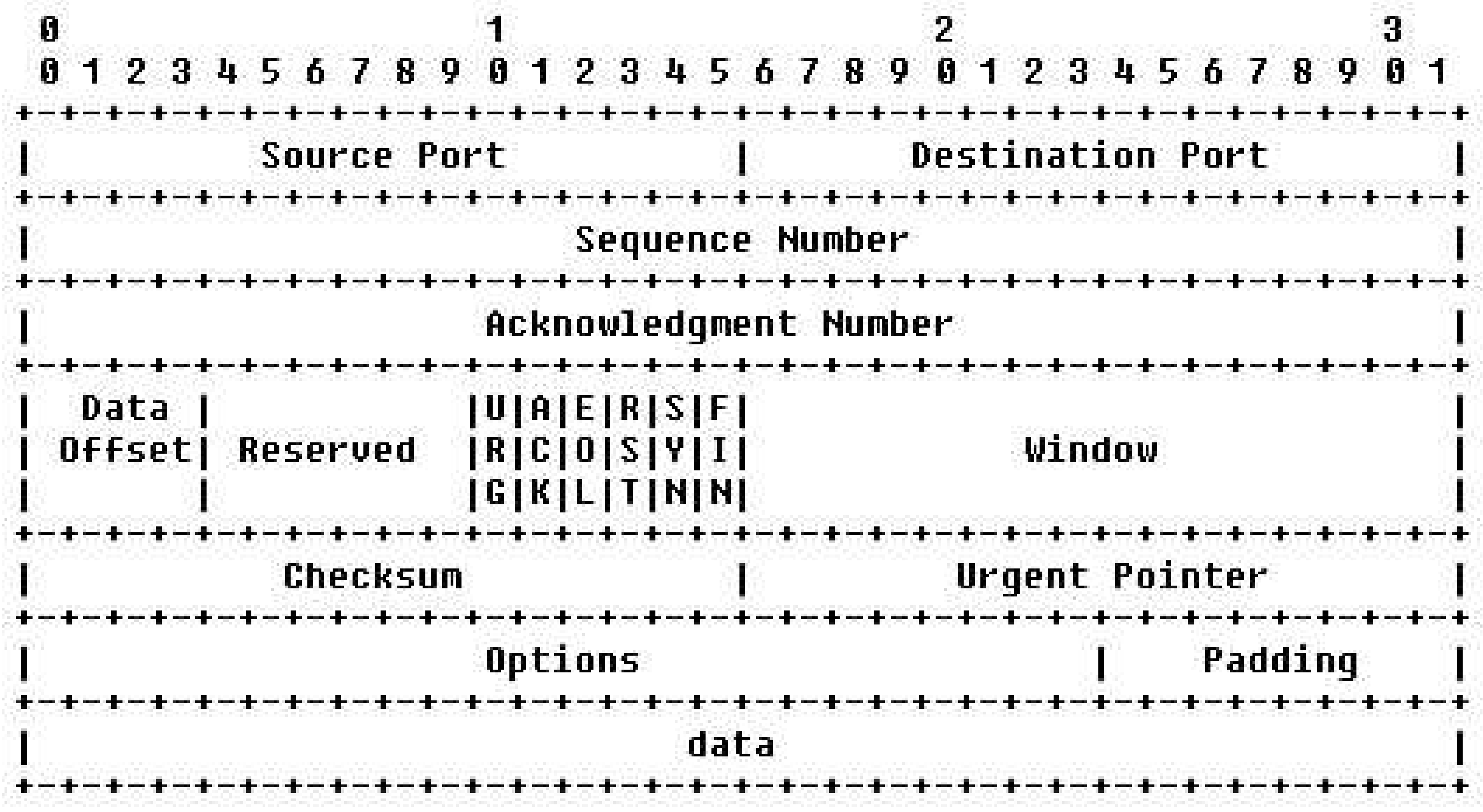
SHARKFEST 2015

WIRESHARK DEVELOPER AND USER CONFERENCE

Advanced TCP stuff – we're not in RFC793 anymore

Jasper Bongertz
Airbus Defence and Space CyberSecurity

Blast from the Past – RFC 761

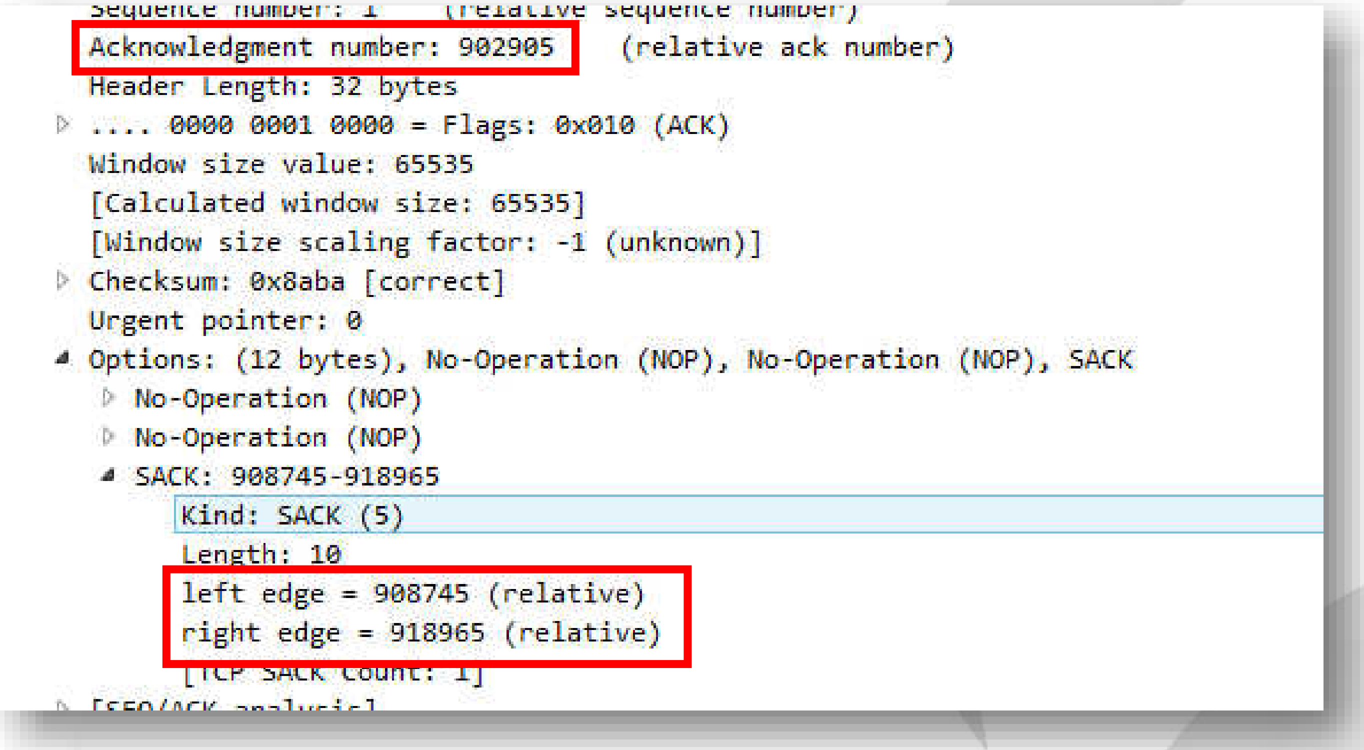


```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Source Port          |       Destination Port        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Acknowledgment Number                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Data |           |U|A|E|R|S|F|                               |
| Offset| Reserved  |R|C|O|S|Y|I|            Window             |
|       |           |G|K|L|T|N|N|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Checksum            |         Urgent Pointer        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Options                    |    Padding    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             data                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

TCP Header Format

Selective Acknowledgements

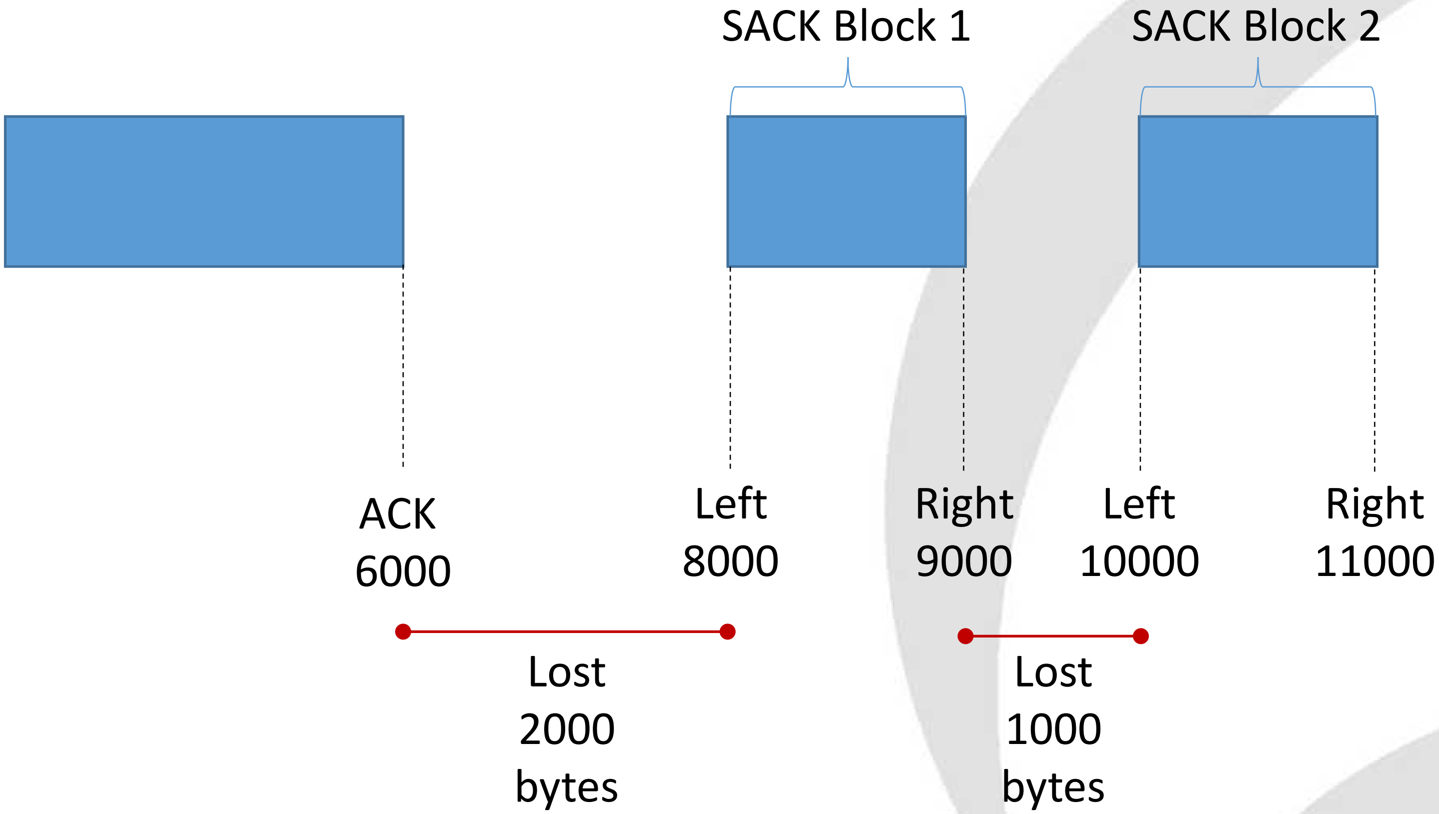
- SACK is used to signal packet loss more precisely
  - SACK edges indicate what was received after the missing segment



```
  Sequence number: 1    (relative sequence number)
  Acknowledgment number: 902905    (relative ack number)
  Header Length: 32 bytes
▷ .... 0000 0001 0000 = Flags: 0x010 (ACK)
  Window size value: 65535
  [Calculated window size: 65535]
  [Window size scaling factor: -1 (unknown)]
▷ Checksum: 0x8aba [correct]
  Urgent pointer: 0
◢ Options: (12 bytes), No-Operation (NOP), No-Operation (NOP), SACK
  ▷ No-Operation (NOP)
  ▷ No-Operation (NOP)
  ◢ SACK: 908745-918965
      Kind: SACK (5)
      Length: 10
      left edge = 908745 (relative)
      right edge = 918965 (relative)
      [TCP SACK Count: 1]
▷ [SEQ/ACK analysis]
```

Selective Acknowledgements

- The ACK number is lower than the left edge values



SACK Block 1

SACK Block 2

ACK
6000

Left
8000

Right
9000

Left
10000

Right
11000

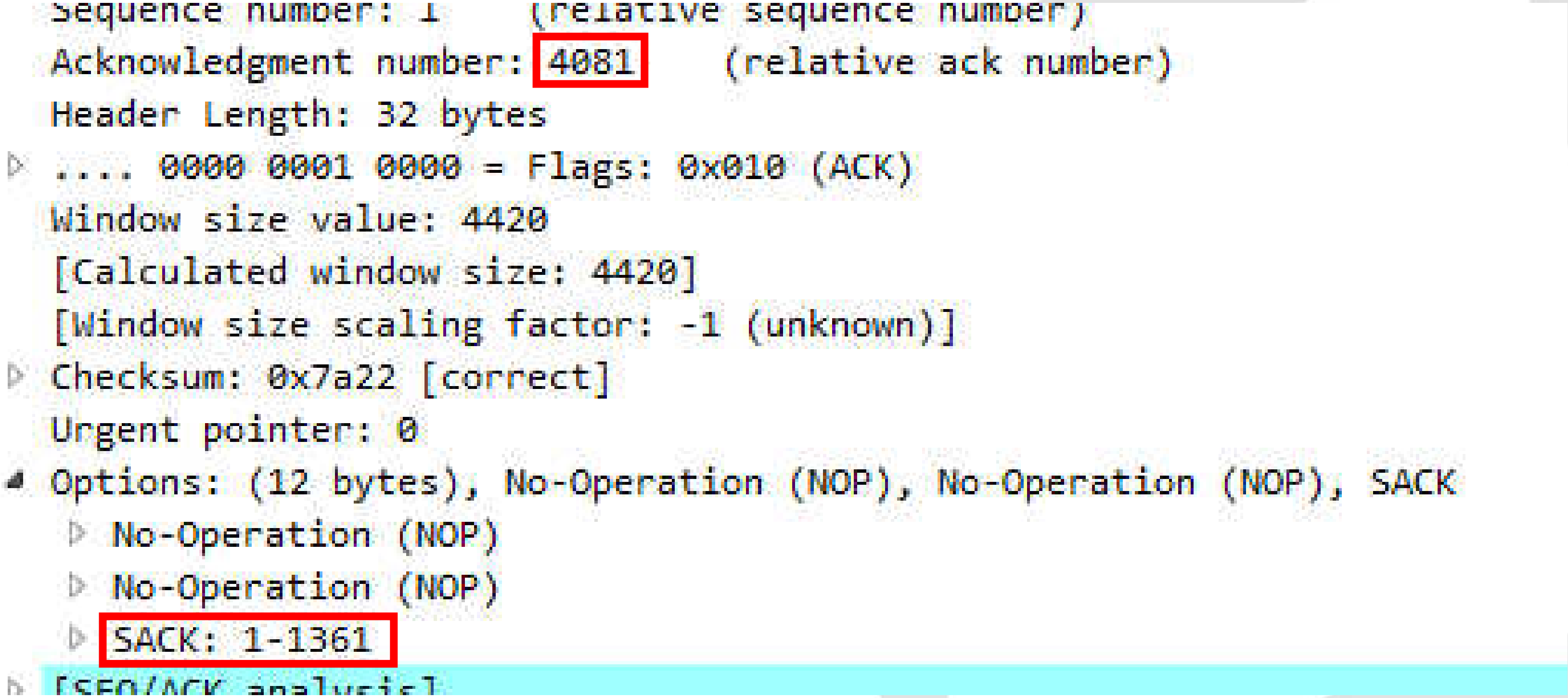
Lost
2000
bytes

Lost
1000
bytes

Demo

D-SACK

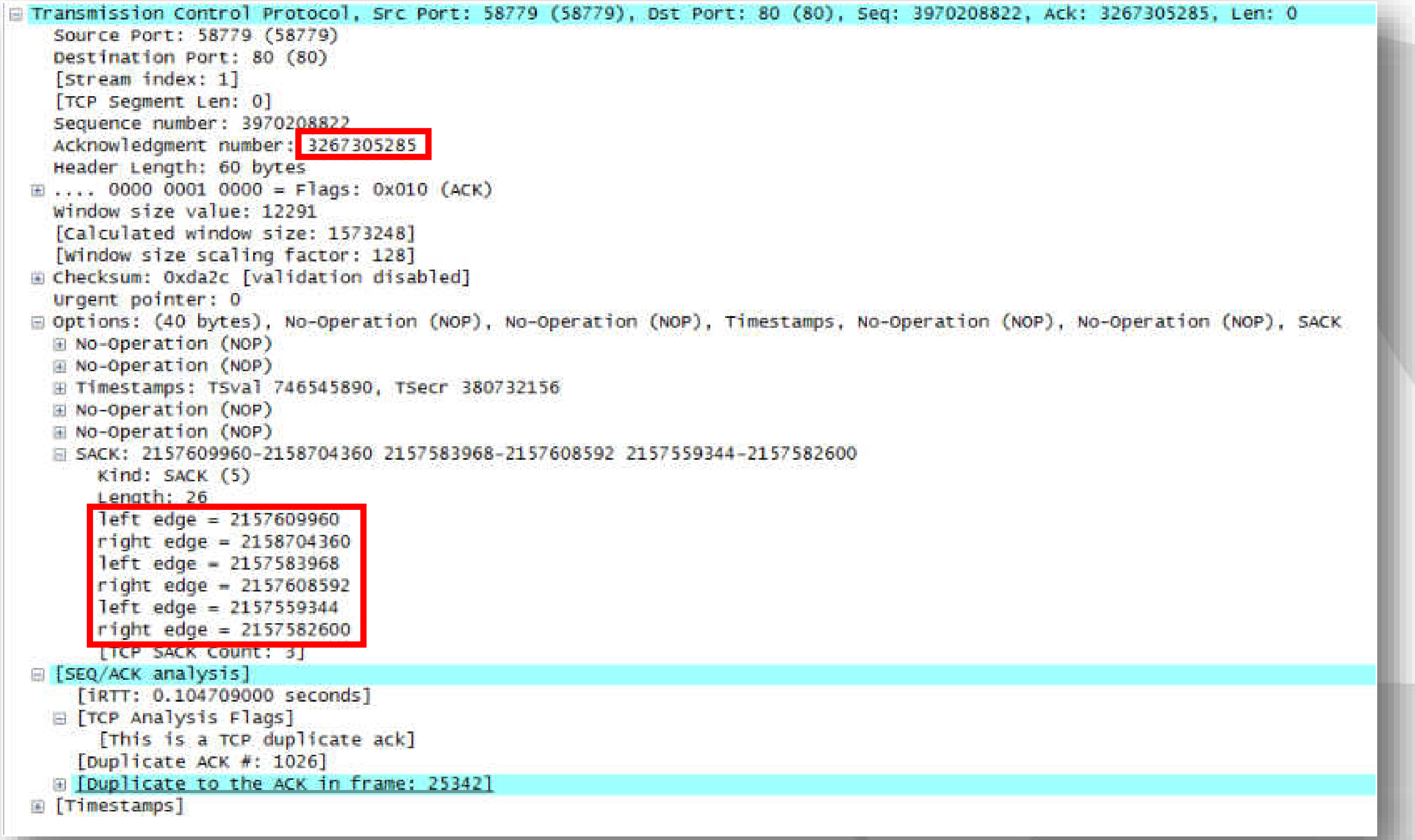
- Special SACK blocks:



```
  Sequence number: 1    (relative sequence number)
  Acknowledgment number: 4081    (relative ack number)
  Header Length: 32 bytes
▷ .... 0000 0001 0000 = Flags: 0x010 (ACK)
  Window size value: 4420
  [Calculated window size: 4420]
  [Window size scaling factor: -1 (unknown)]
▷ Checksum: 0x7a22 [correct]
  Urgent pointer: 0
◢ Options: (12 bytes), No-Operation (NOP), No-Operation (NOP), SACK
   ▷ No-Operation (NOP)
   ▷ No-Operation (NOP)
   ▷ SACK: 1-1361
▷ [SEQ/ACK analysis]
```

Demo

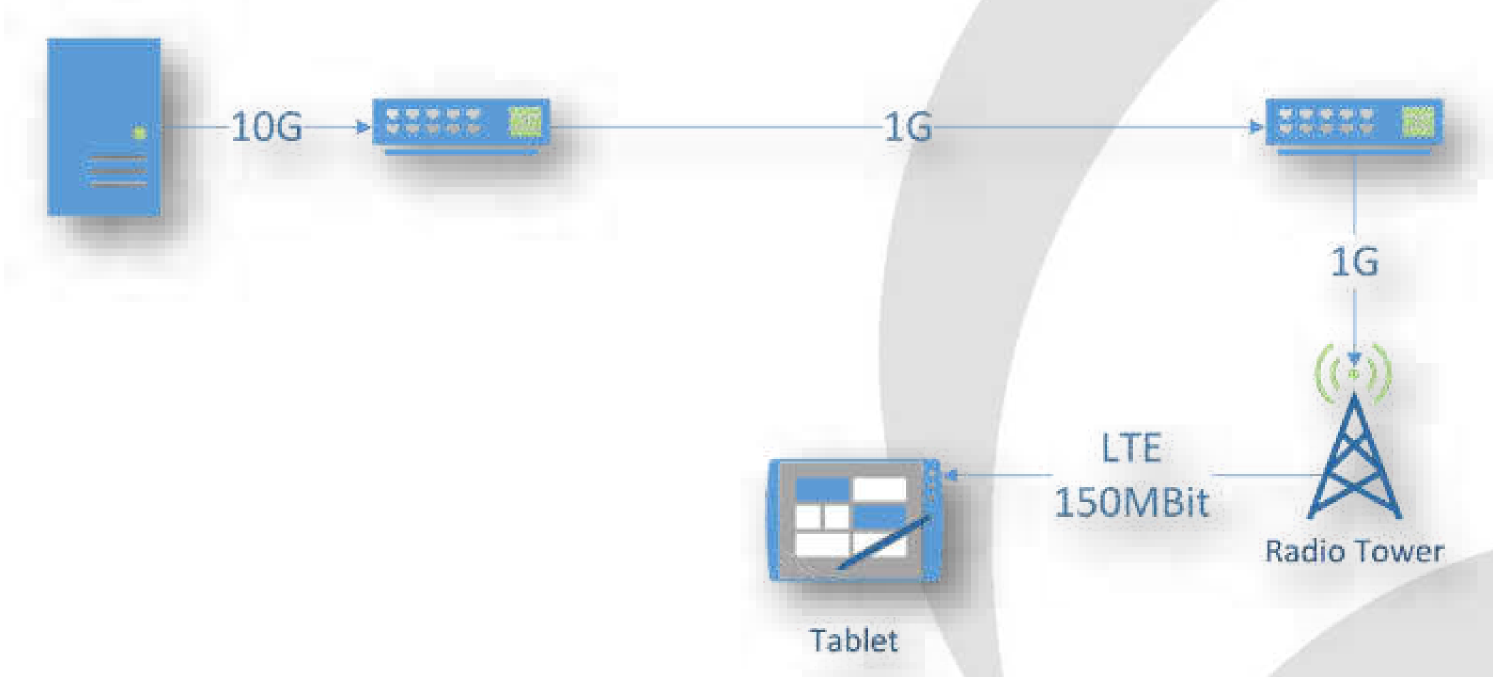
D-SACK or no D-SACK?

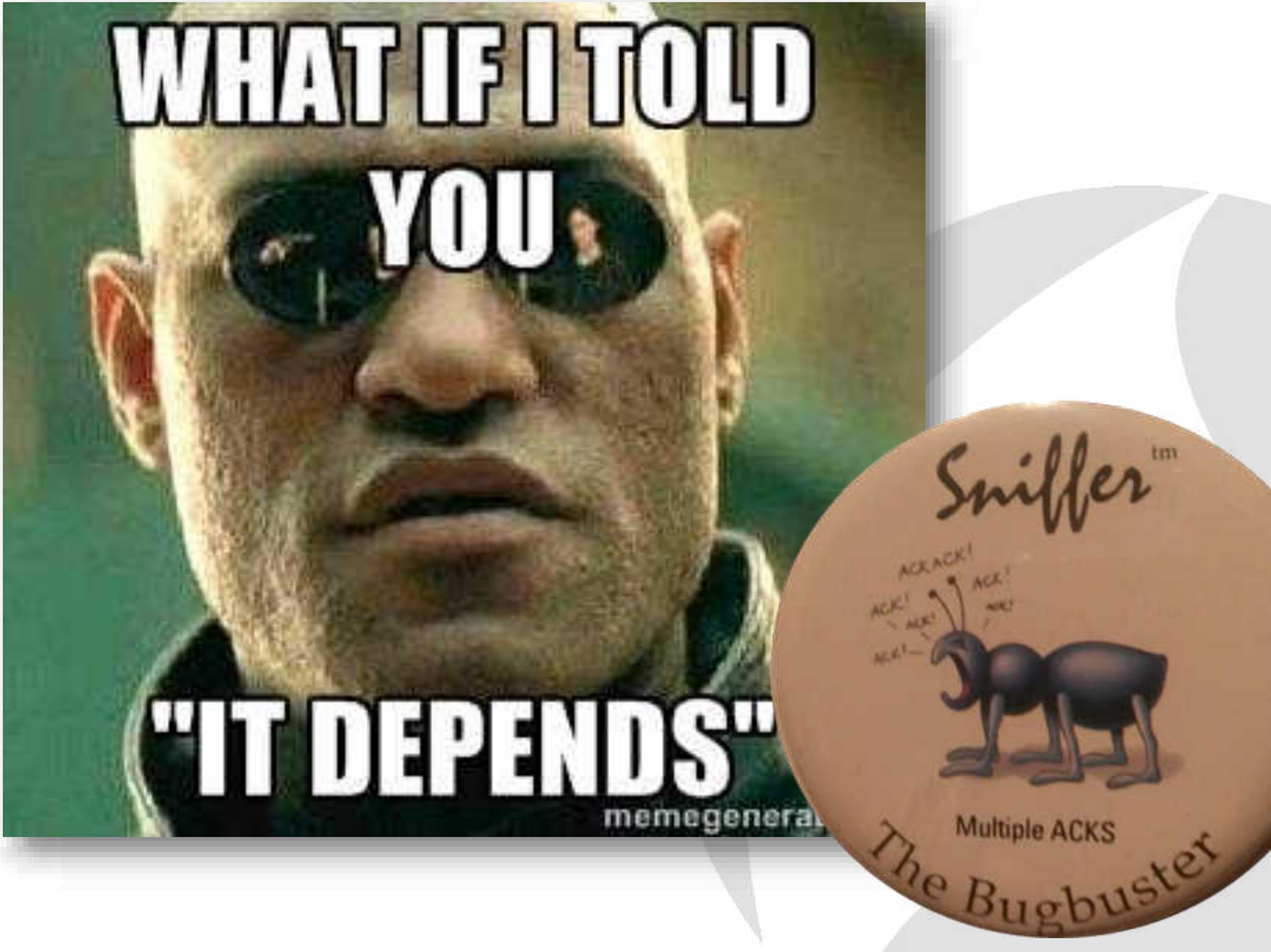


```
Transmission Control Protocol, Src Port: 58779 (58779), Dst Port: 80 (80), Seq: 3970208822, Ack: 3267305285, Len: 0
  Source Port: 58779 (58779)
  Destination Port: 80 (80)
  [Stream index: 1]
  [TCP Segment Len: 0]
  Sequence number: 3970208822
  Acknowledgment number: 3267305285
  Header Length: 60 bytes
  .... 0000 0001 0000 = Flags: 0x010 (ACK)
  Window size value: 12291
  [Calculated window size: 1573248]
  [Window size scaling factor: 128]
  Checksum: 0xda2c [validation disabled]
  Urgent pointer: 0
  Options: (40 bytes), No-Operation (NOP), No-Operation (NOP), Timestamps, No-Operation (NOP), No-Operation (NOP), SACK
    No-Operation (NOP)
    No-Operation (NOP)
    Timestamps: TSval 746545890, TSecr 380732156
    No-Operation (NOP)
    No-Operation (NOP)
    SACK: 2157609960-2158704360 2157583968-2157608592 2157559344-2157582600
      Kind: SACK (5)
      Length: 26
      left edge = 2157609960
      right edge = 2158704360
      left edge = 2157583968
      right edge = 2157608592
      left edge = 2157559344
      right edge = 2157582600
      [TCP SACK Count: 3]
  [SEQ/ACK analysis]
    [iRTT: 0.104709000 seconds]
    [TCP Analysis Flags]
      [This is a TCP duplicate ack]
    [Duplicate ACK #: 1026]
    [Duplicate to the ACK in frame: 25342]
  [Timestamps]
```

Duplicate ACKs and Elephants

- LFN = Long Fat Network (="Elephan")
- Assume you have a network setup like this, what maximum throughput can you achieve?




WHAT IF I TOLD YOU
"IT DEPENDS"
memegenera
Sniffer tm
ACK ACK!
ACK!
ACK!
ACK!
ACK!
ACK!
Multiple ACKS
The Bugbuster

Demo

TCP Fast Open

- Idea: request data already in the SYN packet
  - saves one full round trip time
- Problem:
  - connection isn't established yet
  - this could lead to very effective SYN flooding attacks
- Solution:
  - using "Fast Open Cookies"

Demo

MultiPath TCP

- Idea: open multiple TCP sessions to transport data between two nodes
    - connections use different IPs
    - allows roaming without connection loss
    - data segments have additional sequence numbers
- Challenge: in the future, analyzing isolated TCP connections is not good enough
    - you need to look at all TCP session that are part of the conversation

Thanks! Questions?

eMail: jasper@packet-foo.com
blog: https://blog.packet-foo.com
Twitter: @packetjay