200 [ACK] Seq=1 Ack=1 Win=57344 Len=0 TSval=2278948 TSecr=2278947
88	3.739117	10.0.0.4	10.0.0.1	MPTCP	673	80→59200 [PSH, ACK] Seq=1 Ack=1 Win=57344 Len=585 TSval=2279011 TSecr=2278947[Packet size limited during capture]
95	3.761874	10.0.0.4	10.0.0.1	MPTCP	76	80→59200 [FIN, ACK] Seq=586 Ack=2 Win=57344 Len=0 TSval=2279016 TSecr=2279015

> Frame 11: 88 bytes on wire (704 bits), 88 bytes captured (704 bits)  
 > Linux cooked capture  
 > Internet Protocol Version 4, Src: 10.0.0.1, Dst: 10.0.0.4  
 > **Transmission Control Protocol, Src Port: 59200, Dst Port: 80, Seq: 0, Len: 0**

```

0000  00 04 00 01 00 06 82 69 01 be f0 18 00 00 08 00      .....I.....
0010  45 00 00 48 b1 ba 40 00 40 06 74 f1 0a 00 00 01      E..H..@.@.t....
0020  0a 00 00 04 e7 40 00 50 b6 56 77 a5 00 00 00 00      ....@.P.Vw.....
0030  d0 02 72 10 14 3f 00 00 02 04 05 b4 04 02 08 0a      ..f.?.....
0040  00 22 c6 1f 00 00 00 00 01 03 03 09 1e 0c 10 02      ".:.....
0050  d3 c5 3d 2c 34 8f 58 c8                                ..=4.X.

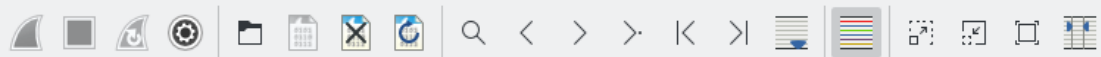
```

# Implementation

- Curl : abstract socket
- Overloaded *recv* function of that abstraction
- Memorizing the amount of data transferred over the socket
- Threshold ?

# Implementation (2)

- Minimum threshold : 16 kB (congestion window)
- Currently set with (configurable) 64 kB
- Only relevant for much larger transfers
- « If over 64 kB, then the transfer will be large »



Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Len	Info
1	0.000000	36:54:26:b3:26:47		ARP	44	Who has 10.0.0.1? Tell 10.0.0.2
2	0.000007	fe:59:85:6b:57:c1		ARP	44	10.0.0.1 is at fe:59:85:6b:57:c1
3	6.268449	10.0.0.1	10.0.0.2	MPTCP	88	41955→80 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM=1 TSval=2089799 TSecr=0 WS=512
4	6.280941	10.0.0.2	10.0.0.1	MPTCP	88	80→41955 [SYN, ACK] Seq=0 Ack=1 Win=28560 Len=0 MSS=1460 SACK_PERM=1 TSval=2089802 TSecr=2089799 WS=512
5	6.280970	10.0.0.1	10.0.0.2	MPTCP	96	41955→80 [ACK] Seq=1 Ack=1 Win=29696 Len=0 TSval=2089803 TSecr=2089802
6	6.281337	10.0.0.1	10.0.0.2	HTTP	164	GET / HTTP/1.1 [Packet size limited during capture]
7	6.301753	10.0.0.2	10.0.0.1	MPTCP	76	80→41955 [ACK] Seq=1 Ack=77 Win=28672 Len=0 TSval=2089805 TSecr=2089803
8	6.301774	10.0.0.2	10.0.0.1	HTTP	1328	HTTP/1.1 200 OK [Packet size limited during capture]
9	6.301787	10.0.0.1	10.0.0.2	MPTCP	76	41955→80 [ACK] Seq=77 Ack=1241 Win=32256 Len=0 TSval=2089808 TSecr=2089805
10	6.301991	10.0.0.1	10.0.0.2	MPTCP	88	41955→80 [FIN, ACK] Seq=77 Ack=1241 Win=32256 Len=0 TSval=2089808 TSecr=2089805
11	6.312145	10.0.0.2	10.0.0.1	MPTCP	88	80→41955 [FIN, ACK] Seq=1241 Ack=78 Win=28672 Len=0 TSval=2089810 TSecr=2089808
12	6.312159	10.0.0.1	10.0.0.2	MPTCP	76	41955→80 [ACK] Seq=78 Ack=1242 Win=32256 Len=0 TSval=2089811 TSecr=2089810
13	6.312163	10.0.0.1	10.0.0.2	MPTCP	76	[TCP Dup ACK 12#1] 41955→80 [ACK] Seq=78 Ack=1242 Win=32256 Len=0 TSval=2089811 TSecr=2089810
14	6.322387	10.0.0.2	10.0.0.1	TCP	56	80→41955 [RST] Seq=1242 Win=0 Len=0

> - Frame 1: 44 bytes on wire (352 bits), 44 bytes captured (352 bits)  
 > - Linux cooked capture  
 > - Address Resolution Protocol (request)

```

0000 00 00 00 01 00 06 36 54 26 b3 26 47 00 00 08 06 .....6T &&G....
0010 00 01 08 00 06 04 00 01 36 54 26 b3 26 47 0a 00 ..... 6T&&G..
0020 00 02 00 00 00 00 00 00 0a 00 00 01 .....
    
```





Apply a display filter ... <Ctrl-/> Expression...

No.	Time	Source	Destination	Protocol	Lenç	Info
1612	6.958002	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1359457 Win=353792 Len=0 TSval=1397849 TSecr=1397845
1613	6.959205	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
1614	6.959213	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1360885 Win=353792 Len=0 TSval=1397849 TSecr=1397845
1615	6.960368	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
1616	6.960376	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1362313 Win=353792 Len=0 TSval=1397849 TSecr=1397846
1617	6.961626	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
1618	6.961635	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1363741 Win=353792 Len=0 TSval=1397850 TSecr=1397847
1619	6.962838	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
1620	6.962847	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1365169 Win=353792 Len=0 TSval=1397850 TSecr=1397847
1621	6.964049	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
1622	6.964058	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1366597 Win=353792 Len=0 TSval=1397850 TSecr=1397848
1623	6.965259	10.0.0.2	10.0.0.3	MPTCP	1516	80→41938 [ACK] Seq=27133 Ack=1 Win=57344 Len=1428 TSval=1397848 TSecr=1397847[Packet size limited during capture]
1624	6.965267	10.0.0.3	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=1 Ack=28561 Win=353792 Len=0 TSval=1397850 TSecr=1397848
1625	6.966472	10.0.0.2	10.0.0.3	MPTCP	1516	80→41938 [ACK] Seq=28561 Ack=1 Win=57344 Len=1428 TSval=1397848 TSecr=1397847[Packet size limited during capture]
1626	6.966481	10.0.0.3	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=1 Ack=29989 Win=353792 Len=0 TSval=1397851 TSecr=1397848
1627	6.967688	10.0.0.2	10.0.0.3	MPTCP	1516	80→41938 [ACK] Seq=29989 Ack=1 Win=57344 Len=1428 TSval=1397849 TSecr=1397847[Packet size limited during capture]
1628	6.967697	10.0.0.3	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=1 Ack=31417 Win=353792 Len=0 TSval=1397851 TSecr=1397849
1629	6.968893	10.0.0.2	10.0.0.3	MPTCP	1516	80→41938 [ACK] Seq=31417 Ack=1 Win=57344 Len=1428 TSval=1397849 TSecr=1397848[Packet size limited during capture]
1630	6.968902	10.0.0.3	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=1 Ack=32845 Win=353792 Len=0 TSval=1397851 TSecr=1397849
1631	6.970105	10.0.0.2	10.0.0.3	MPTCP	1516	80→41938 [ACK] Seq=32845 Ack=1 Win=57344 Len=1428 TSval=1397849 TSecr=1397848[Packet size limited during capture]
1632	6.970113	10.0.0.3	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=1 Ack=34273 Win=353792 Len=0 TSval=1397852 TSecr=1397849
1633	6.971316	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
1634	6.971324	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1368025 Win=353792 Len=0 TSval=1397852 TSecr=1397849
1635	6.972476	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]
1636	6.972484	10.0.0.1	10.0.0.2	MPTCP	76	41938→80 [ACK] Seq=77 Ack=1369453 Win=353792 Len=0 TSval=1397852 TSecr=1397850
1637	6.973738	10.0.0.2	10.0.0.1	HTTP	1516	Continuation[Packet size limited during capture]

> - Frame 1624: 76 bytes on wire (608 bits), 76 bytes captured (608 bits)  
 > - Linux cooked capture  
 > - Internet Protocol Version 4, Src: 10.0.0.3, Dst: 10.0.0.2  
 > - Transmission Control Protocol, Src Port: 41938, Dst Port: 80, Seq: 1, Ack: 28561, Len: 0

```

0000  00 04 00 01 00 06 0a 62 0b b3 40 32 00 00 08 00      .....b ..@2...
0010  45 00 00 3c f9 62 40 00 40 06 2d 55 0a 00 00 03      E..<.b@. @.-U...
0020  0a 00 00 02 a3 d2 00 50 27 cc aa 7e 8d 93 4d 2c      .....P '...M,
0030  a0 10 02 b3 14 33 00 00 01 01 08 0a 00 15 54 5a      ....3.. ....TZ
0040  00 15 54 58 1e 08 20 01 8f a4 c3 a7                ..TX.. ....
    
```

# Limitations

- Except in protocol-specific cases, client doesn't know how much data will be transferred
- Decision threshold is arbitrary

# API Improvement proposal

- MPTCP uses « Add Address » to announce other server interfaces
- Current API doesn't provide access to that info
  - Would be useful to extend the API and provide candidate subflow addresses