

0.1 CVE-2021-24086

According to Microsoft[5] CVE-2021-24086 is a denial of service vulnerability with a CVSS:3.0 score of 7.5 / 6.5, that is a base score metrics of 7.5 and a temporal score metrics of 6.5. The vulnerability affects all supported versions of Windows and Windows Server. According to an accompanied blog post published by Microsoft [7] at the same time as the patch was released, details that the vulnerable component is the Windows TCP/IP implementation, and that the vulnerability revolves around IPv6 fragmentation. The Security Update guide and the blog post also present a workaround that can be used to temporarily mitigate the vulnerability by disabling IPv6 fragmentation.

Figure out if this should be here

0.1.1 Public information

Due to the Microsoft Active Protections Program (MAPP)[6] security software providers are given early access to vulnerability information. This information often include Proof of Concept (PoC)s for vulnerabilities to be patched, in order to aid security software providers to create valid detections for exploitation of soon-to-be patched vulnerabilities. Due to MAPP, some security software providers publish relevant information regarding recently patched vulnerabilities. However, the information is usually very vague in details, and can therefore only aid in the initial exploration of the vulnerability. For CVE-2021-24086, both McAfee[10] and Palo Alto[9] posted public information about CVE-2021-24086. However, both articles contained very limited details, and is therefore far from sufficient to reproduce the vulnerability. Before trying to rediscover the vulnerability, the following information is available:

- The vulnerability lies within the handling of fragmented packets in IPv6
- The relevant code lies within the `tcpip.sys` drivers
- The root cause of the vulnerability is a NULL pointer dereference in `Ipv6ReassembleDatagram` of `tcpip.sys`
- The reassembled packet should contain around 0xFFFF (65535) bytes of extension headers, which is usually not possible

0.1.2 Binary diffing

The usage of binary diffing to gather information about patched vulnerabilities is well described in current research[8][12], and has been made popular and easy to do by tools such as Bindiff[13] and Diaphora[4].

If we look at figure 1 we can compare the function changes of the patched and not-patched `tcpip.sys`. Looking at `tcpip!Ipv6pReassembleDatagram` we can see that the similarity factor is only 0.38 telling us that a significant amount of code has been changed.

write a little about how bindiffing works. Or don't idc.

0.1. CVE-2021-24086

Similarity	Confid	Change	EA Primary	Name Primary	EA Secondary	Name Secondary
0.16	0.27	GI--E--	00000001C018D794	sub_00000001C018D794	00000001C015A1D6	sub_00000001C015A1D6
0.27	0.42	GI--EL-	00000001C01905B5	sub_00000001C01905B5	00000001C01568FC	lppCleanupPathPrimitive
0.31	0.73	GI--E--	00000001C0190F38	Ipv4pReassembleDatagram	00000001C0190F68	Ipv4pReassembleDatagram
0.38	0.98	GI--E--	00000001C0199FAC	Ipv6pReassembleDatagram	00000001C019A0AC	Ipv6pReassembleDatagram
0.42	0.62	-I--E--	00000001C0154959	sub_00000001C0154959	00000001C0001E42	sub_00000001C0001E42
0.54	0.96	GI-----	00000001C019A658	Ipv6pReceiveFragment	00000001C019A7F8	Ipv6pReceiveFragment

Figure 1: Primary matched functions of `tcpip.sys`

Diving into the binary diff of `tcpip!Ipv6pReassembleDatagram` as seen on listing 1, we can clearly see a change. The first many changes from line *5-39* are simply register changes and other insignificant changes due to how the compiler works. However, on line *41-42* a new comparison is made to ensure that the value of the register `edx` is less than `0xFFFF`. This matches the statement given in subsection 0.1.1 (Public information), that the vulnerability is triggered by a packet of around `0xFFFF` bytes.

```
1  --- "a/.\unpatched tcpip.sys"
2  +++ "b/.\patched tcpip.sys"
3  @@ -1,6 +1,4 @@
4  -sub     rsp, 58h          ; Integer Subtraction
5  +sub     rsp, 60h          ; Integer Subtraction
6  movzx   r9d, word ptr [rdx+88h] ; Move with Zero-Extend
7  mov     rdi, rdx
8  mov     edx, [rdx+8Ch]
9  -mov     bl, r8b
10 +mov     r13b, r8b
11 add     edx, r9d          ; Add
12 -mov     byte ptr [rsp+98h+var_70], 0
13 -and     [rsp+98h+var_78], 0 ; Logical AND
14 mov     [rsp+98h+length], edx
15 lea     eax, [rdx+28h]    ; Load Effective Address
16 -mov     rdx, rdi
17 mov     [rsp+98h+var_68], eax
18 lea     eax, [r9+28h]     ; Load Effective Address
19 mov     [rsp+98h+BytesNeeded], eax
20 -xor     r9d, r9d         ; Logical Exclusive OR
21 mov     rax, [rcx+0D0h]
22 -lea     rcx, IppReassemblyNetBufferListsComplete ; Load Effective
    ↪ Address
23 -mov     r13, [rax+8]
24 -mov     rax, [r13+0]
25 +mov     r12, [rax+8]
26 +mov     rax, [r12]
27 mov     r15, [rax+28h]
28 mov     eax, gs:1A4h
29 mov     r8d, eax
30 -mov     rax, [r13+388h]
31 +mov     rax, [r12+388h]
32 lea     rbp, [r8+r8*2]    ; Load Effective Address
33 -mov     r12, [rax+r8*8]
34 -xor     r8d, r8d         ; Logical Exclusive OR
35 +mov     rcx, [rax+r8*8]
36 shl     rbp, 6           ; Shift Logical Left
37 -add     rbp, [r15+4728h] ; Add
38 +add     rbp, [r15+4728h] ; Add
39 +mov     [rsp+98h+var_58], rcx
40 +cmp     edx, 0FFFFh      ; Compare Two Operands
41 +jbe     short loc_1C019A186 ; Jump if Below or Equal (CF=1 | ZF=1)
```

Listing 1: Diff of patched and vulnerable Ipv6pReassembleDatagram

Looking at the raw assembly without any knowledge of what the registers contain or what parameters are passed to the function can be very confusing. To make it easier for the reader to follow, listing 2 contains the annotated

decompiled code of the vulnerable and patched `tcpip!Ipv6pReassembleDatagram` function. Here the patch is easy to spot, as the call to `tcpip!NetioAllocateAndReferenceNetBufferAndNetBufferList` is replaced with the check that we also observed in listing 1. The check is there to ensure that the total packet size is less than `0xFFFF`, which is the largest 16 bit value. The packet size is calculated on line 4-6 using the fragmentable and unfragmentable parts of the reassembled packet.

```
1  --- "a/.\\unpatched tcpip.sys"
2  +++ "b/.\\patched tcpip.sys"
3  void __fastcall Ipv6pReassembleDatagram(__int64 a1, struct_datagram
   ↳ *datagram, char a3) {
4  unfragmentableHeaderLength = datagram->unfragmentableHeaderLength;
5  packetSize = unfragmentableHeaderLength + datagram->fragmentableLength;
6  BytesNeeded = unfragmentableHeaderLength + 40;
7  v6 = *(_QWORD *)((_QWORD *) (a1 + 208) + 8i64);
8  v7 = *(_QWORD *)((_QWORD *) v6 + 40i64);
9  LockArray_high = HIDWORD(KeGetPcr()[1].LockArray);
10 -v11 = NetioAllocateAndReferenceNetBufferAndNetBufferList(IppReassembly_
   ↳ NetBufferListsComplete, datagram, 0i64, 0i64, 0,
   ↳ 0);
11 +if ( packetSize > 0xFFFF )
```

Listing 2: Diff of patched and vulnerable `Ipv6pReassembleDatagram`

At this stage of the vulnerability rediscovery process, the following requirements are now available:

- We have to abuse IPv6 fragmentation in `tcpip!Ipv6pReassembleDatagram`
- We have to construct a single packet with around `0xFFFF` bytes of extension headers
- We have to trigger a null dereference somewhere in `tcpip!Ipv6pReassembleDatagram`

The next section will give a primer into how IPv6 fragmentation works to better understand how we can fulfill the above-mentioned requirements.

0.1.3 IPv6 fragmentation primer

When the size of a packet is larger than the Maximum transmission unit (MTU) of the outbound interface, IPv6 fragmentation is used. The MTU of most standard network equipment and desktop computers is 1500 bytes. Therefore if you have an IPv6 packet that is larger than 1500 bytes, the packet must be fragmented. This is done by splitting the packet into a number of fragments, that each has to be decorated with the IPv6 fragment header. This header is a

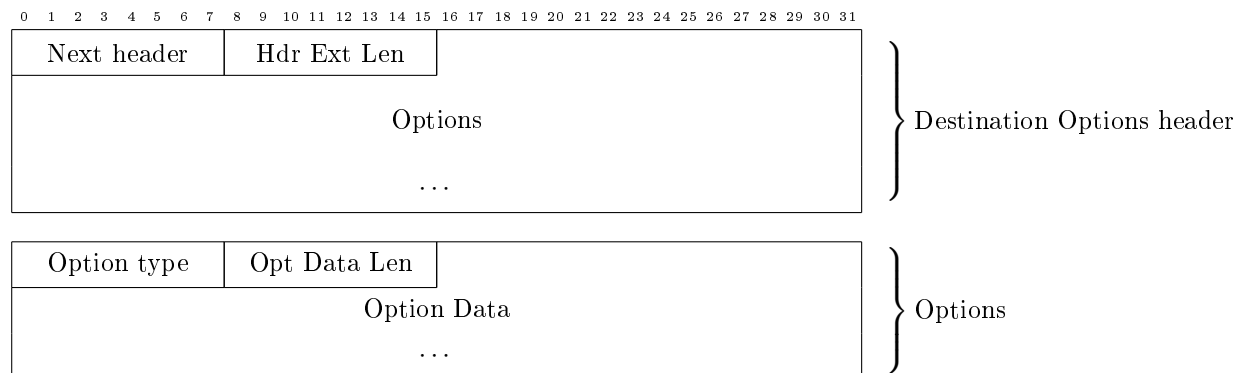
part of the specification for IPv6 Extension Headers[3, sec. 4.5]. The IPv6 Extension Headers specification specify a number of headers situated between the IPv6 header and the upper-layer header in a packet. The full list of extension headers can be seen in the following list:

1. Hop-by-Hop Options
2. *Fragment*
3. *Destination Options*
4. Routing
5. Authentication
6. Encapsulating Security Payload

As mentioned in section 0.1.1, the vulnerability is triggered when around 0xFFFF bytes of extension headers are present in the packet. Therefore, the following sections will describe both the *Destination Options* and *Fragment* extension headers in enough detail to support the exploitation of CVE-2021-24086.

IPv6 Destination Options extension header

IPv6 Destination Options are a way of defining options that should be handled by the destination node. In our case this would be the device that we are trying to attack using CVE-2021-24086. The specification can be seen on Figure 2 (IPv6 Destination Options Header [3, sec. 4.6]). The header is essentially structured as a list of options, where it is up to the receiver of a packet to support certain options.



Where

Next Header is an 8-bit selector identifying the initial header type of the Fragmentable part of the original packet.

Hdr Ext Len is an 8-bit unsigned integer describing the length of the Destination Option header in 8-octets units excluding the first 8 octets

Options is a variable-length field. See below

And

Option Type is an 8-bit identifier of the option type

Opt Data Len is an 8-bit unsigned integer describing the length of the *Data Option* field in octets

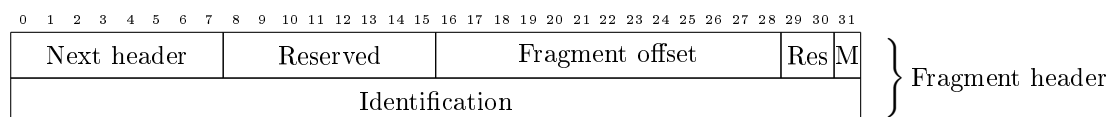
Options is a variable-length field with data specified by the option type

Figure 2: IPv6 Destination Options Header [3, sec. 4.6]

By default, only one option exist, the *PadN option*[3, sec. 4.2] which is used to create padding between two options. While this may not seem overly exciting, it is a very important part of how we can exploit CVE-2021-24086. Most other extension headers contain data that must be valid, such as routing options, which makes it hard to create a valid packet with around 0xFFFF bytes of extension headers. Destination Options does not have this limitation, as we can simply fill it with an arbitrary number of *PadN* options.

IPv6 Fragment extension header

Moving on to the IPv6 Fragment extension header, which, as mentioned earlier, is a header placed when you split an IPv6 packet into smaller fragments. IPv6 fragments are mostly used to send packets larger than the configured MTU, on either the sender or receiver side. The specification is detailed on figure Figure 3 (IPv6 Fragment Header [3, sec. 4.5]). The header contains an offset that points to where the fragment data fits into the entire packet.



Where

Next Header is an 8-bit selector identifying the initial header type of the Fragmentable part of the original packet.

Reserved is an 8-bit reserved field. Initialized to zero.

Fragment Offset is a 13-bit unsigned integer stating the offset.

Res is a 2-bit reserved field that is initialized to zero by the transmitter and ignored by the receiver.

M flag is a 1-bit boolean field describing if this is the last fragment. 1 = more fragments, 0 = last fragment.

Identification is a 32-bit identifier that is unique to fragments from the same package.

Figure 3: IPv6 Fragment Header [3, sec. 4.5]

Every packet that is fragmented has an unique identification, as specified in Figure 3 (IPv6 Fragment Header [3, sec. 4.5]). According to the specification[3, sec. 4.5], this identification must be different than any other fragmented packet sent recently¹.

A packet destined to be fragmented goes through two different processes, fragmentation and reassembly. Fragmentation happens on the sender side whereas reassembly is handled by the recipient of the packet.

¹Recently is very loosely defined by RFC 8200[3] as the "*maximum likely lifetime of a packet, including transit time from source to destination and time spent awaiting reassembly with other fragments of the same packet.*"[3, sec. 4.5]

Fragmentation is done by the sender and is a fairly simple concept. Looking at figure Figure 4 (IPv6 fragmentation[2]), it can be seen that an IPv6 packet contains two parts, an unfragmentable and a fragmentable part. The unfragmentable part is the IPv6 headers and the following two IPv6 extension headers, as they are processed by nodes en route:

- Hop-by-Hop Options Headers
- Routing Header

The rest of the IPv6 packet, including the Destination Options header, is handled as a fragmentable part.

Reassembly Reassembling the fragmented packet is done by the receiver and is essentially the fragmentation process in reverse. So here the receiver will convert a number of fragments into a single packet that can be handled as a standard IPv6 packet. The split of a fragmented packet can be seen on figure Figure 4 (IPv6 fragmentation[2]). Here it is easy to see that every fragment contains the unfragmentable part before any fragmented data.

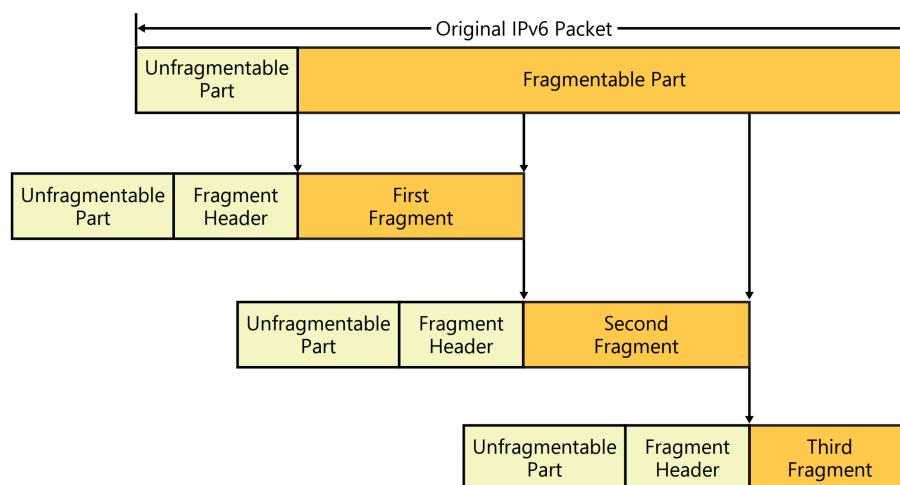


Figure 4: IPv6 fragmentation[2]

0.1.4 Root-cause analysis

At this point in the analysis the following relevant information has been presented to the reader:

1. The vulnerability happens when `tcpip.sys` reassembles a fragmented packet
2. The root cause of the vulnerability is a NULL pointer dereference in `Ip_v6ReassembleDatagram` of `tcpip.sys`

3. The packet should contain around 0xFFFF bytes of extension headers
4. Extension headers can be present both in the unfragmentable and the fragmentable part of the packet
5. The MTU limits how many bytes the unfragmentable part of the packet can contain
6. The Destination Options extension header is a good candidate for reaching 0xFFFF bytes
7. The Fragment extension header is needed to fragment the packet

To understand the root-cause of CVE-2021-24086 we must first understand how the fragmentable and unfragmentable data of the fragmented packet is handled in `Ipv6pReceiveFragment` and `Ipv6ReassembleDatagram`. If we start with `Ipv6pReceiveFragment`, we can see that a packet is reassembled when the total length of all fragment matches the expected length of the packet:

```
1  RtlCopyMdlToBuffer(netBuffer->MdlChain, netBuffer->DataOffset, v55,  
    ↪ netBuffer->DataLength, &v53);  
2  IppReassemblyInsertFragment(datagram, ippReassemblyLocation, NewIrql);  
3  IppIncreaseReassemblySize((struct_a1 *) (Blink + 20304), datagram,  
    ↪ netBuffer->DataLength + 256, netBuffer->DataLength);  
4  
5  if ( datagram->dataLength == datagram->fragmentableLength ) {  
6      Ipv6pReassembleDatagram(a1, datagram, v21);  
7  }  
8  else {  
9      IppCheckReassemblyQuota((PKSPIN_LOCK) (Blink + 20304));  
10 }
```

Listing 3: `Ipv6pReceiveFragment` packet reassembly logic

The check can be seen on line (5) of listing 3 where line (6) shows the call to `Ipv6ReassembleDatagram`. Once inside `Ipv6pReceiveFragment` we can see that both the unfragmentable and fragmentable lengths are saved to local variables as seen on listing 4

```
1 void __fastcall Ipv6pReassembleDatagram(__int64 a1, struct_datagram
   ↪ *datagram, char a3)
2 {
3     int unfragmentableHeaderLength; // er9
4     ulong BytesNeeded; // [rsp+48h] [rbp+10h]
5     int length; // [rsp+B8h] [rbp+20h]
6
7     ...
8
9     unfragmentableHeaderLength = datagram->unfragmentableHeaderLength;
10    length = unfragmentableHeaderLength + datagram->fragmentableLength;
11    BytesNeeded = unfragmentableHeaderLength + 40;
12
13    ...
14 }
```

Listing 4: Ipv6pReassembleDatagram length calculation

It's also important to notice the `BytesNeeded` variable which is equal to the size of unfragmentable header and the size of the Ipv6 header which is 40 bytes as seen on line (11). To understand the root cause, it is important to understand what will happen if the unfragmentable part of the header contains around 0xFFFF bytes. The calculation of `BytesNeeded` on line 11 also shows why it is only necessary to have *around* 0xFFFF bytes in the unfragmentable part.

Tracking down where `BytesNeeded` is used leads us to the code found in listing 5. This listing contains the code for obtaining a buffer to store the data for the unfragmentable part of the header. As it can be seen on line (9) and 19, this is where the `BytesNeeded` variable is used.

```
1 NetBufferList = (_NET_BUFFER_LIST *)NetioAllocateAndReferenceNetBufferA_
  ↳ ndNetBufferList(IppReassemblyNetBufferListsComplete, datagram,
  ↳ 0i64, 0i64, 0, 0);
2 if ( !NetBufferList )
3 {
4     ...
5     goto failure;
6 }
7
8 netBuffer = NetBufferList->FirstNetBuffer;
9 if ( NetioRetreatNetBuffer(netBuffer, (unsigned __int16)BytesNeeded, 0)
  ↳ < 0 )
10 {
11     IppRemoveFromReassemblySet((PKSPIN_LOCK)(v7 + 20304),
  ↳ (__int64)datagram, a3);
12     NetioDereferenceNetBufferList(NetBufferList, 0i64);
13
14     ...
15
16     goto memory_failure;
17 }
18
19 buffer = NdisGetDataBuffer(netBuffer, BytesNeeded, 0i64, 1u, 0);
```

Listing 5: Ipv6pReassembleDatagram NetBuffer null reference logic

The logic for listing 5 can be explained as such:

1. The NetBufferList is retrieved by NetioAllocateAndReferenceNetBufferA_ ndNetBufferList and checked for validity
2. The first NetBuffer is retrieved using NetioRetreatNetBuffer
 - Notice the cast to a unsigned 16 bit integer on line (9) wich will truncate the BytesNeeded.
3. NdisGetDataBuffer is used to retrieve a buffer.
 - Notice that BytesNeeded is *not* cast in this call on line 10.

Now the question is, what happens when NetioRetreatNetBuffer is invoked with a smaller value than NdisGetDataBuffer? The answer to that question is that NdisGetDataBuffer returns null. Later on in the function this buffer, which is null, is written to which will demonstrate that this indeed is a null pointer dereference. At this point we are presented with the root cause of the vulnerability, and can therefore move on to the process of triggering the vulnerability by sending a packet with about 0xFFFF extension headers in the unfragmentable part of the packet.

0.1.5 Triggering the vulnerability

To trigger CVE-2021-24086 a raw IPv6 packet has to be constructed which might not conform completely with the IPv6 specification. For this reason, it was decided to build the PoC using a combination of custom Python code and Scapy[11], which is a Python package used to craft network packets.

As explained in subsection 0.1.4 (Root-cause analysis) the unfragmentable part of the Ipv6 packet header is constrained by the size of the MTU, which is usually around 1500 bytes. In 2012 Antonios Atlasis highlighted a number of security issues present in implementations of IPv6 across different operating systems such as Windows, CentOS, Ubuntu and others[1]. In his paper, Antonios explain how two create *nested fragments* that allow one to embed a fragment inside another fragment. Figure 5 (Nested fragments[1]) shows how such a packet can be constructed.

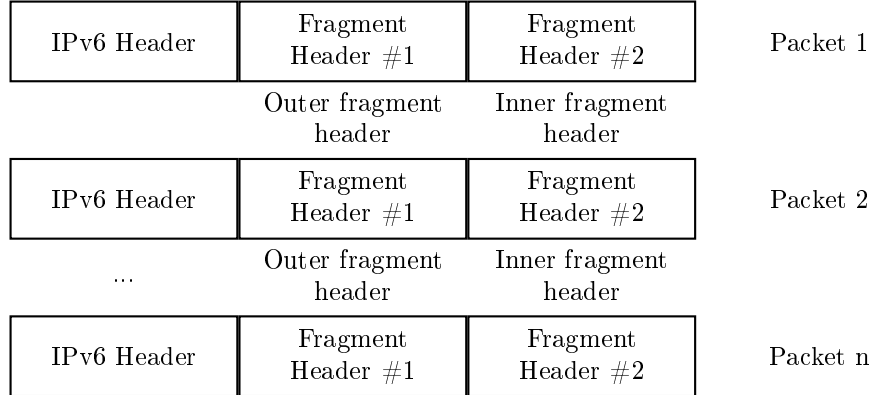


Figure 5: Nested fragments[1]

If we combine all the knowledge gained from the previous section the following IPv6 packet structure should produce a PoC that can be used to trigger CVE-2021-24086:

1. Create a long packet, *packet₁* with around 0xFFFF bytes of *Destination Option* header data. This packet should be fragmented using IPv6 fragments.
2. Construct *packet₂* as a IPv6 packet containing one fragment header with a unique fragment header id. The packet should also contain some data.
3. Add a fragment header to the end of the headers for *packet₁* with the fragment header id set to the fragment header id used in *packet₂*.
4. Send all fragments for *packet₁*

5. Send *packet*₂ which should trigger a reassembly of the nested fragments leading to Denial-of-Service (DoS)

Pseudo-code for this PoC can be seen in listing 6

```
1 first_fragment_id = random.randint()
2 second_fragment_id = random.randint()
3
4 packet1 = IPv6Header + IPv6ExtHdrDestOpt1 + IPv6ExtHdrDestOpt2 +
   ↪ IPv6ExtHdrDestOpt.. + IPv6ExtHdrDestOptN
5 packet1 += IPv6ExtHdrFragment(fragment_header_id = second_fragment_id)
6 packet1 += UDPPacket
7
8 packet1_fragments = fragment(packet1)
9
10 packet2 = IPv6Header + IPv6ExtHdrFragment + 'Dummy data'
11
12 send(packet1_fragments)
13 send(packet2)
```

Listing 6: Pseudo-code PoC for triggering CVE-2021-24086

An implementation of the pseudo-code can be found in Appendix ?? . As the PoC is hardcoded to run against the IPv6 multicast address, *ff02::1*, it will run against any machines present on the current IPv6 subnet. Running the PoC with a vulnerable Windows machine present will result in a Blue screen of death (BSoD) on the vulnerable machine as seen on Figure 6 (BSoD when running PoC), where it can also be seen that the crash originated from `tcpip.sys`.

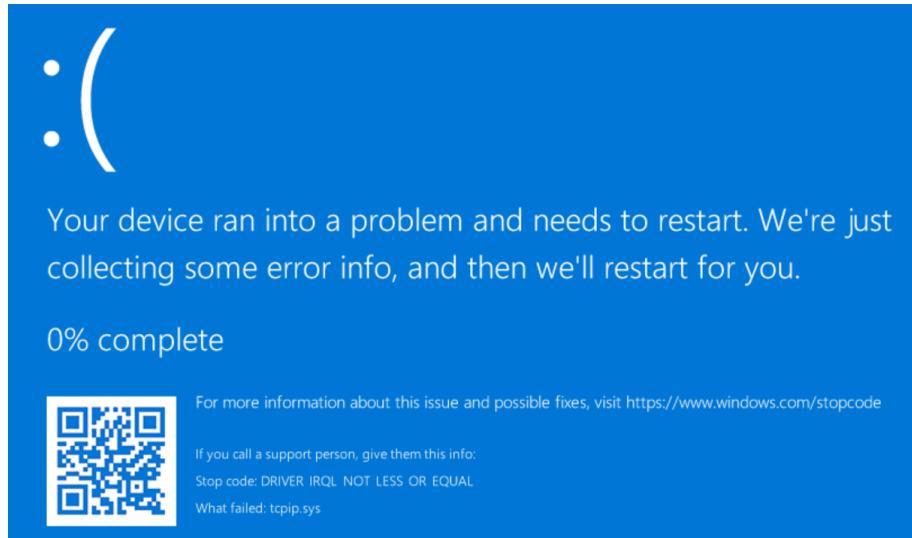


Figure 6: BSOD when running PoC

If the same PoC is run with a debugger attached, we get the output seen on listing 7 showing the details around the crash.

```

1  tcpip!Ipv6pReassembleDatagram+0x14f:
2  fffff801`0a2a937b 0f1100          movups  xmmword ptr [rax],xmm0 ds:00000000`00000000=????????????????????
3  Resetting default scope
4
5  STACK_TEXT:
6  tcpip!Ipv6pReassembleDatagram+0x14f          : fffff80a`9437e000 00000000`00000000
7  tcpip!Ipv6pReceiveFragment+0x84a             : fffff80a`9c1e3560 fffff801`0a300008
8  tcpip!Ipv6pReceiveFragmentList+0x42          : fffff80a`00000003 fffff801`0ab6ad00
9  tcpip!Ipv6pReceiveHeaderBatch+0x7f0b5        : fffff801`0a303000 fffff80a`943d18e0
10 tcpip!Ipv6pReceivePacketsCore+0x32f           : fffff80a`94a16690 fffff80a`94bcf510
11 tcpip!Ipv6pReceivePackets+0xc                 : fffff80a`94bcf510 00000000`00000000
12 tcpip!Ipv6pReceiveNonPreValidatedNetBufferListChain+0x26f : fffff801`0ab6b101 fffff80a`94cd4800
13 tcpip!FltReceiveNetBufferListChainCalloutRoutine+0x17c    : fffff80a`943c4c60 00000000`00000002
14 nt!KeExpandKernelStackAndCalloutInternal+0x78            : fffff801`0a18e420 fffff801`0ab6b358
15 nt!KeExpandKernelStackAndCalloutEx+0x1d                 : 00000000`c00000b5 fffff801`0ab6b528
16 tcpip!NetioExpandKernelStackAndCallout+0x8d              : 00000000`00000401 fffff801`0ab6b3c0
17 tcpip!FltReceiveNetBufferListChain+0x46d                : fffff80a`9408c801 00000000`00000001
18 NDIS!NdisMIndicateNetBufferListsToOpen+0x140            : fffff80a`94cd58a0 00000000`0000dd01
19 NDIS!NdisMTopReceiveNetBufferLists+0x22b                 : fffff80a`94b211a0 fffff801`0b586101
20 NDIS!NdisCallReceiveHandler+0x60                       : fffff80a`94bcf510 fffff801`0ab6b7a1
21 NDIS!NdisInvokeNextReceiveHandler+0x1df                 : 00000000`000078f5 00000000`00000401
22 NDIS!NdisMIndicateReceiveNetBufferLists+0x104           : fffff80a`94b7b520 fffff80a`94b7b520
23 kdnic!RxReceiveIndicateDpc+0x1e5                      : 00000000`00000002 fffff801`00003333
24 nt!KiProcessExpiredTimerList+0x172                    : 00000000`00000000 00000000`00000000
25 nt!KiRetireDpcList+0x5dd                                : 00000000`00000000 fffff801`026b0180
26 nt!KiIdleLoop+0x9e                                     : fffff801`0ab6c000 fffff801`0ab66000

```

Listing 7: Stacktrace when triggering CVE-2021-24086

Examining the stacktrace seen in listing 7 we can see that the crash happens at `Ipv6pReassembleDatagram+0x14f` coming from `Ipv6pReceiveFragment+0x84a` which matches the root-cause found in subsection 0.1.4 (Root-cause analysis). Looking at line (1) also shows the NULL pointer dereference as the instruction `movups` attempts to write to the address NULL.