

A-PAWS: Alternative Approach for PAWS
draft-nishida-tcpm-apaws-01

Yoshifumi Nishida

Background

- RFC7323 requires putting timestamps in all segments

Once TSopt has been successfully negotiated, TSopt **MUST be sent** in every non-<RST> segment for the duration of the connection

- Time Stamp consumes 10-12 bytes in option space
 - 25-30% option space cannot be used for other purposes

A-PAWS: An Alternative for PAWS

- Don't need to put TS in every segment
 - TS will be used for only RTTM
- Provide the same level of protection as PAWS
 - Fallback to PAWS when there is a risk

A-PAWS logic

- Basic rules

- Sender doesn't put TS for PAWS under certain conditions
 - There is no risk for sequence number wrapping
 - There is no risk for seeing reincarnate segments
- Receiver mustn't drop segments without TS if A-PAWS is activated
- Receiver sends back TS in ACK if it receives TS in Data
- If there is a risk, don't activate A-PAWS or fallback to PAWS

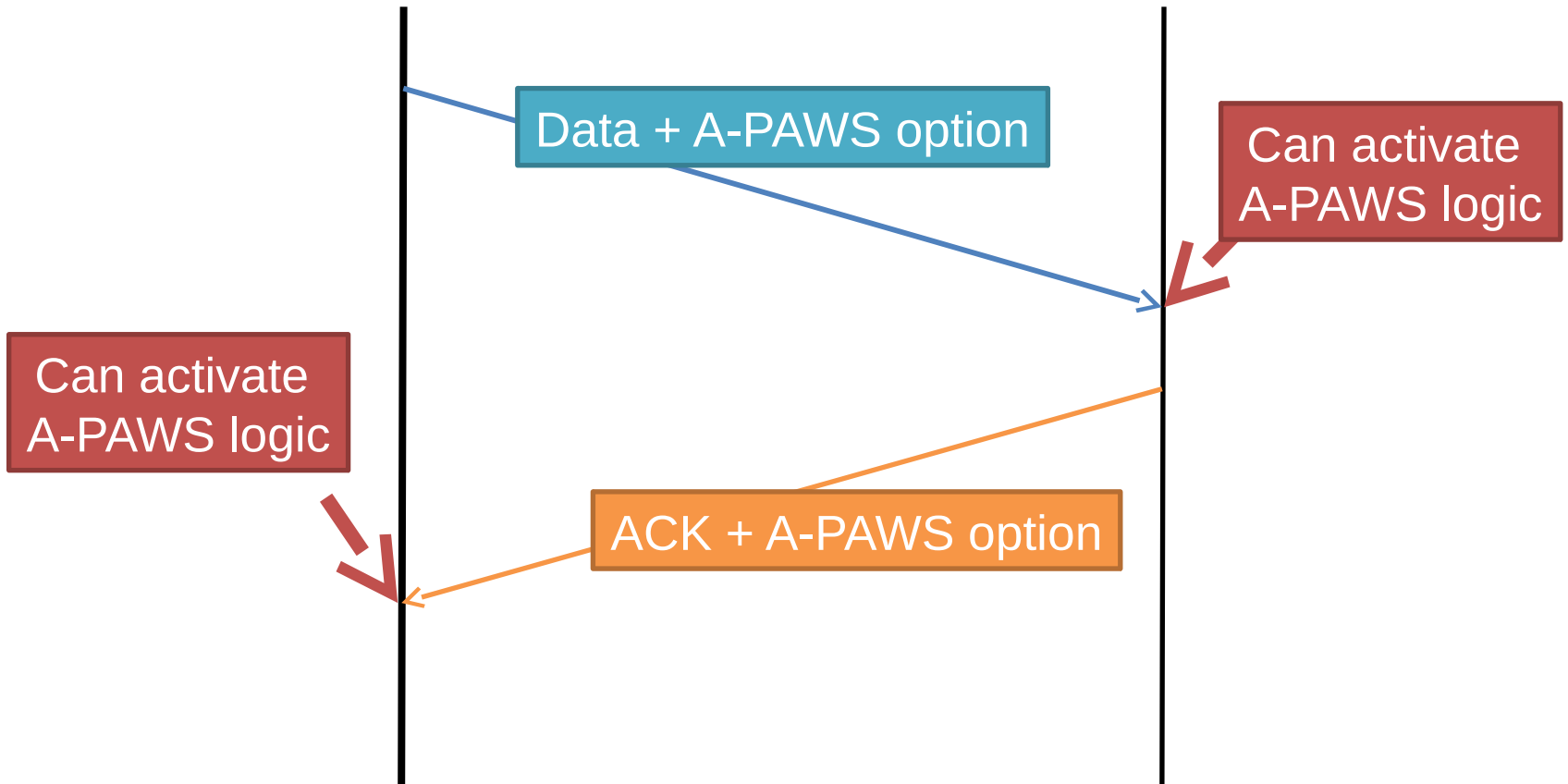
A-PAWS Signaling (1)

- Use new TCP option in non-SYN segments
 - Option space in SYN is very scarce resource
 - Loose synchronization
 - It is basically unidirectional signaling
 - No negotiation, retransmission is necessarily

A-PAWS Signaling (2)

- Set A-PAWS option in ANY data segments
 - Frequency/interval is implementation decision
- When it receives A-PAWS option, it can send a packet with A-PAWS logic (i.e. can omit TS)
- A node MAY set A-PAWS option in ACK when it receives A-PAWS option in data segments

Signaling Example



Sending A-PAWS option means you're OK to omit TS if necessary

Some updates from -00

- Even if TCP sends more than 4GB, we might not need to fallback to normal TCP
 - It will be safe if 2MSL has passed
- If `SO_REUSEADDR` is used, don't send A-PAWS signal, but it can activate A-PAWS if the peer supports it
 - I don't allow you to omit TS, but I will omit TS
 - Useful for client-server style applications

Preliminary Implementation

- Implemented in Linux 3.18.20
 - Implementation specific logic
 - If peer doesn't respond A-PAWS, it will retry after 120 * backoff seconds
 - If 5 times attempts failed, it stop sending A-PAWS option
 - If A-PAWS is activated, it puts TS once in 3 segments when flightsize ≥ 4
 - linux code doesn't drop segments without TS
 - Don't need to update this part

A-PAWS traffic example

```
55782 > B.10000: Flags [S], seq 3972504301, win 29200, options [mss 1460,sackOK,TS val 170830315 ecr 0,nop,wscale 7], length 0
10000 > A.55782: Flags [S.], seq 2832076323, ack 3972504302, win 28960, options [mss 1460,sackOK,TS val 362979174 ecr 170830315,nop,wscale 7], len
55782 > B.10000: Flags [.], ack 1, win 229, options [nop,nop,TS val 170830326 ecr 362979174], length 0
55782 > B.10000: Flags [P.], seq 1:1025, ack 1, win 229, options [nop,nop,TS val 170830326 ecr 362979174,Unknown Option 254ab02], length 1024
55782 > B.10000: Flags [.], seq 1025:2473, ack 1, win 229, options [nop,nop,TS val 170830326 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 2473:3921, ack 1, win 229, options [nop,nop,TS val 170830326 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 3921:5369, ack 1, win 229, options [nop,nop,TS val 170830327 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 5369:6817, ack 1, win 229, options [nop,nop,TS val 170830327 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 6817:8265, ack 1, win 229, options [nop,nop,TS val 170830327 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 8265:9713, ack 1, win 229, options [nop,nop,TS val 170830327 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 9713:11161, ack 1, win 229, options [nop,nop,TS val 170830327 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 11161:12609, ack 1, win 229, options [nop,nop,TS val 170830327 ecr 362979174], length 1448
55782 > B.10000: Flags [.], seq 12609:14057, ack 1, win 229, options [nop,nop,TS val 170830327 ecr 362979174], length 1448
10000 > A.55782: Flags [.], ack 1025, win 243, options [nop,nop,TS val 362979186 ecr 170830326,Unknown Option 254ab02], length 0
10000 > A.55782: Flags [.], ack 2473, win 265, options [nop,nop,TS val 362979186 ecr 170830326], length 0
10000 > A.55782: Flags [.], ack 3921, win 288, options [nop,nop,TS val 362979186 ecr 170830326], length 0
10000 > A.55782: Flags [.], ack 5369, win 311, options [nop,nop,TS val 362979186 ecr 170830327], length 0
10000 > A.55782: Flags [.], ack 6817, win 333, options [nop,nop,TS val 362979186 ecr 170830327], length 0
10000 > A.55782: Flags [.], ack 8265, win 356, options [nop,nop,TS val 362979186 ecr 170830327], length 0
10000 > A.55782: Flags [.], ack 9713, win 378, options [nop,nop,TS val 362979186 ecr 170830327], length 0
10000 > A.55782: Flags [.], ack 11161, win 401, options [nop,nop,TS val 362979186 ecr 170830327], length 0
10000 > A.55782: Flags [.], ack 12609, win 424, options [nop,nop,TS val 362979186 ecr 170830327], length 0
10000 > A.55782: Flags [.], ack 14057, win 446, options [nop,nop,TS val 362979186 ecr 170830327], length 0
55782 > B.10000: Flags [.], seq 14057:15501, ack 1, win 229, options [nop,nop,TS val 170830338 ecr 362979186], length 1444
55782 > B.10000: Flags [.], seq 15501:15505, ack 1, win 229, length 4
55782 > B.10000: Flags [.], seq 15505:16953, ack 1, win 229, length 1448
55782 > B.10000: Flags [.], seq 16953:18401, ack 1, win 229, options [nop,nop,TS val 170830338 ecr 362979186], length 1448
55782 > B.10000: Flags [.], seq 18401:19849, ack 1, win 229, length 1448
55782 > B.10000: Flags [.], seq 19849:21297, ack 1, win 229, length 1448
55782 > B.10000: Flags [.], seq 21297:22745, ack 1, win 229, options [nop,nop,TS val 170830338 ecr 362979186], length 1448
55782 > B.10000: Flags [.], seq 22745:24193, ack 1, win 229, length 1448
55782 > B.10000: Flags [.], seq 24193:25641, ack 1, win 229, length 1448
55782 > B.10000: Flags [.], seq 25641:27089, ack 1, win 229, options [nop,nop,TS val 170830338 ecr 362979186], length 1448
55782 > B.10000: Flags [P.], seq 27089:28537, ack 1, win 229, length 1448
```

Conclusion

- A-PAWS
 - Simple logic to allow TCP not to put TS options in every segments
 - Unidirectional signaling with non-SYN segment
- Please read the draft for more info
 - Feedbacks are welcome!