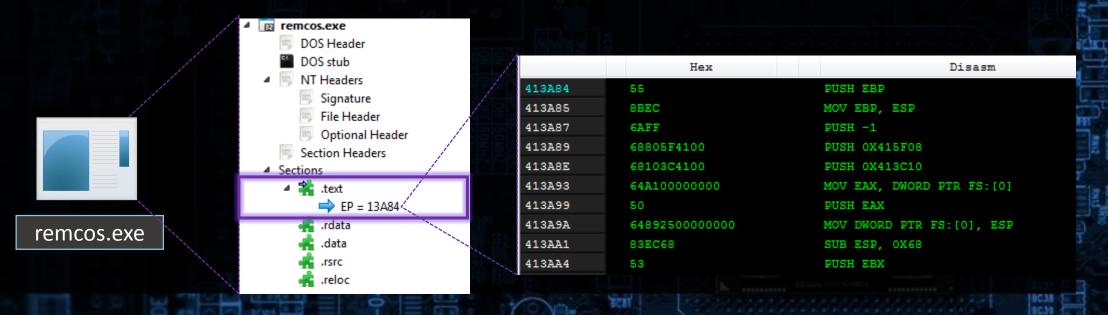


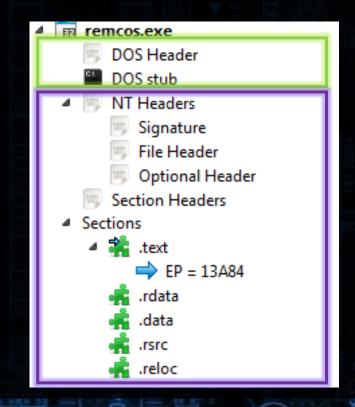
- PE (Portable Executable) is a native executable format on Windows
- PE files:
 - user mode: EXE, DLL
 - kernel mode: driver (.sys), kernel image (ntoskrnl.exe)
 - UEFI (run in SMM System Managemant Mode)
 - Also OBJ files have structures similar to PE

- PE (Portable Executable) contains information:
 - What to execute: the compiled code
 - How to execute: headers with data necessary for loading it



- PE format is based on a Unix format COFF that was used in VAX/VMS
- It was introduced as a part of specification Win32
- Throughout many years, the core of the format didn't change, only some new fields of some structures have been added
- Since introduction of 64 bit environment, PE needed to be adjusted to it: 64 bit PE was introduced
- Also, new variants have been introduced, like .NET PE containing additional structures with intermediate code and metadata

• PE file structure: the DOS part (legacy) and the Windows Part





• DOS Header: only e_magic, and e_lfnew must be filled:

```
typedef struct _IMAGE_DOS_HEADER {
                                    // DOS .EXE header
                                    // Magic number -----
        e_magic;
 WORD
 WORD
        e_cblp;
                                    // Bytes on last page of file
                                    // Pages in file
 WORD
        e_cpi
        e_crlci
 WORD
                                    // Relocations
        e_cparhdr;
 WORD
                                    // Size of header in paragraphs
        e_minalloci
                                    // Minimum extra paragraphs needed
 WORD
 WORD
        e_maxalloc;
                                    // Maximum extra paragraphs needed
                                    // Initial (relative) SS value
 WORD
        e_ssi
 WORD
        e_spi
                                    // Initial SP value
        e_csum:
 WORD
                                    // Checksum
 WORD
        e_ip;
                                    // Initial IP value
 WORD
        e_csi
                                    // Initial (relative) (S value
        e_lfarlc;
 WORD
                                    // File address of relocation table
 WORD
        e ovnoi
                                    // Overlay number
        e_res[4];
 WORD
                                    // Reserved words
        e_oemid;
                                    // OEM identifier (for e_oeminfo)
 WORD
 WORD
        e oeminfo;
                                    // OEM information; e_oemid specific
        e_res2[10];
 WORD
                                    // Reserved words
        e_lfanew;
                                    // File address of new exe header -----> Points to the NT header
 LONG
} IMAGE_DOS_HEADER, *PIMAGE_DOS_HEADER;
```

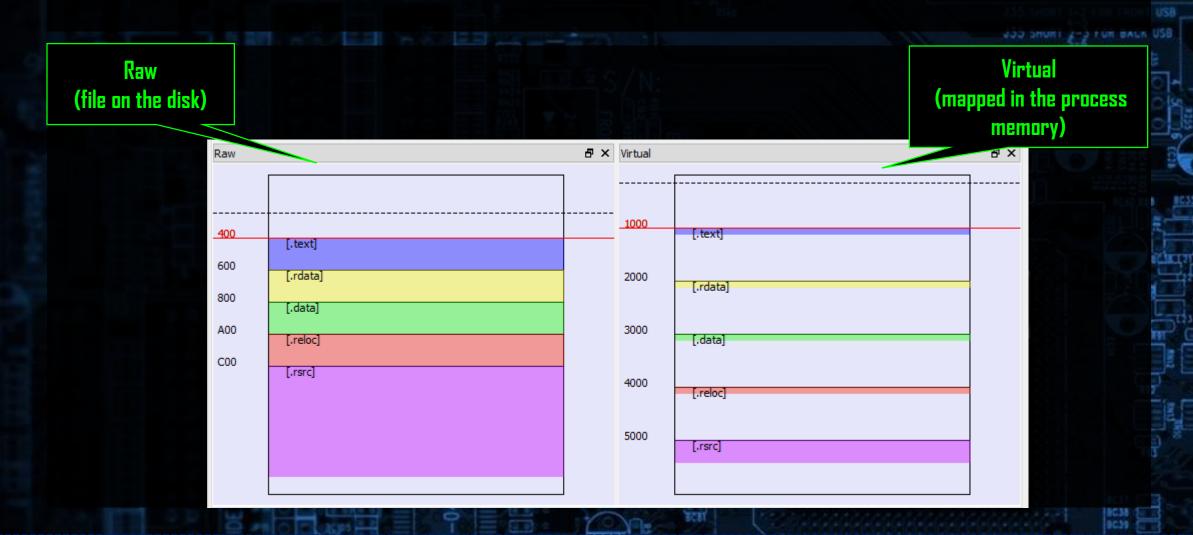
- PE is divided into sections with different permissions
- Sections introduce a logical layout of the binary, that compilers/linkers can follow
- Dividing PE on section improves security: the code is isolated from the data
- HOWEVER:
 - if DEP is disabled, page without execution permission can still be executed
 - The section containing the Entry Point will always be treated as executable

• PE sections are defined by sections header

Dis	asm: [.r	src] to [.relo	c] Gener	al DOS Hdr	Rich Hdr	File Hdr	Optional	Hdr	Section	Hdrs	Impo	rts	Res	ources
+ 23														
Nar	ne	Raw Addr.	Raw size	Virtual Addr.	Virtual Size	Chara	cteristics	Ptr t	o Reloc.	Num	. of Reloc.	Nur	m. of Line	enum.
A	.text	400	DA00	1000	D84E	60000	020	0		0		0		
	>		٨		٨	r-x								
A	.rdata	DE00	2C00	F000	2A40	40000	040	0		0		0		
	>		٨		٨	r								
A	.data	10A00	1000	12000	2C84	C0000	040	0		0		0		
	>		٨		٨	rw-								
A	.rsrc	11A00	200	15000	1E0	40000	040	0		0		0		
	>		٨		٨	r								
A	.reloc	11C00	1000	16000	F3E	42000	040	0		0		0		
	>		٨		٨	r								

- on the disk PE is stored in a raw format (the unit is defined by File Alignment)
- In memory PE is mapped to its virtual format (the unit is defined by Section Alignment) usually of the granularity of one page (0x1000)

Disasm: [.rsrc]	to [.reloc]	General	DOS Hdr	Rich Hdr		File Hdr	Optional Hdr	
Offset	Name			Value Value				
118	Base of Da	ta			F000			
11C	Image Base	2			400000			
120	Section Ali	gnment		1000				
124	File Alignn	nent			200	J		

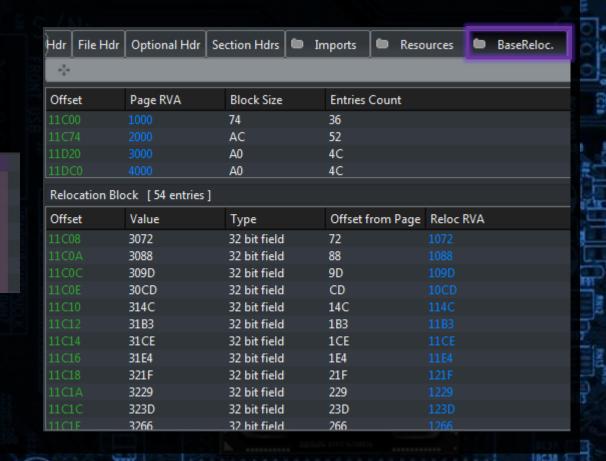


• The most information lies in data directories

Disas	m: [.rsrc]	to [.reloc]	General	DOS Hdr	Rich Hdr	File Hdr	Optional Hdr	Section Hdrs		
Offse	t	Name		-	Value		Value			
		Size of Hea	p Reserve		100000					
	54	Size of Hea	p Commi	t	1000					
		Loader Flag	gs		0					
15		Number of	RVAs and	Sizes	10					
A		Data Direct	tory		Addres	S	Size			
	160						0 28 1E0			
	168 Import Directory									
		Exception I	Directory				0			
	180	Security Di	rectory				0			
		Base Reloc	ation Tabl	e			9AC			
	190	Debug Dire	ectory				0			
		Architectu		Data						
		RVA of Glo					0			
		TLS Directo	ory				0			
	1B0	Load Confi		Directory			40			
	1B8		-	ory in head	ers 0		0			
		Import Add	dress Table	e			138 0			
		Delay Load	Import D	escriptors						
	1D0	.NET heade					0			

Relocation Table

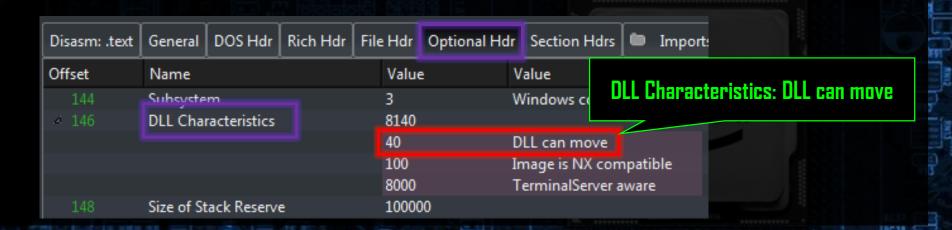
۵		Data Directory	Address	Size
	160	Export Directory		0
		Import Directory		28
		Resource Directory		1E0
		Exception Directory		0
	180	Security Directory		0
	188	Base Relocation Table		9AC
	190	Debug Directory		0



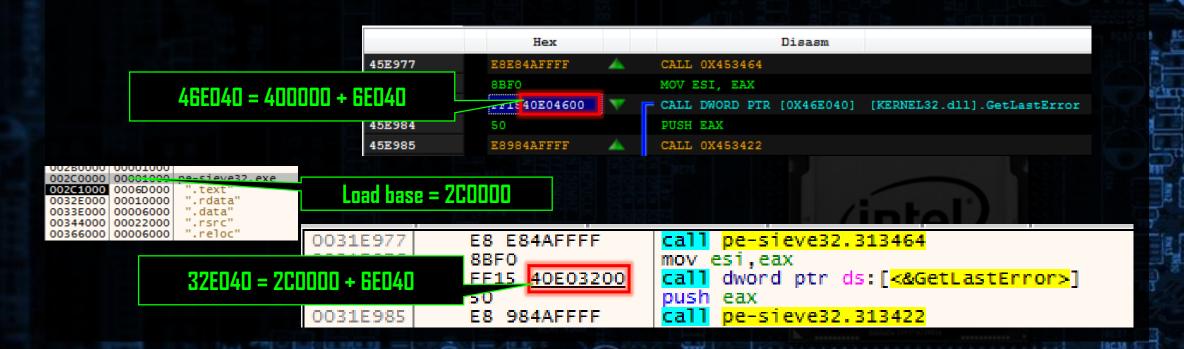
- 1. PE comes with some default base address in the header
- 2. All the absolute addresses inside the PE assume that it was loaded at this base



- In the past EXEs were usually loaded at their default base (only DLLs didn't have to)
- Nowadays most PEs load at a dynamic base (due to ASLR)
- A flag in the header determines if a dynamic base will be used



 If the PE was loaded at a different base than the one defined in the header, all its fields using absolute addresses must be recalculated (rebased)



•How does PE know where are the fields that needs to be rebased?

- •How does PE know **where** are the fields that needs to be **rebased**?
- •They are listed in the Relocation Table!

- Let's open one of our sample PEs in PE-bear and see the relocation table
- Check the code snippet to see how the relocation table is processed

Exercise time...

Basics of a PE file: Imports & Exports

Most executables use some functions exported by other modules (external libraries)

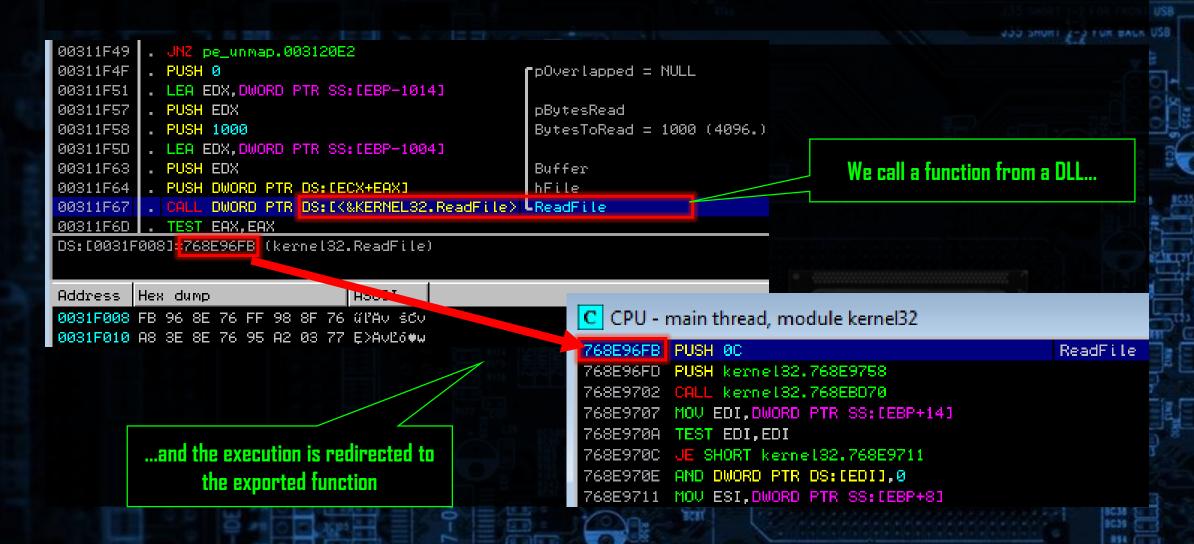
- 1. If we use a static library, the linker will automatically add the external code into our PE
- 2. If we use a dynamic library (DLL), the used functions will be listed in the Import Table of our PE, and dynamic linking will be done when the PE is loaded
- 3. Alternatively, we can load a DLL by ourselves using LoadLibrary and fetch the exported function via GetProcessAddress

• Export Table

				-1			me :	-	the s		
Disasm: .tex	t General	DOS Hdr	Rich Hdr	File Hdr	Optional Hdr	Section Hdrs	Exports	Imports	Reso		
+**											
Offset	Name		,	Value	Meaning						
B47C4	Characteri	stics	(0							
B47C8	TimeDate9	Stamp		4CE78B54	sobota, 20	0.11.2010 08:48:	20 UTC				
B47CC				0							
B47CE	-			0							
B47D0	Name			B8502	KERNEL32	2.dll					
B47D4	Base			1							
B47D8	NumberOfFunctions			54F							
B47DC	NumberOfNames			54F	F						
B47E0	AddressOf	Functions		B4FEC	4FEC						
B47E4	AddressOf	Names		6528							
B47E8	AddressOf	NameOrdinals	I	B7A64							
Exported Fund	ctions [1359	entries]									
Offset C	Ordinal	Function RVA	Name RV	/A Name			Forwarder				
B481C D)	BBFFB	B85E3	AddRe	fActCtx						
B4820 E	1	36399	B85F0	AddSI)ToBoundaryDes	scriptor					
B4824 F	8BD00 B860B			AddSecureMemoryCacheCallback							
B4828 1	0 BEE06 B8628			AddVe	ctoredContinue	Handler	NTDLL.RtlAddVectoredContinueHandler				
B482C 1	11 BEE2A B8643			AddVe	ctoredException	Handler	NTDLL.RtIAddV	ectoredExceptio	nHandler		
B4830 1	12 5DF86 B865F				Calendar Date			·			
B4834 1	3	AC328	B8672	•	AllocConsole						
B4838 1	4 9	97CA8	B867F	Allocat	AllocateUserPhysicalPages						

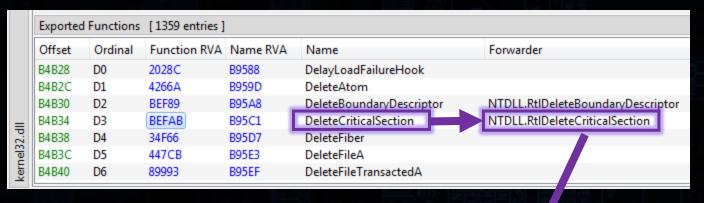
- 1. DLLs are libraries of functions for other PEs to use
- 2. An Export Table is a catalogue allowing to find and use a particular function

Offset	Ordinal	Function RVA	Name RVA	Name ReadConsoleW				
B56E0	3BE	60E73	BCF4F					
B56E4	3BF	42C62	BCF5C	ReadDire	ectoryChangesW			
B56E8	3C0	496FB	BCF72	ReadFile				
B56EC	3C1	63D99	PCF7B		Hex		Disasm	
B56F0	3C2	31B14	BCF86	77E296FB	◆ 6A0C		PUSH OXC	ReadFil
B56F4	3C3	3C1CE	BCF96	77E296FD	685897E277		PUSH 0X77E29758	
B56F8	3C4	9851F	BCFA8			W		
B56FC	3C5	4CB4F	BCFC0	77E29702	E869260000	W	CALL 0X77E2BD70	
B5700	3C6	42D7C	BCFCC	77E29707	8B7D14		MOV EDI, DWORD PTR [EBP + 0X	.14]
B5704	3C7	40D25	BCFDC	77E2970A	85FF		TEST EDI, EDI	
B5708	3C8	A8CD5	BCFEC	77E2970C	7403	A.	JE SHORT 0X77E29711	
B570C	3C9	36644	BCFFC	77E2970E	832700		AND DWORD PTR [EDI], 0	
25.00		300.7	20.10	77E29711	887508		MOV ESI, DWORD PTR [EBP + 8]	



- 1. Functions can be exported by a name or by ordinal (a number)
- 2. Some exports can be forwarded (pointing to other functions, in other DLLs)

• Forwarded functions



Exported Fu	inctions [1990 e	ntries]		
Offset	Ordinal	Function RVA	Name RVA	Name
3614C	2E6	556D1	3ECD1	RtlDeleteBoundaryDescriptor
36150	2E7	59AC5	3ECED	RtlDeleteCriticalSection
36154	2E8	5DD50	3ED06	RtlDeleteElementGenericTable
36158	2E9	DC18	3ED23	RtIDeleteElementGenericTableAvI
3615C	2EA	76E87	3ED43	RtlDeleteHashTable
36160	2EB	B6994	3ED56	RtlDeleteNoSplay
4			!!!	

	Disasm	Hint
i	🏫 MOV EDI, EDI	${\tt RtlDeleteCriticalSection}$
	PUSH EBP	
	MOV EBP, ESP	
L.	PUSH -2	
:	PUSH 0X77F10DE8	
	PUSH 0X77EDE0ED	

59AC5 59AC7 59AC8 59ACA 59ACC 59AD1

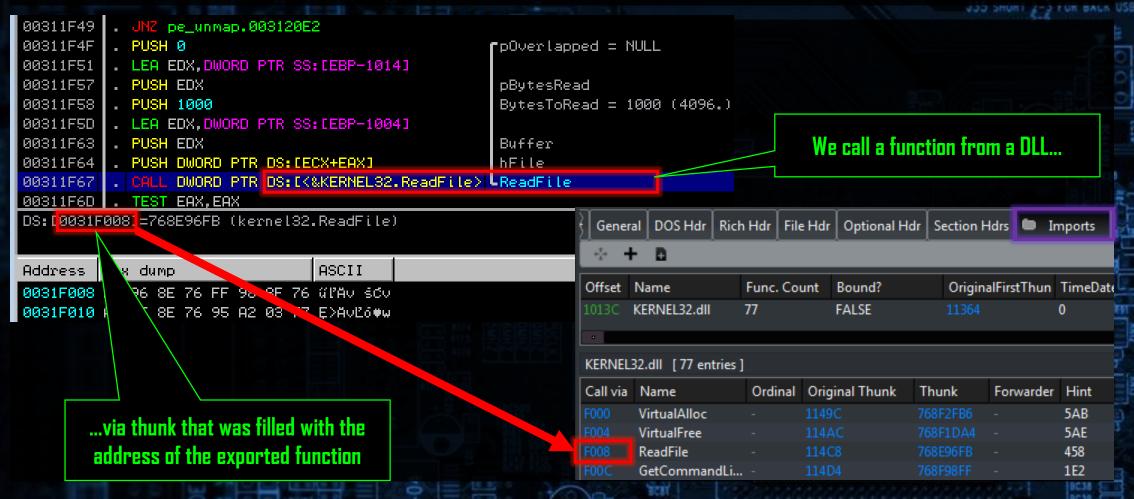
• Import Table

Disasm: .rd	lata General	DOS Hdr	Rich Hdr	File Hdr	Optional Ho	lr Se	ction Hdrs	Imports	Resources	BaseRel		
	D							-				
Offset	Name	Func. Cour	nt Bound?	Orig	inalFirstThun	TimeD	ateStamp	Forwarder	NameRVA	FirstThunk		
22474	KERNEL32.dll	93	FALSE	2309	С	0		0	23328	1D000		
KERNEL32.dl	[93 entries]											
Call via	Name	Ordinal	Original Thu	ınk Th	unk For	warder	Hint					
1D000	${\sf CreateDirectoryA}$	-	23214	232	214 -		C1					
1D004	CloseHandle	-	23228	232	228 -		8E					
1D008	GetLastError	-	23236	232	236 -		26A					
1D00C	OpenProcess	-	23246	232	246 -		408					
1D010	VirtualFree	-	23254	232	254 -		5AE					
1D014	CreateToolhelp	-	23262	232	.62 -		10A					
1D018	Module32First	-	2327E	232	?7E -		3DF					
1D01C	Module32Next	-	2328E	232	.8E -		3E1					
1D020	CreateFileA	-	2329E	232	.9E -		CE					
1D024	GetFileSize	-	232AC	232	AC -		254					
1D028	MapViewOfFile	-	232BA	232	2BA -		3DB					
1D02C	${\sf UnmapViewOfF}$	-	232CA	232	2CA -		593					
1D030	CreateFileMapp	-	232DC	232	DC -		CF					

https://github.com/hasherezade/malware training vol1/blob/main/exercises/module1/lesson2 pe/pe snippets /imports load.h

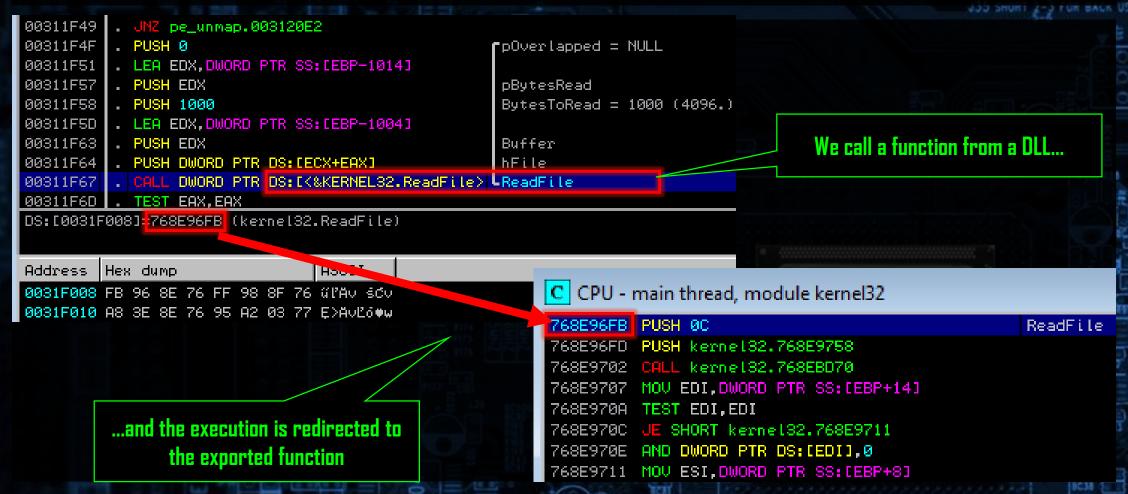
- Dynamic linking is done when a PE is loaded
- The loader walks through the Import Table of the PE
 - loads needed DLLs
 - searches the imported functions in the export table of the DLL
 - fills the thunks via which the PE is going to make calls to the exported functions with appropriate addresses

Basics of a PE file: Imports & Exports

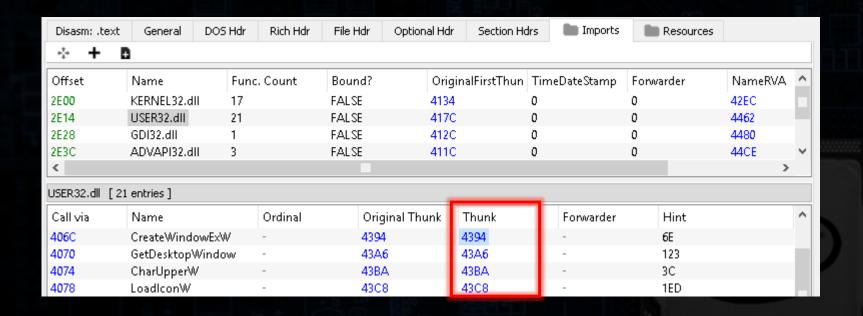


BBC2

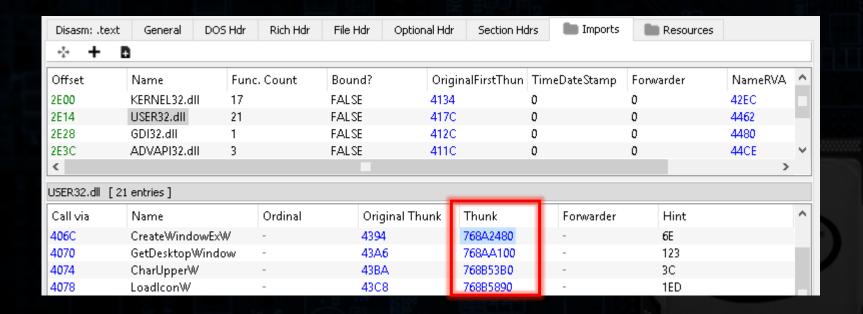
Basics of a PE file: Imports & Exports

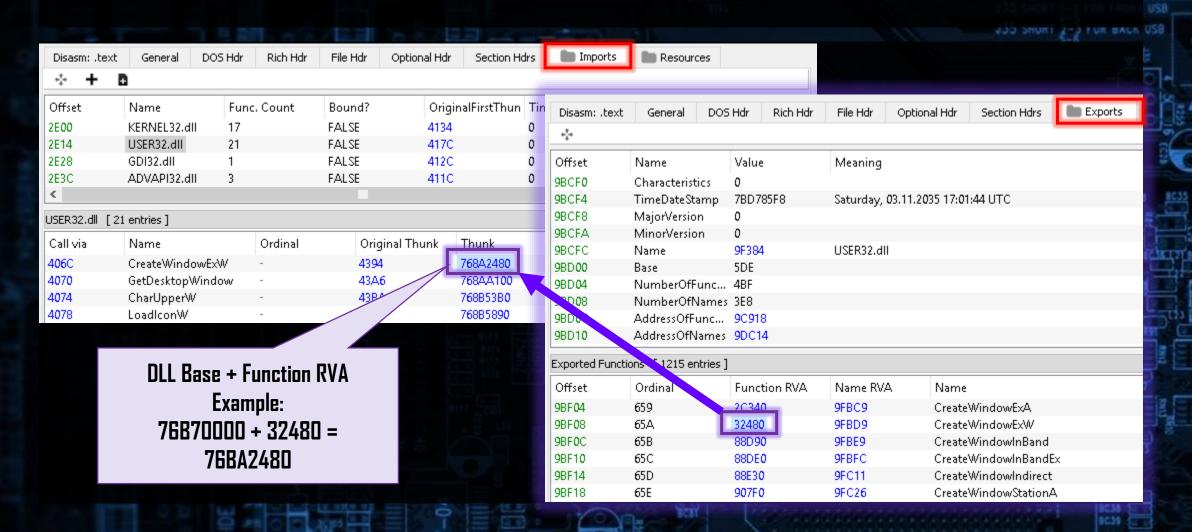


• Raw: before filling imports

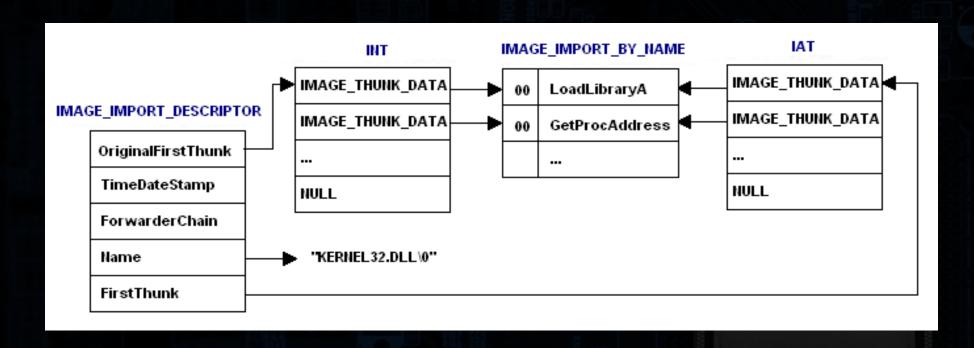


Loaded: after filling imports – thunks are filled with addresses of exported functions





• Import Table: structure



- Let's open one of our sample PEs in PE-bear and see the import table. Find the corresponding DLLs and their exports.
- Check the code snippets to see how the import and export tables are processed

Exercise time...



- Compile the given code of a custom PE loader and get familiar with it
 - https://github.com/hasherezade/malware_training_vol1/tree/main/exercises/module1/lesson2 pe



- Classic articles about PE by Matt Pietrek:
 - https://bytepointer.com/resources/pietrek_in_depth_look_into_pe_format_ptl.htm -
 - https://bytepointer.com/resources/pietrek in depth look into pe format pt2.htm
 - https://docs.microsoft.com/en-us/previous-versions/ms809762(v=msdn.10)?redirectedfrom=MSDN