

# NYU Poly Reverse Engineering Lecture

Aaron Portnoy

TippingPoint Security Research

Peter Silberman

Mandiant Engineering and Research

# Outline

## Agenda

Software auditing and reverse engineering on Windows

### **SESSION ONE**

Auditing methodologies

Tools of the trade

Disassembling and IDA Pro

### **SESSION TWO**

Reversing styles and techniques

Vulnerability classes

Vulnerability analysis and debugging

Automation (if time permits)

Focus on vulnerability discovery and analysis

# Introduction to RE

## Why is this skill set valuable?

- Source is not often available for Windows applications

- Eliteness factor

  - Often exponentially more difficult than source auditing

  - Find bugs where few others are comfortable

- Homebrew patches

- Patch analysis

  - e.g. Reversing Microsoft patches to discover root cause

  - 1<sup>st</sup> with a new exploit or Metasploit module?

  - 1<sup>st</sup> with a new signature for an AV or IPS?

- Binary code is a \$ goldmine \$

  - Bugs exist for long periods of time in binary code

# Introduction to RE (cont.)

## **Why is this skill set valuable (cont.)?**

- Knowledge is portable

  - Apply techniques to wide array of tasks

  - New architectures become approachable

- Developing countries rely on it

  - Why engineer from scratch when you can copy

  - Reverse for security, ensure there are no backdoors

  - In the US we take these things for granted

# Introduction to RE (cont.)

## It's not always about the assembly

Reversing is the process by which you attempt to understand the system

- Operating system
- Software
- Hardware
- Plane, train, auto, anything that was engineered

Reverse the *system* as a whole, helps locate trust boundaries

## Malkovich malkovich, malkovich?

GOAL: Get inside the developer's head

*Reverse engineer intentional behavior to determine how to deviate execution from that intention*

# Overview of Approach

## **General Steps, pre-disassembling**

- Examine system behavior
- Enumerate components
- Determine relationships
- Determine trust
- Locate and probe inputs

**Many tools to aid in this process**

We'll cover some of these

# Step by step

## **Examining system behavior**

Documentation

Install & use the product!

Support forums

Exploitable bugs might be “annoyances” to regular users

# Step by step (cont.)

## Enumerating components

### Documentation

Product prerequisites (Java, .NET, ...)

### MSRPC

Process Explorer  
rpcdump, rpcinfo  
mIDA

ActiveX, Codecs, File formats, Protocol Handlers, ...

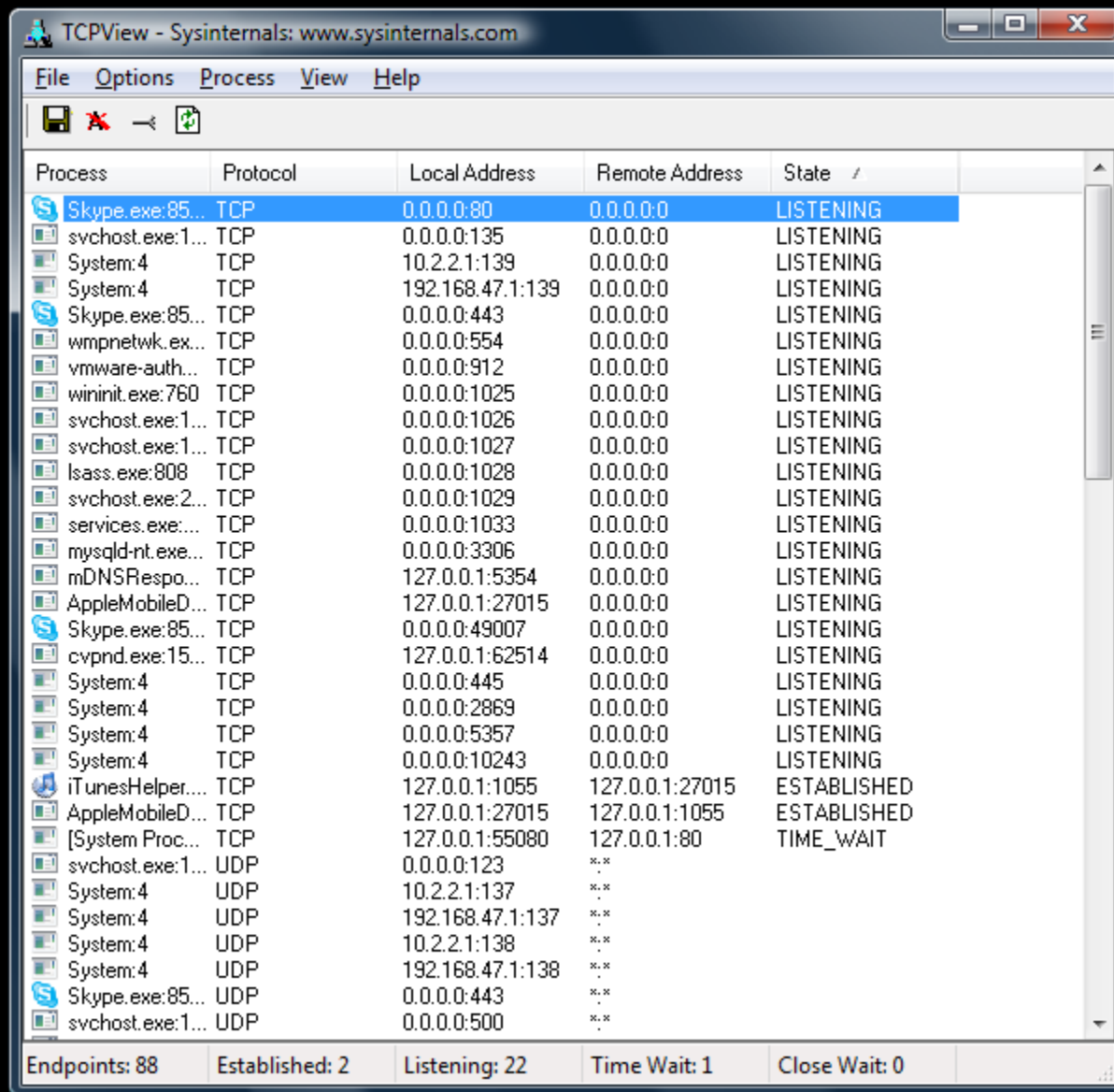
RegMon  
FileMon  
ProcessMon

### Services

services.msc / TCPView



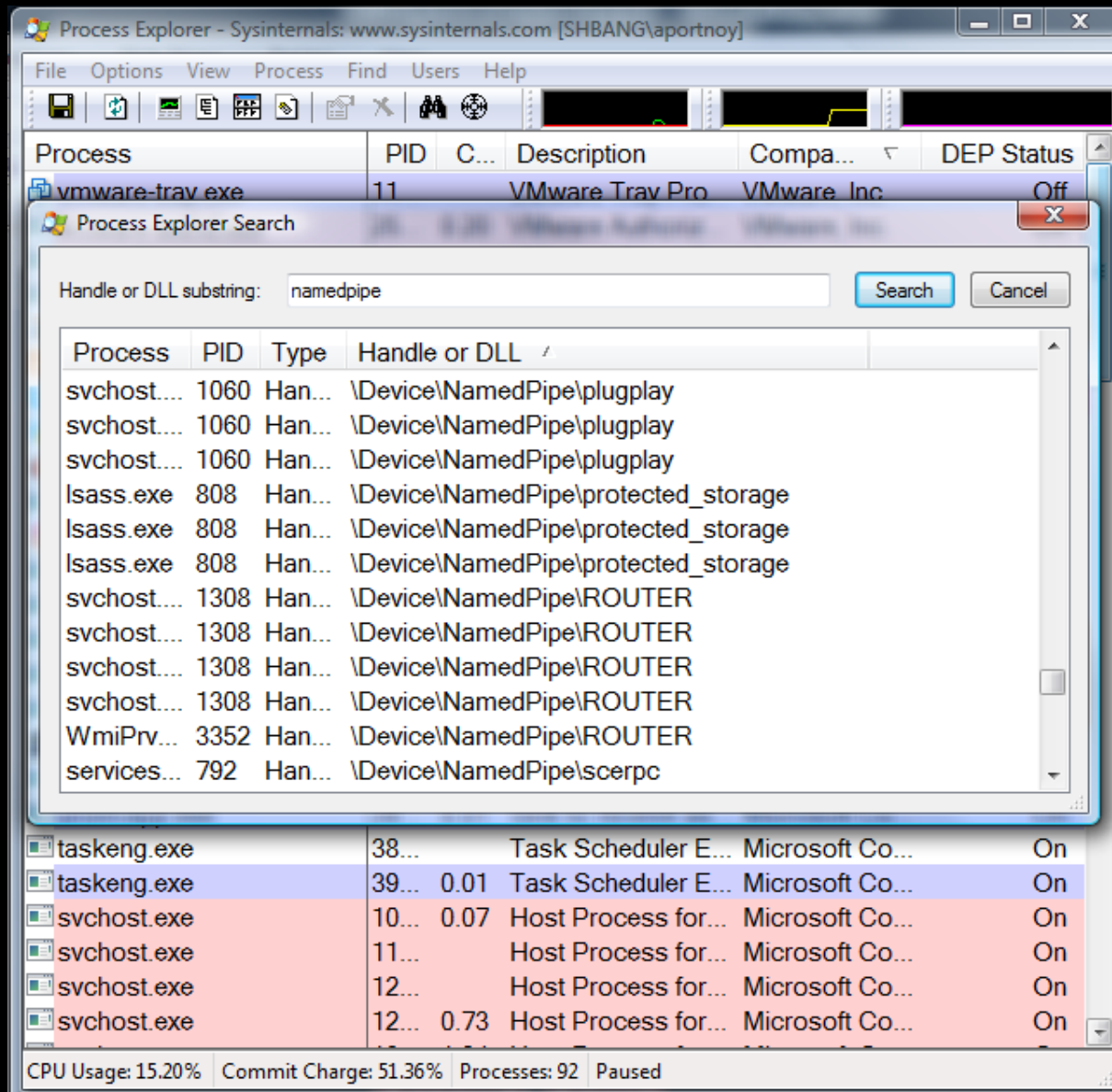
# TCP View from MS



Process	Protocol	Local Address	Remote Address	State
Skype.exe:85...	TCP	0.0.0.0:80	0.0.0.0	LISTENING
svchost.exe:1...	TCP	0.0.0.0:135	0.0.0.0	LISTENING
System:4	TCP	10.2.2.1:139	0.0.0.0	LISTENING
System:4	TCP	192.168.47.1:139	0.0.0.0	LISTENING
Skype.exe:85...	TCP	0.0.0.0:443	0.0.0.0	LISTENING
wmpnetwk.ex...	TCP	0.0.0.0:554	0.0.0.0	LISTENING
vmware-auth...	TCP	0.0.0.0:912	0.0.0.0	LISTENING
wininit.exe:760	TCP	0.0.0.0:1025	0.0.0.0	LISTENING
svchost.exe:1...	TCP	0.0.0.0:1026	0.0.0.0	LISTENING
svchost.exe:1...	TCP	0.0.0.0:1027	0.0.0.0	LISTENING
lsass.exe:808	TCP	0.0.0.0:1028	0.0.0.0	LISTENING
svchost.exe:2...	TCP	0.0.0.0:1029	0.0.0.0	LISTENING
services.exe...	TCP	0.0.0.0:1033	0.0.0.0	LISTENING
mysqld-nt.exe...	TCP	0.0.0.0:3306	0.0.0.0	LISTENING
mDNSRespo...	TCP	127.0.0.1:5354	0.0.0.0	LISTENING
AppleMobileD...	TCP	127.0.0.1:27015	0.0.0.0	LISTENING
Skype.exe:85...	TCP	0.0.0.0:49007	0.0.0.0	LISTENING
cvpnd.exe:15...	TCP	127.0.0.1:62514	0.0.0.0	LISTENING
System:4	TCP	0.0.0.0:445	0.0.0.0	LISTENING
System:4	TCP	0.0.0.0:2869	0.0.0.0	LISTENING
System:4	TCP	0.0.0.0:5357	0.0.0.0	LISTENING
System:4	TCP	0.0.0.0:10243	0.0.0.0	LISTENING
iTunesHelper....	TCP	127.0.0.1:1055	127.0.0.1:27015	ESTABLISHED
AppleMobileD...	TCP	127.0.0.1:27015	127.0.0.1:1055	ESTABLISHED
[System Proc...	TCP	127.0.0.1:55080	127.0.0.1:80	TIME_WAIT
svchost.exe:1...	UDP	0.0.0.0:123	..	
System:4	UDP	10.2.2.1:137	..	
System:4	UDP	192.168.47.1:137	..	
System:4	UDP	10.2.2.1:138	..	
System:4	UDP	192.168.47.1:138	..	
Skype.exe:85...	UDP	0.0.0.0:443	..	
svchost.exe:1...	UDP	0.0.0.0:500	..	

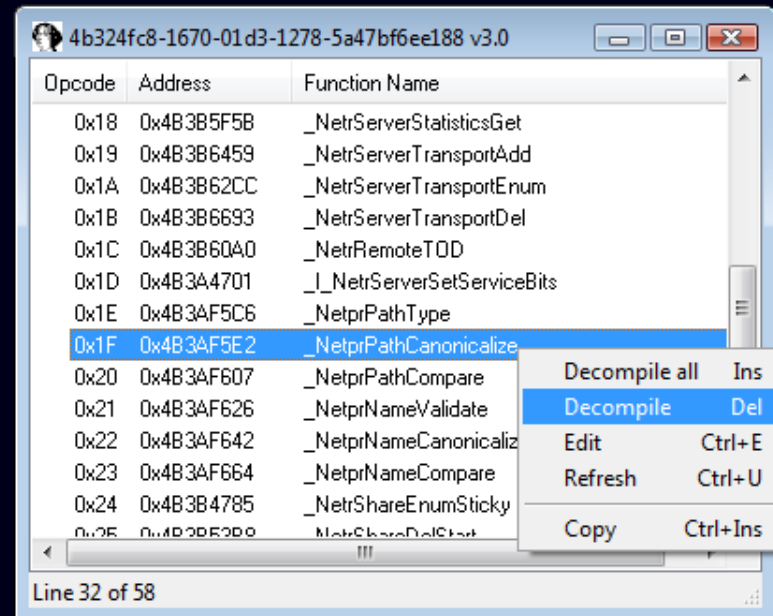
Endpoints: 88   Established: 2   Listening: 22   Time Wait: 1   Close Wait: 0

# Process Explorer from MS

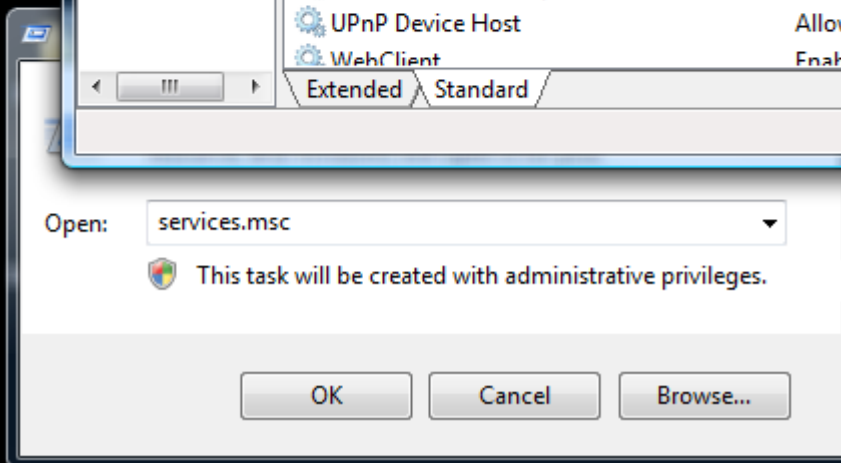
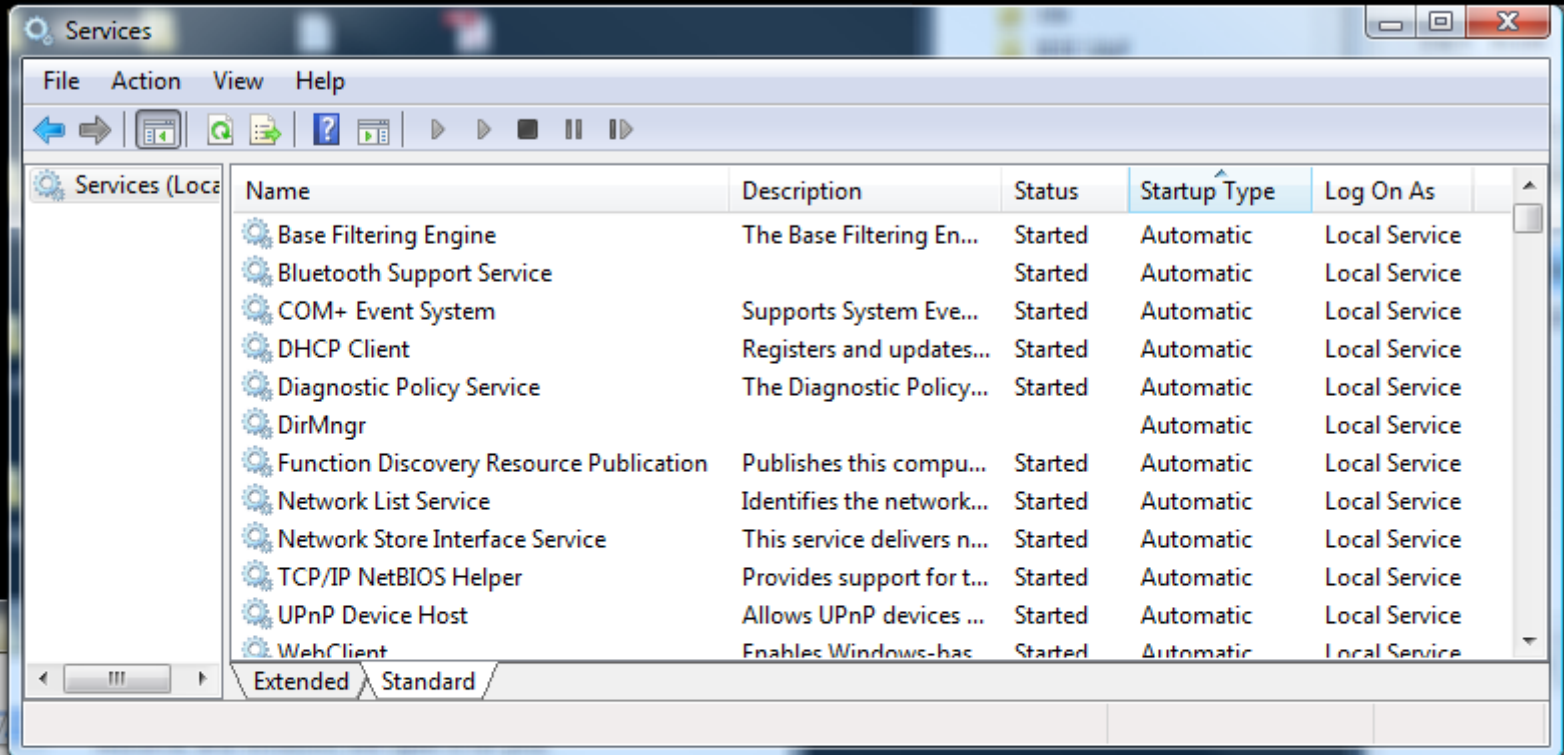


# mIDA from Tenable (MSRPC)

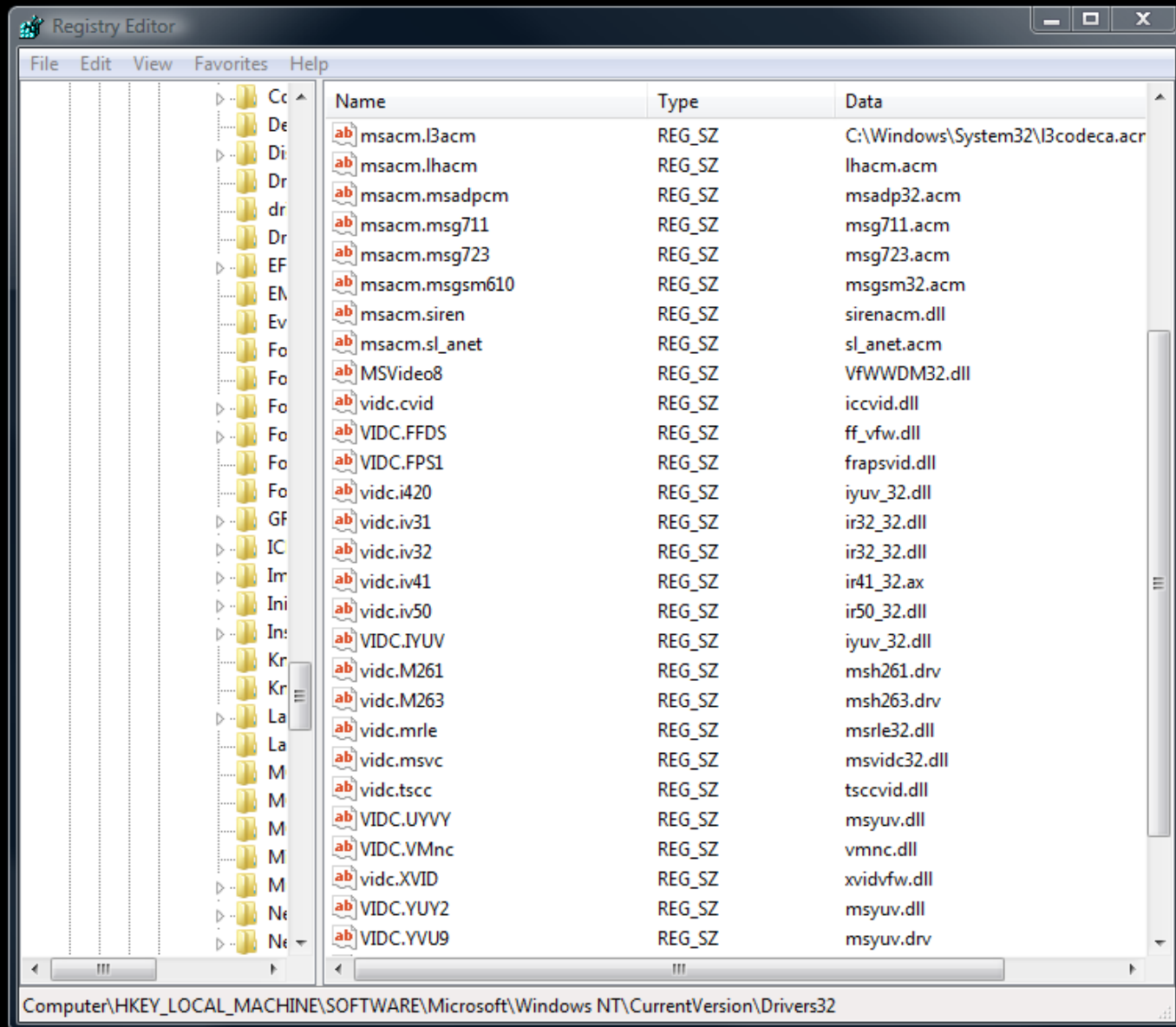
```
.text:4B3AF5E2
.text:4B3AF5E2 ; ===== S U B R O U T I N E =====
.text:4B3AF5E2
.text:4B3AF5E2 ; Attributes: bp-based frame
.text:4B3AF5E2
.text:4B3AF5E2 ; __stdcall NetprPathCanonicalize(x, x, x, x, x, x, x)
.text:4B3AF5E2 _NetprPathCanonicalize@28 proc near ; DATA XREF: .text:4B3A175C↑o
.text:4B3AF5E2
.text:4B3AF5E2 arg_4 = dword ptr 0Ch
.text:4B3AF5E2 arg_8 = dword ptr 10h
.text:4B3AF5E2 arg_C = dword ptr 14h
.text:4B3AF5E2 arg_10 = dword ptr 18h
.text:4B3AF5E2 arg_14 = dword ptr 1Ch
.text:4B3AF5E2 arg_18 = dword ptr 20h
.text:4B3AF5E2
.text:4B3AF5E2 mov edi, edi
.text:4B3AF5E4 push ebp
.text:4B3AF5E5 mov ebp, esp
.text:4B3AF5E7 push [ebp+arg_18]
.text:4B3AF5EA push [ebp+arg_14]
.text:4B3AF5ED push [ebp+arg_10]
.text:4B3AF5F0 push [ebp+arg_C]
.text:4B3AF5F3 push [ebp+arg_8]
.text:4B3AF5F6 push [ebp+arg_4]
.text:4B3AF5F9 call _NetpwPathCanonicalize@24 ; NetpwPathCanonicalize(x, x,
```



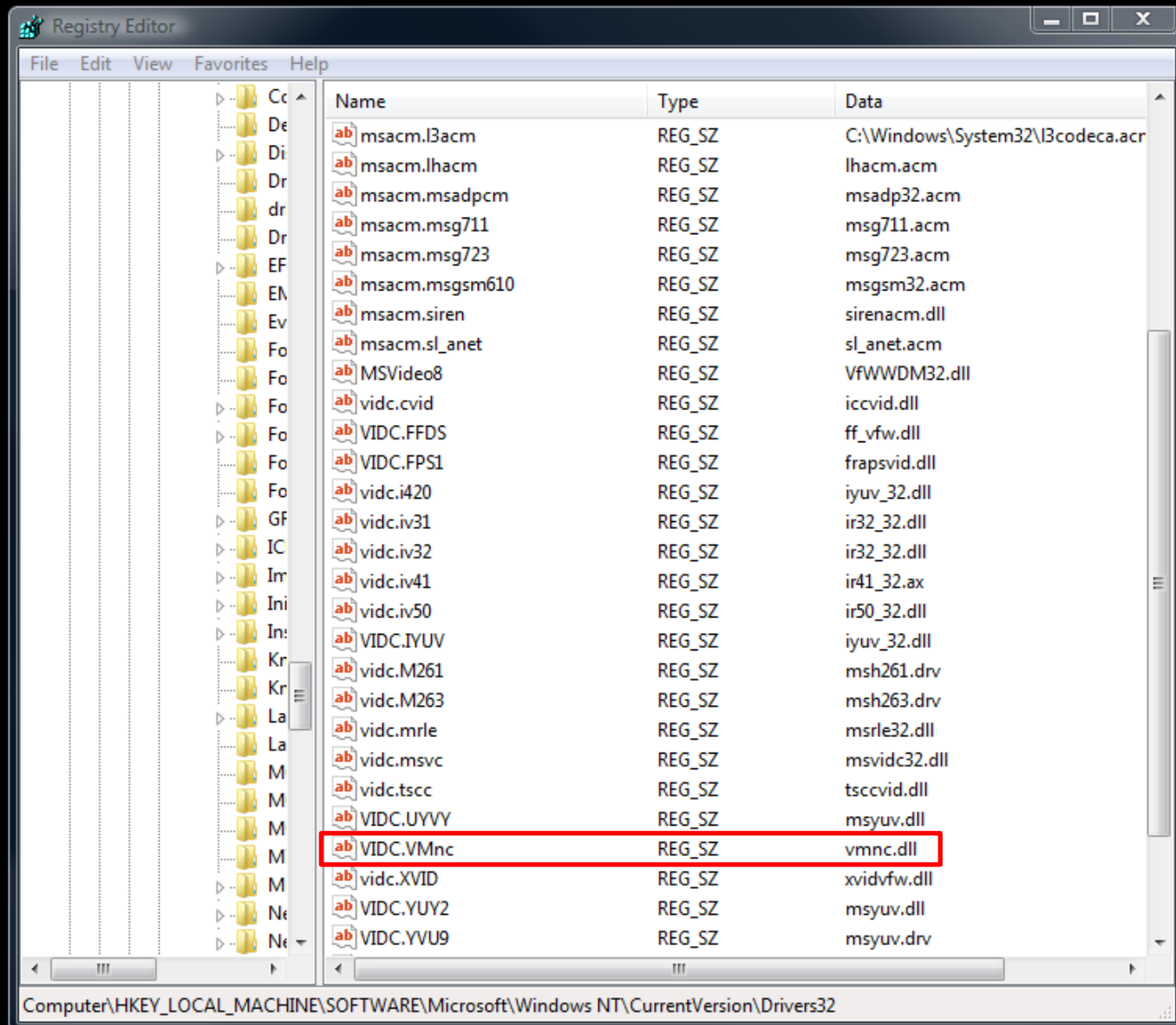
# Services



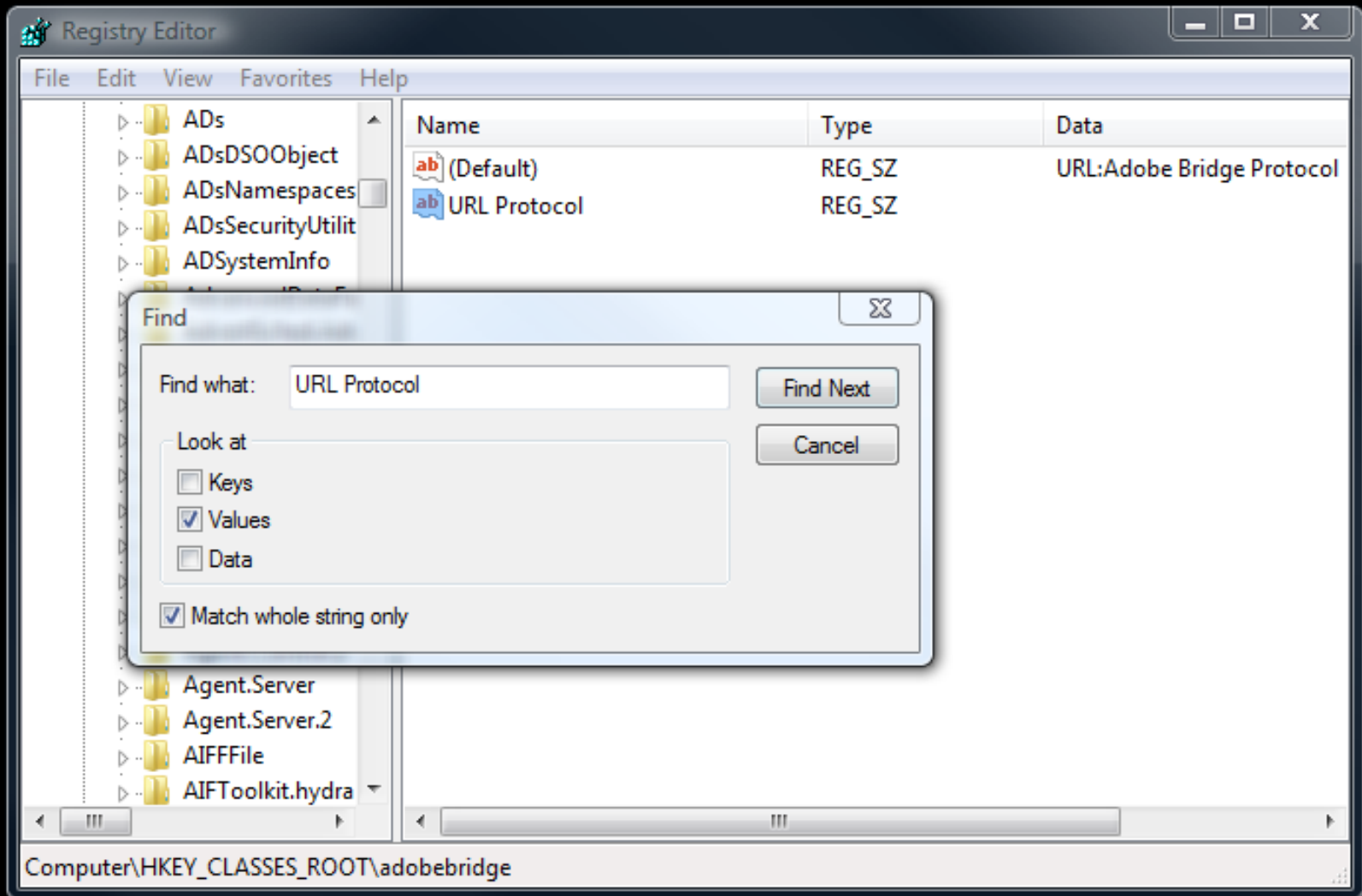
# Registered Codecs



# Registered Codecs



# Registered Protocol Handlers





# ActiveX Controls

ComRaider - [142 classes returned]

IDEFENSE  
A VeriSign Company

LABS

Start

View

Date	GUID	ProgID	Server	Description
11.24...	{0006F023-000...	RECIP.RecipCtl.1	ole32.dll	Microsoft Office Outlook Recipient Control
11.24...	{0006F024-000...	DOCSITE.Doc...	ole32.dll	Microsoft Office Outlook Rich Format Control
11.24...	{0149EEDF-D0...	MSVidCtl.MSVI...	C:\Windows\S...	XDS Feature Segment
11.24...	{02BF25D5-58C...	QuickTime.Qui...	C:\Program File...	QuickTime Object
11.24...	{0369B4E5-45...	BDATuner.Cha...	C:\Windows\S...	BDA Tuning Model Channel Tune Request
11.24...	{0369B4E6-45...	BDATuner.ATS...	C:\Windows\S...	BDA Tuning Model ATSC Channel Tune Request
11.24...	{03C06416-D1...	BDATuner.Digit...	C:\Windows\S...	BDA Tuning Model Digital Cable Locator
11.24...	{055CB2D7-29...	BDATuner.MP...	C:\Windows\S...	BDA Tuning Model MPEG2 Component Class (Broadcast Substream)
11.24...	{06290BD8-48...	ScriptletHandler...	C:\Windows\sy...	Constructor for Scriptlet Automation Handler
11.24...	{06290BD9-48...	ScriptletHandler...	C:\Windows\sy...	Constructor for Scriptlet Event Handler
11.24...	{06290BDB-48...	ScriptletHandler...	C:\Windows\sy...	Constructor for Scriptlet Behavior Handler
11.24...	{0955AC62-8F...	BDATuner.MP...	C:\Windows\S...	BDA Tuning Model MPEG2 Tune Request
11.24...	{10072CEC-8C...	PeerDraw.Peer...	%CommonProgr...	PeerDraw Class
11.24...	{11556518-F20...	NvCpl.NvCplLa...	C:\Windows\sy...	NvCplLateBound Class
11.24...	{15D6504A-54...	BDATuner.DVB...	C:\Windows\S...	BDA Tuning Model DVB Tune Request
11.24...	{18E49F30-0E...	BDATuner.Lan...	C:\Windows\S...	BDA Tuning Model Language Component Type Class (Broadcast Substream Type)
11.24...	{1C15D484-91...	MSVidCtl.MSVI...	C:\Windows\S...	Legacy Analog TV Tuner Device Segment
11.24...	{1C82EAD9-50...	WindowsMail.M...	%SystemRoot%	Windows Mail Mime Editor
11.24...	{1DF7D126-40...	BDATuner.DVB...	C:\Windows\S...	BDA Tuning Model DVB Satellite Locator
11.24...	{1EFB6596-857...	MSComctlLib.T...	C:\Windows\sy...	Microsoft TabStrip Control, version 6.0
11.24...	{20DD1B9E-87...	MSComctl2.DT...	C:\Windows\sy...	Microsoft Date and Time Picker Control 6.0 (SP4)
11.24...	{22A85CE1-F0...	CAPICOM.Utiliti...	C:\Program File...	Utilities Class
11.24...	{232E456A-87...	MSComctl2.Mo...	C:\Windows\sy...	Microsoft MonthView Control 6.0 (SP4)
11.24...	{24DC3975-09...	MSVidCtl.MSVI...	C:\Windows\S...	Video Mixing Renderer 9 Device Segment
11.24...	{26EC0B63-AA...	BDATuner.Digit...	C:\Windows\S...	BDA Tuning Model Digital Cable Tune Request
11.24...	{2C247F23-859...	MSComctlLib.I...	C:\Windows\sy...	Microsoft ImageList Control, version 6.0
11.24...	{2C63E4EB-4C...	BDATuner.MP...	C:\Windows\S...	BDA Tuning Model MPEG2 Tune Request Factory
11.24...	{21411122-58...	MSVidCtl.MSVI...	C:\Windows\S...	Legacy Analog TV Tuner Device Segment

Search Tools

Search: ☒ Show Active ☐ File Name ☐ Guid ☐ Date ☐ Contains

☐ Show Hidden ☐ Description ☐ ProgID ☐ Highlighted

Date Added: 11.24.08 Selected File: C:\Windows\System32\msvidctl.dll  

ProgID: MSVidCtl.MSVidXDS.1 GUID: {0149EEDF-D08F-4142-8D73-D23903D21E90}

Description: XDS Feature Segment

Audit Notes & Safety Report

Kill Bit is Set

RegKey Safe for Script: True  
RegKey Safe for Init: False



# Step by step (cont.)

## Determining relationships

### Documentation

Loaded modules

Do different processes share 3rd party DLL files?

Local ports, Named pipes

IPC

Check shared handles

Wireshark

Use the product, sniff

# Step by step (cont.)

## Determining trust

### Documentation

Ethernet interfaces bound

Local? Remote? TCP? UDP? ...

Named pipe restrictions

Authentication

ActiveX

Safe for scripting

Safe for init

Privileges of users running processes

Permissions on resources, directories, handles, ...

# Step by step (cont.)

## Locating inputs

- Documentation

- Registry entries

  - File formats, codecs, protocol handlers

- TCP View /Process Explorer

## Probing inputs

- Create your own “clients”

  - MSRPC

    - Impacket (Core), PyMSRPC (Myself and Cody Pierce)

  - ActiveX

    - COMRaider (David Zimmer), Axman (HD Moore)

  - TCP/UDP/...

    - Socket code (py, pl, C, take your pick)

- Subverting client code

  - Don't bother implementing an encryption if you can steal theirs

Questions?

# Disassembling and IDA Pro

# Intro to binary code structure

## Modules

process.exe, library.dll

## Functions

At least one basic block, can be called

## Basic Blocks

Groups of instructions terminated at a branch or return

```
mov ebx, dword_0x400400  
test ebx, ebx  
jz fail
```

## Instructions

Atomic

```
mov eax, [ebp+0x4c]
```

# Func/Basic block/Instruction

```

; Attributes: bp-based frame

; void __stdcall MIDL_user_free(void *)
__stdcall MIDL_user_free(x) proc near ; CODE XREF: SsRecreateStickyShares()+59↓p ...

arg_0= dword ptr 8

; FUNCTION CHUNK AT 4B3AE4C1 SIZE 00000004 BYTES

mov     edi, edi
push    ebp
mov     ebp, esp
cmp     [ebp+arg_0], 0
jz      loc_4B3AE4C1
```

```

; NLL
pop     ebp
jmp     ds:LocalFree(x)
__stdcall MIDL_user_free(x) endp
```

```

; NLL
; START OF FUNCTION CHUNK FOR _MIDL_user_free@4

loc_4B3AE4C1:                ; CODE XREF: MIDL_user_free(x)+9↑j
pop     ebp
retn    4
; END OF FUNCTION CHUNK FOR _MIDL_user_free@4
```

# Func/Basic block/Instruction

```

; Attributes: bp-based frame

; void __stdcall MIDL_user_free(void *)
__stdcall MIDL_user_free(x) proc near ; CODE XREF: SsRecreateStickyShares()+59↓p ...

arg_0= dword ptr 8

; FUNCTION CHUNK AT 4B3AE4C1 SIZE 00000004 BYTES

mov     edi, edi
push    ebp
mov     ebp, esp
cmp     [ebp+arg_0], 0
jz      loc_4B3AE4C1
```

**FUNCTION**

```

pop     ebp
jmp     ds:LocalFree(x)
__stdcall MIDL_user_free(x) endp
```

```

; START OF FUNCTION CHUNK FOR _MIDL_user_free@4

loc_4B3AE4C1: ; CODE XREF: MIDL_user_free(x)+9↑j
pop     ebp
retn    4
; END OF FUNCTION CHUNK FOR _MIDL_user_free@4
```



# Func/Basic block/Instruction

```

; Attributes: bp-based frame

; void __stdcall MIDL_user_free(void *)
__stdcall MIDL_user_free(x) proc near, CODE_XREF: $S$recreateStickyShares()+59↓p ...

arg_0= dword ptr 8

; FUNCTION CHUNK AT 4B3AE4C1 SIZE 00000004 BYTES

mov     edi, edi
push    ebp
mov     ebp, esp
cmp     [ebp+arg_0], 0
jz      loc_4B3AE4C1

```

**BASIC BLOCK**

```

pop     ebp
jmp     ds:local_free(x)
__stdcall MIDL_user_free(x) endp

```

**BASIC BLOCK**

```

; START OF FUNCTION CHUNK FOR _MIDL_user_free@4

loc_4B3AE4C1: ; CODE_XREF: MIDL_user_free(x)+9↑j
pop     ebp
retn    4
; END OF FUNCTION CHUNK FOR _MIDL_user_free@4

```

**BASIC BLOCK**

# Func/Basic block/Instruction

```

; Attributes: bp-based frame

; void __stdcall MIDL_user_free(void *)
__stdcall MIDL_user_free(x) proc near ; CODE XREF: SsRecreateStickyShares()+59↓p ...

arg_0= dword ptr 8

; FUNCTION CHUNK AT 4B3AE4C1 SIZE 00000004 BYTES

mov     edi, edi
push    ebp
mov     ebp, esp
cmp     [ebp+arg_0], 0
jz      loc_4B3AE4C1
```

## INSTRUCTIONS

```

pop     ebp
jmp     ds:LocalFree(x)
__stdcall MIDL_user_free(x) endp
```

```

; START OF FUNCTION CHUNK FOR _MIDL_user_free@4

loc_4B3AE4C1:                ; CODE XREF: MIDL_user_free(x)+9↑j
pop     ebp
retn    4

; END OF FUNCTION CHUNK FOR _MIDL_user_free@4
```

# Graphing

**Code can be represented as a graph, as shown previously**

Graph traversal code is applicable here

Assuming no dynamic transfers of execution, like:

*call [edx+0x20]*

## **Graphing tricks**

Reachability (Function & Basic Block)

Upgraph/Downgraph/Intersection

Discover new vectors for attacks

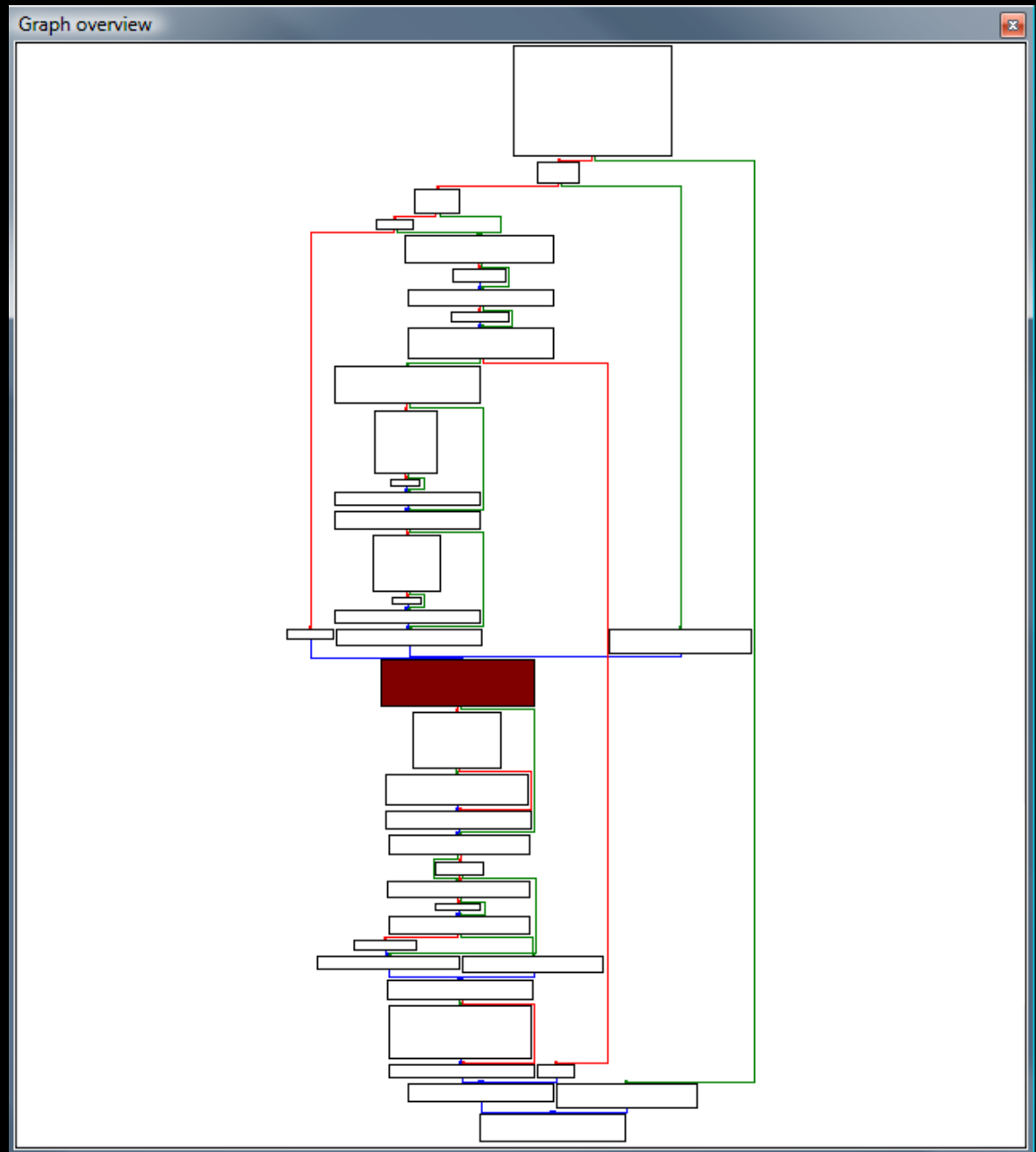
Discover paths to interesting code

Locate recursive functions programmatically

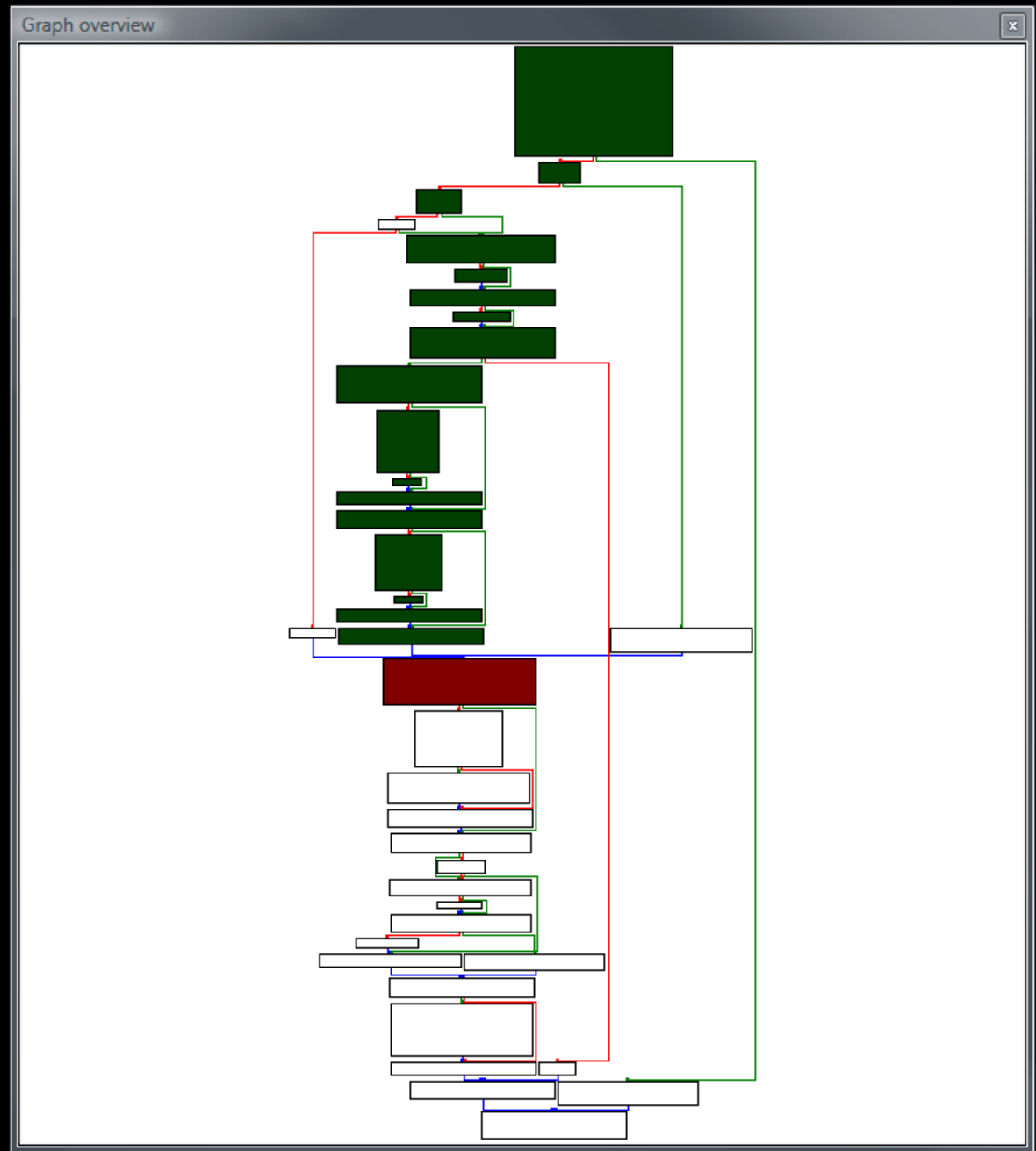
Loop detection

Binary diffing (BinDiff from Zynamics)

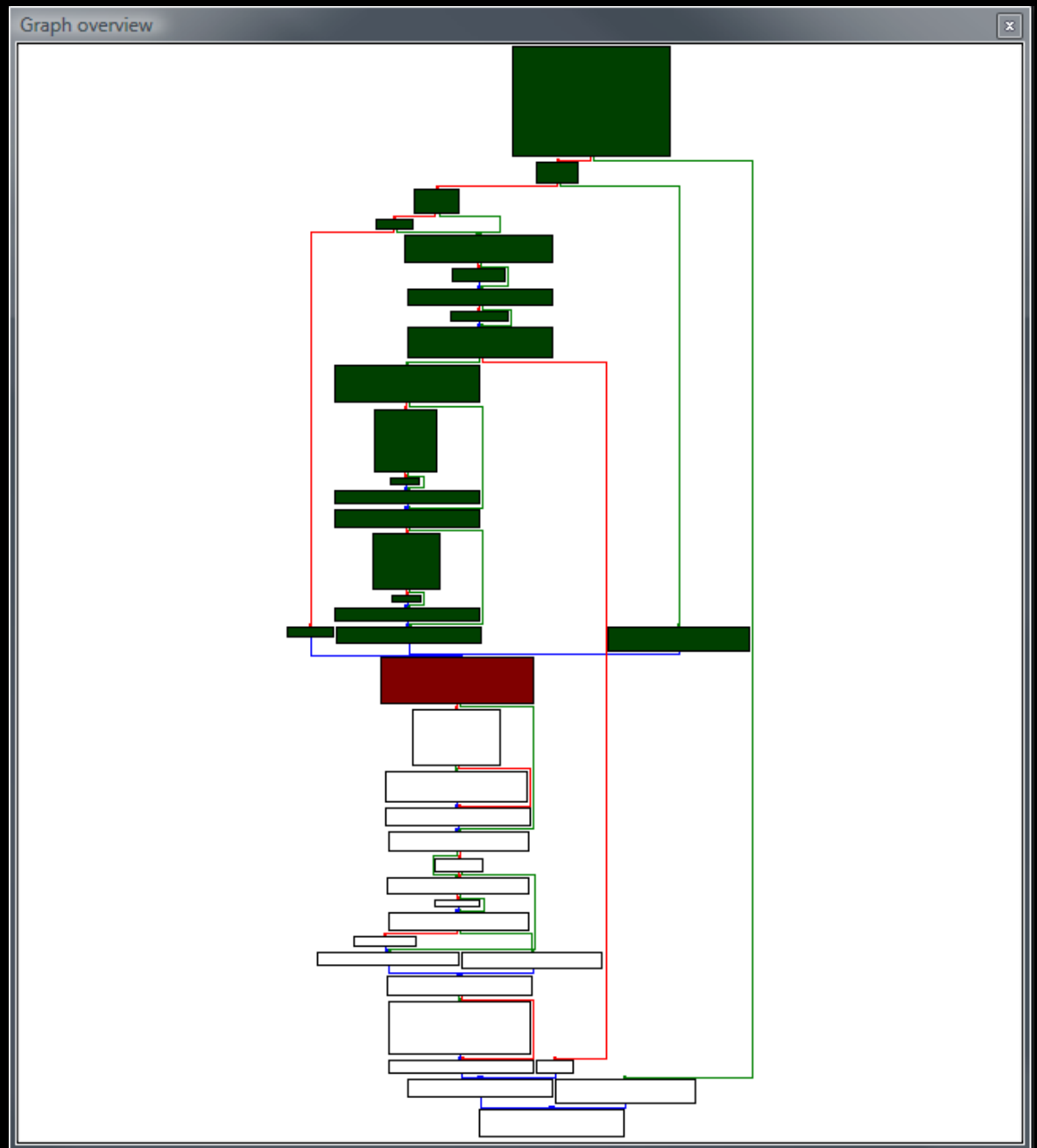
# Graphing



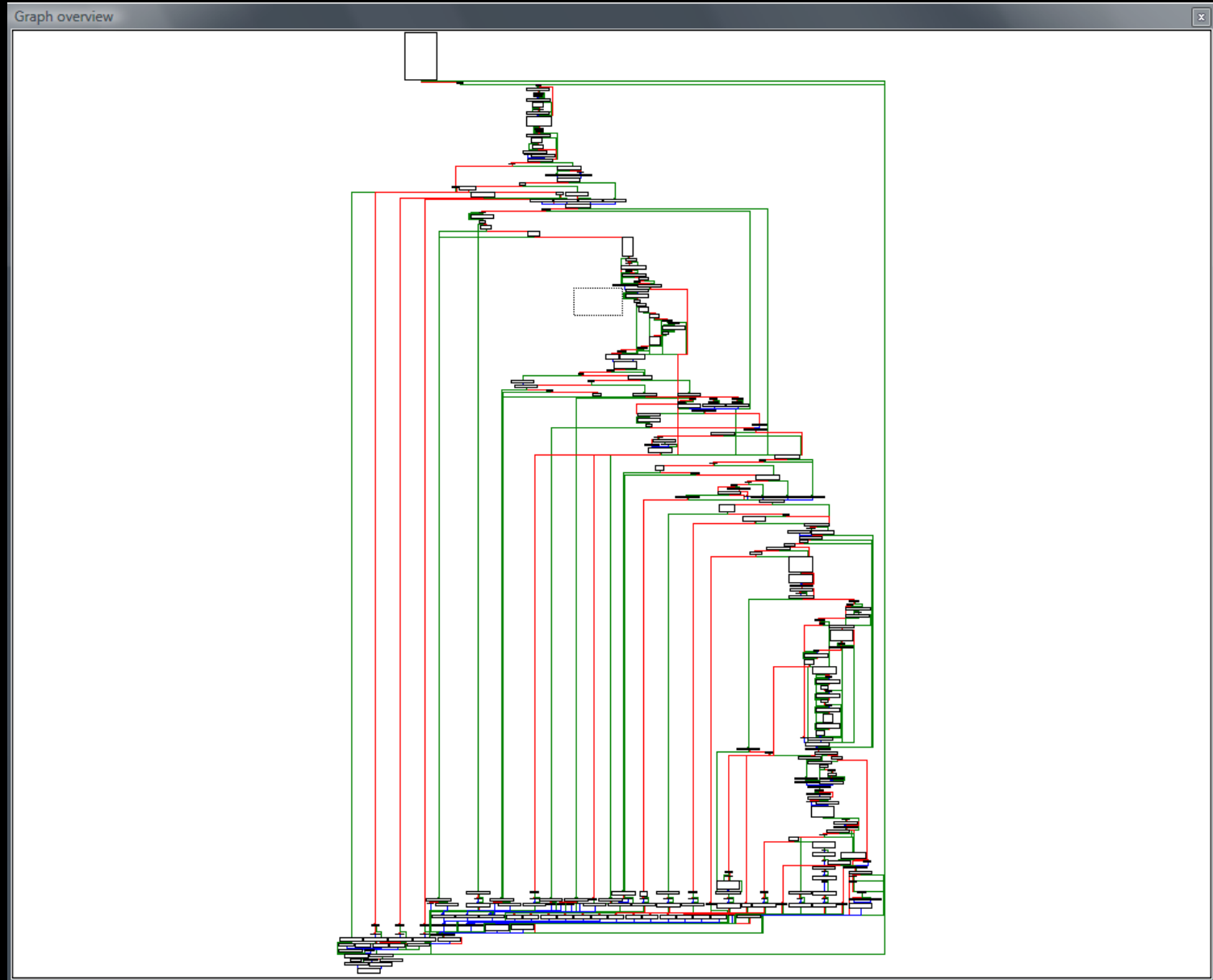
# Graphing



# Graphing



# Graphing



Questions?



# Intro to binary data structures

## Objects (think object-oriented, C++, ...)

- @ecx

- Constructors

- Destructors

- Function tables

  - Methods

- Inheritance

## Variables

- Local

- Global

- Structures

- Defined on Stack vs. Heap

  - Important for exploitation

# Introduction to IDA Pro

**How many here have used a disassembler? IDA?**

**Important facilities to a reverser provided by IDA**

- FLIRT

- Strings

  - assert() calls

  - debug functions

- Cross referencing

- Imports/Exports

- Segments

**IDA SDK, IDC, IDAPython, IDA Debugger**

- Plugins

- Automated analysis (we'll get to this later)

# RE – Static Analysis

## **Important to locate sources of user input**

No runtime info available (besides sometimes RTTI)

## **Cross referencing and graphing is key**

C++ can make this aggravating

## **Pattern matching is helpful**

IDC/IDAPython

Find me all “movsx” from this function down

Find me all “add reg32, x” followed by malloc()

Loop detection

Unsafe library calls

\*cpy

\*alloc

# RE – Dynamic Analysis

**Breakpoints allow for jump start on analysis**

e.g. Memory breakpoint on recv() buffer

**Ability to resolve...**

Object structure and relationships

Type information

Input from other processes/systems/configs/...

Global variables

**Ability to populate .idb with runtime information**

**Crucial to exploit development**

e.g. Analyze heap layout dynamically

Questions?

# RE – Debuggers

	Pros	Cons
<b>WinDBG</b>	Mature piece of software. Great symbol support. Allows for neat tricks like heap walking and integrity checks. Kernel!	Steep learning curve. Poor plugin API.
<b>OllyDBG</b>	Intuitive user interface. Large community of users. Nice plugin API.	Flakey symbol support. Only supports 32-bit. <b>Default install exposes exploitable vulnerabilities!</b>
<b>PyDBG</b>	Scriptable and easily extensible.	Python is slow. Only supports 32-bit. Designed to be event-based.
<b>IDA debugger</b>	Contains 9 debugging engines. Built-in to IDA.	Multiple module support can be tricky to get the hang of. UI sketchy.

# To Recap

## **Reverse from the top down**

Understand the system to understand it's parts

## **Use the proper tools to aid you**

Saves time and focus

## **Use every technique you have available**

Mixture of static and dynamic analysis

# Conclusion of Session One

**Questions?**

**E-mail the mailing list if you have additional questions**

I am subscribed as well

**Alternatively**

My gmail username is aportnoy

**Thanks!**