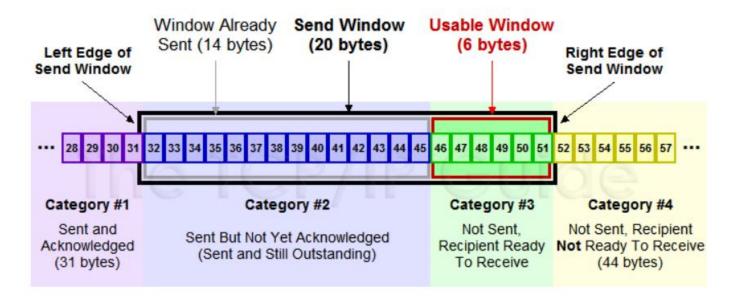
# CS 305 Lab Tutorial Lab 8 TCP Sliding Window

Dept. Computer Science and Engineering Southern University of Science and Technology



### Part A.1 Sliding window system(1)

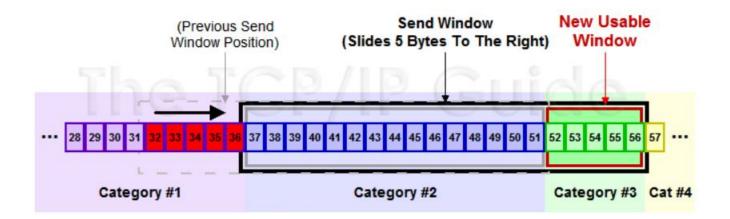


The *send window* is the key to the entire TCP sliding window system: it represents the maximum number of unacknowledged bytes a device is allowed to have outstanding at once.

The **usable window** is the number of bytes that the sender is still allowed to send at any point in time; it is equal to the size of the send window less the number of unacknowledged bytes already transmitted.



## Sliding window system(2)



When the sending device **receives new acknowledgment**, it will be able to transfer some of the bytes from Category #2 to Category #1, since they have now been acknowledged. When it does so, something interesting will happen. Since five bytes have been acknowledged, and the window size didn't change, the sender is allowed to send five more bytes. In effect, the window shifts, or **slides**, over to the right in the timeline.

At the same time five bytes move from Category #2 to Category #1, five bytes move from Category #4 to Category #3, **creating a new usable window for subsequent transmission**.



#### ACK number, Sequence number, len

```
Time
                                 Destination
No.
                 Source
                                                 Protocol Info
                                                        http(80) → 54861 [ACK] Seq=65701 Ack=333 Win=30336 Len=1460
    94 8,574280
                 gaia.cs.umass... 192.168.88.149
                                                 TCP
    95 8.576343
                 gaia.cs.umass... 192.168.88.149
                                                TCP
                                                        http(80) → 54861 [ACK] Seq=67161 Ack=333 Win=30336 Len=1460
                                                        http(80) → 54861 [ACK] <u>Seq=68621</u> Ack=333 Win=30336 Len=1460
    96 8.576345
                 gaia.cs.umass... 192.168.88.149
                                                TCP
                                                        http(80) → 54861 [ACK] Seq=70081 Ack=333 Win=30336 Len=1460
    97 8.576345
                 gaia.cs.umass... 192.168.88.149
                                                TCP
                                                        54861 → http(80) [ACK] Seq=333 Ack=71541 Win=65536 Len=0
    98 8.576516
                192.168.88.149 gaia.cs.umass.... TCP
    Source Port: 54861 (54861)
    Destination Port: http (80)
    [Stream index: 0]
    [TCP Segment Len: 0]
    Sequence number: 333
                            (relative sequence number)
                                  (relative sequence number)]
     [Next sequence number: 333
    Acknowledgment number: 71541
                                    (relative ack number)
                                                                   ack num (71541) =
    0101 .... = Header Length: 20 bytes (5)
                                                                   seg (70081) +len (1460)

√ Flags: 0x010 (ACK)

       000. .... = Reserved: Not set
       ...0 .... = Nonce: Not set
       .... 0... = Congestion Window Reduced (CWR): Not set
       .... .0.. .... = ECN-Echo: Not set
       .... ..0. .... = Urgent: Not set
       .... 1 .... = Acknowledgment: Set
       .... .... 0... = Push: Not set
       .... .... .0.. = Reset: Not set
       .... .... ..0. = Syn: Not set
       .... .... ...0 = Fin: Not set
       [TCP Flags: ······A····]
```



### Changes of window

While the size of usable window turn to be 0, it means the sender will not send any segment at this moment. Wireshark mark the segment with "[Tcp Window Full]"

$$Seq (135781) + len (1280) - ack (135781) = win (1280)$$

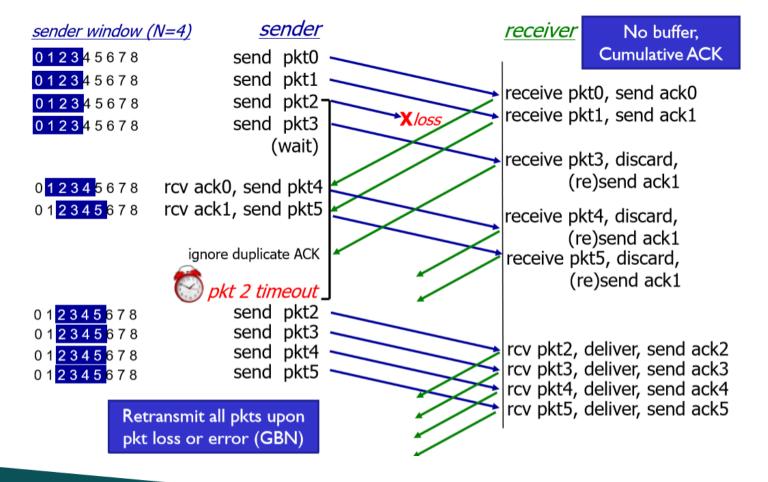
Nο.		Time	Source	Destination	Protocol	Info
	307	19.117577	LAPTOP-RITC8	gaia.cs.umass	TCP	54861 → http(80) [ACK] Seq=333 Ack=135781 Win=1280 Len=0
	309	20.576025	gaia.cs.umas…	LAPTOP-RITC8FU	тср →	• [TCP Window Full] http(80) → 54861 [PSH, ACK] Seq=135781 Ack=333 Win=30336 Len=1280 [TCP

While the recv window turn to be zero, sender will stop to send packet, it will send "[TCP Keep-Alive]" to keep the TCP connection, waiting for the changing of recv window.

```
gaia.cs.umass.edu TCP → [TCP ZeroWindow] 54861 → http(80) [ACK] Seq=333 Ack=137061 Win=0 Len=0
175 19.366403
               192.168.88.149
176 20.862900
                                                     TCP → [TCP Keep-Alive] http(80) → 54861 [ACK] Seq=137060 Ack=333 Win=30336 Len=0
                                    192.168.88.149
177 20.862992
               192.168.88.149
                                    gaia.cs.umass.edu TCP
                                                              [TCP ZeroWindow] 54861 → http(80) [ACK] Seq=333 Ack=137061 Win=0 Len=0
                                    192.168.88.149
                                                              [TCP Keep-Alive] http(80) → 54861 [ACK] Seq=137060 Ack=333 Win=30336 Len=0
178 26.701220
               gaia.cs.umass.edu
                                                              [TCP ZeroWindow] 54861 → http(80) [ACK] Seq=333 Ack=137061 Win=0 Len=0
179 26.701357
               192.168.88.149
                                    gaia.cs.umass.edu TCP
               192.168.88.149
                                    gaia.cs.umass.edu TCP → [TCP Window Update] 54861 → http(80) [ACK] Seq=333 Ack=137061 Win=65536 Len=0
180 31.323807
```

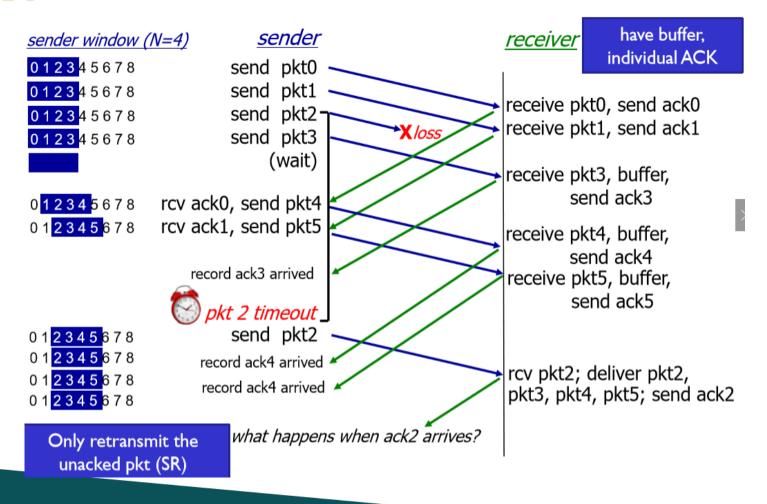


#### Part A.2 Retransmission: GBN





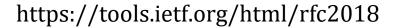
#### SR





#### TCP SACK

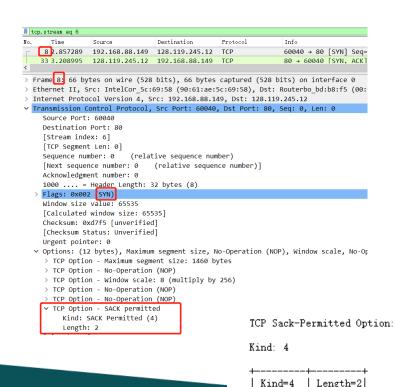
- A Selective Acknowledgment (SACK) mechanism, combined with a selective repeat retransmission policy, can help the sender retransmit only the missing data segments.
  - The receiving TCP sends back SACK packets to the sender informing the sender of data that has been received.
  - With selective acknowledgments, the data receiver can inform the sender about all segments that have arrived successfully, so the sender need retransmit only the segments that have actually been lost.
- The selective acknowledgment extension uses two TCP options:
  - SACK-permitted option
  - SACK option





#### **SACK-permitted option**

• "SACK-permitted", which may be sent in a SYN segment to indicate that the SACK option can be used once the connection is established.



```
8 2.857289 192.168.88.149 128.119.245.12 TCP
                                                                60040 → 80 [SYN] Seq=0 Win=65535 Len=6
  333.208995 128.119.245.12 192.168.88.149 TCP
                                                                80 → 60040 [SYN, ACK] Seq=0 Ack=1 Win:
> Frame 33 66 bytes on wire (528 bits), 66 bytes captured (528 bits) on interface 0
> Ethernet II, Src: Routerbo bd:b8:f5 (00:0c:42:bd:b8:f5), Dst: IntelCor 5c:69:58 (90:61:ae:5c:69:58)
> Internet Protocol Version 4, Src: 128.119.245.12, Dst: 192.168.88.149
Transmission Control Protocol, Src Port: 80, Dst Port: 60040, Seq: 0, Ack: 1, Len: 0
    Source Port: 80
    Destination Port: 60040
    [Stream index: 6]
    [TCP Segment Len: 0]
    Sequence number: 0 (relative sequence number)
                               (relative sequence number)]
    [Next sequence number: 0
    Acknowledgment number: 1
                               (relative ack number)
    1000 .... = Header Length: 32 bytes (8)
    Flags: 0x012 SYN, ACK)
    Window size value: 29200
    [Calculated window size: 29200]
    Checksum: 0x1a9f [unverified]
    [Checksum Status: Unverified]
    Urgent pointer: 0
  Options: (12 bytes), Maximum segment size, No-Operation (NOP), No-Operation (NOP), SACK permitted,
    > TCP Option - Maximum segment size: 1460 bytes
    > TCP Option - No-Operation (NOP)
    > TCP Option - No-Operation (NOP)

→ TCP Option - SACK permitted

         Kind: SACK Permitted (4)
```



Wireshark tips: TCP.option\_kind==4

#### SACK option(1)

The SACK option is to be sent by a data receiver to inform the data sender of non-contiguous blocks of data that have been received and queued.

The data receiver awaits the receipt of data (perhaps by means of retransmissions) to fill the gaps in sequence space between received blocks.

When missing segments are received, **the data receiver acknowledges the data normally** by advancing the left window edge in the Acknowledgement Number Field of the TCP header.

The SACK option does not change the meaning of the Acknowledgement Number field.

This option contains a list of **some of the blocks of contiguous sequence space occupied by data that has been received and queued within the window.** 

Each contiguous block of data queued at the data receiver is defined in the SACK option by **two 32-bit unsigned integers** in network byte order:

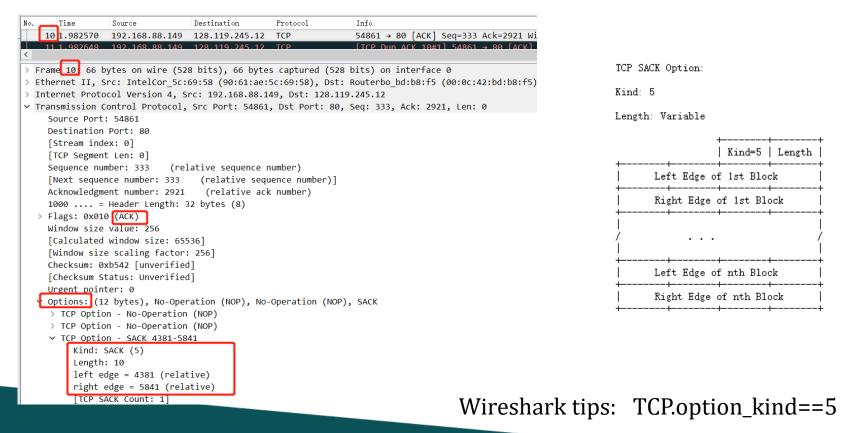
- **Left Edge** of Block This is the first sequence number of this block.
- Right Edge of Block This is the sequence number immediately following the last sequence number of this block.

Each block represents received bytes of data that are **contiguous and isolated**; that is, **the bytes just below the block**, (**Left Edge of Block - 1**), and just above the block, (**Right Edge of Block**), have NOT been received.



### SACK option(2)

• SACK option, which may be sent over an established connection once permission has been given by SACK-permitted.





#### SACK option(3)

```
Destination
                                                              http(80) → 54861 [ACK] Seq=151841 Ack=333 Win=30336 Len=1460 TCP segment of a reassembled PDU]
199 46.985513 gaia.cs.umass... 192.168.88.149 TCP
                                                              54861 → http(80) [ACK] Seq=333 Ack=153301 Win=65536 Len=0
200 46.985600 192.168.88.149 gaia.cs.umass.... TCP
                                                               [TCP Previous segment not captured] http(80) → 54861 [ACK] Seq=156221 Ack=333 Win=30336 Len=1460 [TCP
201 47.595142 gaia.cs.umass.... 192.168.88.149 TCP
202 47.595144 gaia.cs.umass... 192.168.88.149 TCP
                                                              http(80) → 54861 [ACK] Seq=157681 Ack=333 Win=30336 Len=1460 [TCP segment of a reassembled PDU]
                                                              [TCP Dup ACK 200#1] 54861 → http(80) [ACK] Seq=333 Ack=153301 Win=65536 Len=0 SLE=156221 SRE=157681
203 47.595274 192.168.88.149 gaia.cs.umass.... TCP
                                                              [TCP Dup ACK 200#2] 54861 → http(80) [ACK] Seq=333 Ack=153301 Win=65536 Len=0 SLE=156221 SRE=159141
204 47.595443 192.168.88.149 gaia.cs.umass... TCP
                                                               [TCP Retransmission] http(80) → 54861 [ACK] Seq=153301 Ack=333 Win=30336 Len=1460
205 48.207253 gaia.cs.umass... 192.168.88.149 TCP
                                                              54861 → http(80) [ACK] Seq=333 Ack=154761 Win=65536 Len=0 SLE=156221 SRE=159141
206 48.207367 192.168.88.149 gaia.cs.umass.... TCP
207 49.742628 gaia.cs.umass.... 192.168.88.149 TCP
                                                               [TCP Retransmission] http(80) → 54861 [ACK] Seq=154761 Ack=333 Win=30336 Len=1460
208 49.742765 192.168.88.149 gaia.cs.umass... TCP
                                                              54861 → http(80) [ACK] Seq=333 Ack=159141 Win=65536 Len=0
                                                              http(80) → 54861 [ACK] Seq=159141 Ack=333 Win=30336 Len=1460 [TCP segment of a reassembled PDU]
209 50.363845 gaia.cs.umass.... 192.168.88.149 TCP
```

```
#203 and #204 are SACK
#203 tells that 156221~157681 are contiguous and isolated
#204 tells that 156221~159141 are contiguous and isolated
#200 tells that the block before 153301 are acked
So
#205 retransmit 153301 ~ 153301+ 1460 -1 ,#206 ack it with 154761
#207 retransmit 154761 ~ 154761+1460 -1
#208 ack it with 159141 (for 157681~159140 are contiguous but NOT isolated)
```



## SACK option(4)

No.	Time	Source	Destination	Protocol	Info
	199 46.985513	gaia.cs.umass	192.168.88.149		http(80) → 54861 [ACK] Seq=151841 Ack=333 Win=30336 Len=1460 [TCP segment of a reassembled PDU]
	200 46.985600	192.168.88.149	gaia.cs.umass	TCP	54861 → http(80) [ACK] Seq=333 Ack=153301 Win=65536 Len=0
	201 47.595142	gaia.cs.umass	192.168.88.149	TCP	[TCP Previous segment not captured] http(80) → 54861 [ACK] Seq=156221 Ack=333 Win=30336 Len=1460 [TCP
	202 47.595144	gaia.cs.umass	192.168.88.149	TCP	http(80) → 54861 [ACK] Seq=157681 Ack=333 Win=30336 Len=1460 [TCP segment of a reassembled PDU]
	203 47.595274	192.168.88.149	gaia.cs.umass	TCP	[TCP Dup ACK 200#1] 54861 → http(80) [ACK] Seq=333 Ack=153301 Win=65536 Len=0 SLE=156221 SRE=157681
	204 47.595443	192.168.88.149	gaia.cs.umass	TCP	[TCP Dup ACK 200#2] 54861 → http(80) [ACK] Seq=333 Ack=153301 Win=65536 Len=0 SLE=156221 SRE=159141
	205 48.207253	gaia.cs.umass	192.168.88.149	TCP	[TCP Retransmission] http(80) → 54861 [ACK] Seq=153301 Ack=333 Win=30336 Len=1460
	206 48.207367	192.168.88.149	gaia.cs.umass	TCP	54861 → http(80) [ACK] Seq=333 Ack=154761 Win=65536 Len=0 SLE=156221 SRE=159141
	207 49.742628	gaia.cs.umass	192.168.88.149	TCP -	[TCP Retransmission] http(80) → 54861 [ACK] Seq=154761 Ack=333 Win=30336 Len=1460
	208 49.742765	192.168.88.149	gaia.cs.umass	TCP	54861 → http(80) [ACK] Seq=333 Ack=159141 Win=65536 Len=0
	209 50.363845	gaia.cs.umass	192.168.88.149	TCP	http(80) → 54861 [ACK] Seq=159141 Ack=333 Win=30336 Len=1460 [TCP segment of a reassembled PDU]

153300	153301	•••	154761		156221	•••	157681		159141
#200 ACK	200 ACK					#201 Transmit		#202 Transmit	
					#203 ACK		#204 ACK		
	#205 Retran	smit							
	#206 ACK		#207 Retrar #208 ACK	nsmit					
			00 / 1011	į					11200 T





#### Retransmission(1)

```
Time
               Source
                               Destination
                                                Protocol
                                                                Info
                                                                http(80) → 54861 [ACK] Seq=157681 Ack=333 Win=30336 Len=146
202 47.595144 gaia.cs.umass... 192.168.88.149 TCP
203 47.595274 192.168.88.149 gaia.cs.umass.... TCP
                                                                [TCP Dup ACK 200#1] 54861 → http(80) [ACK] Seq=333 Ack=153
204 47.595443 192.168.88.149 gaia.cs.umass.... TCP
                                                                [TCP Dup ACK 200#2] 54861 → http(80) [ACK] Seq=333 Ack=1533
                                                             TCP Retransmission] http(80) → 54861 [ACK] Seq=153301 Ack=
205 48.207253 gaia.cs.umass... 192.168.88.149 TCP
  Sequence number: 153301
                             (relative sequence number)
                                    (relative sequence number)]
  [Next sequence number: 154761
  Acknowledgment number: 333
                                (relative ack number)
  0101 .... = Header Length: 20 bytes (5)
> Flags: 0x010 (ACK)
  Window size value: 237
  [Calculated window size: 30336]
  [Window size scaling factor: 128]
  Checksum: 0x3487 [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0

▼ [SEQ/ACK analysis]
     [iRTT: 0.450320000 seconds]
     [Bytes in flight: 5840]
     [Bytes sent since last PSH flag: 16060]

「TCP Analysis Flags]
     v [Expert Info (Note/Sequence): This frame is a (suspected) retransmission]
          [This frame is a (suspected) retransmission]
          [Severity level: Note]
          [Group: Sequence]
       [The RTO for this segment was: 0.612109000 seconds]
       [RTO based on delta from frame: 202]
```

While RTO timeout, retransmission is triggered



#### Retransmission(2)

```
Destination
    Time
               Source
                                                Protocol
                                                                Info
202 47.595144 gaia.cs.umass... 192.168.88.149 TCP
                                                                http(80) → 54861 [ACK] Seq=157681 Ack=333 Win=30336 Len=1466
203 47.595274 192.168.88.149 gaia.cs.umass.... TCP
                                                                [TCP Dup ACK 200#1] 54861 → http(80) [ACK] Seq=333 Ack=15330
204 47.595443 192.168.88.149 gaia.cs.umass.... TCP
                                                                [TCP Dup ACK 200#2] 54861 → http(80) [ACK] Seq=333 Ack=15336
                                                                [TCP Retransmission] http(80) → 54861 [ACK] Seq=153301 Ack=3
205 48.207253 gaia.cs.umass... 192.168.88.149 TCP
                                                               54861 → http(80) [ACK] Seq=333 Ack=154761 Win=65536 Len=0 SL
206 48.207367 192.168.88.149 gaia.cs.umass... TCP
207 49.742628 gaia.cs.umass... 192.168.88.149 TCP
                                                             FTCP Retransmission http(80) → 54861 [ACK] Seq=154761 Ack=3
  Sequence number: 154761
                             (relative sequence number)
                                   (relative sequence number)]
  [Next sequence number: 156221
  Acknowledgment number: 333
                                (relative ack number)
  0101 .... = Header Length: 20 bytes (5)
> Flags: 0x010 (ACK)
  Window size value: 237
  [Calculated window size: 30336]
  [Window size scaling factor: 128]
  Checksum: 0x595f [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0

▼ [SEQ/ACK analysis]
     [iRTT: 0.450320000 seconds]
     [Bytes in flight: 4380]
     [Bytes sent since last PSH flag: 17520]

「TCP Analysis Flags]
     v [Expert Info (Note/Sequence): This frame is a (suspected) retransmission]
          [This frame is a (suspected) retransmission]
          [Severity level: Note]
          [Group: Sequence]
       [The RTO for this segment was: 2.147484000 seconds]
       [RTO based on delta from frame: 202]
```



#### Fast retransmission

- TCP may generate an immediate acknowledgment (a duplicate ACK) when an out-of-order segment is received. This duplicate ACK should not be delayed. The purpose of this duplicate ACK is to let the other end know that a segment was received out of order, and to tell it what sequence number is expected.
- Since TCP does not know whether a duplicate ACK is caused by a lost segment or just a reordering of segments, it waits for a small number of duplicate ACKs to be received.
  - It is assumed that if there is just a reordering of the segments, there will be only one or two
    duplicate ACKs before the reordered segment is processed, which will then generate a new ACK.
  - If three or more duplicate ACKs are received in a row, it is a strong indication that a segment has been lost.
- TCP then performs a retransmission of what appears to be the missing segment, without waiting for a retransmission timer to expire.

https://tools.ietf.org/html/rfc2001



#### Practise 8.1

- Using Wireshark to capture and analysis the TCP stream.
- Invoke a HTTP request to get http://gaia.cs.umass.edu/wiresharklabs/alice.txt
- Analysis the TCP stream
  - Any duplicate ack, what's the possible reason?
  - Any TCP segment with sack permit option and sack option
    - select a tcp package which contained a sack option, find the segment ranges which is acked in this sack option
  - Any TCP retransmission? Is it retransmission or fast retransmission?
  - Any window size 0 segment, what does it mean? If the window size is 0, what would happened next on this tcp connection?
  - Any TCP window full segment, what does it mean?
- Tips: you can use some tools to cause your network congestion, such as clumsy-0.2-win64, which has been uploaded in Sakai site.



#### Tips on Wireshark(1)

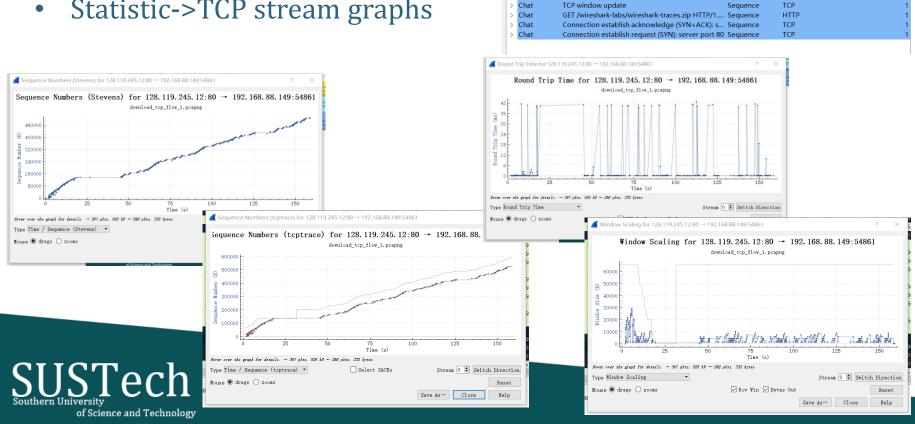
- How to get all the TCP segments of the TCP stream related to a http session in Wireshark:
  - 1<sup>st</sup> step: Find a HTTP packet in the HTTP session
  - 2<sup>nd</sup> step: right click the packet which will invoke a shortcut menu, then choose "follow ->TCP" in the shortcut menu
- How to find a special segment in a TCP stream
  - Using display filter in Wireshark
  - 1<sup>st</sup> step: if a TCP stream is filtered, then a description such as "tcp.stream eq xx" (xx here is the id of this TCP stream) could be found in the display filter
  - 2<sup>nd</sup> step: in the same display filter, make a new filter rule description along with the original one, then press "Enter" key in your keyboard to make the new filter run
    - such as:

tcp.stream eq 1 is the original one, to find a TCP zero window in this TCP stream, tcp.stream eq 1 && tcp.window\_size\_value==0 is the new display filter description



### Tips on Wireshark(2)

- Analysis -> expert info
- Statistic->TCP stream graphs



■ Wireshark · Expert Information · download\_tcp\_flow\_1.pcapng

TCP Zero Window segment

TCP keep-alive segment

Duplicate ACK (#1)

Connection finish (FIN)

TCP window specified by the receiver is now co...

This frame is a (suspected) out-of-order segment

Previous segment(s) not captured (common at c..

This frame is a (suspected) fast retransmission

This frame is a (suspected) retransmission

Summary

Severity

> Warning

Warning

Warning

Warning

Note

Note

Note

Note

Chat

Protocol

TCP

TCP

TCP

TCP

TCP

TCP

TCP

TCP

TCP

Count

Group

Sequence

Sequence

Sequence

Sequence

Sequence

Sequence