



Tasty Malware Analysis with T.A.C.O.

Bringing Cuckoo Reports into IDA Pro

Ruxcon 2015
Jason Jones

Who Am I?

- Sr. Security Research Analyst for Arbor Networks' ASERT
- Attend AHA! in Austin semi-frequently
 - Welcome to the track!
- Speaker at
 - BlackHatUSA / Botconf / AusCERT / REcon
- Research interests
 - RE automation
 - Malware clustering
 - Graph database applications to Reverse Engineering / Threat Intel



Agenda

- Similar Work
- Malware Behaviors
- Cuckoo Sandbox
- TACO
 - Features
 - -UI
 - Demo
 - Future Work







Similar Work

Similar Work

- Nothing (that I know of) uses Cuckoo as it's mechanism for propagating data into an IDB
- Inspired by similar work from many authors
- UI takes inspiration from IDAScope by Daniel Plohmann (@push_pnx)
 - Excellent plugin, in my toolbox



funcap

- https://github.com/deresz/funcap
- IDA Pro script to add some useful runtime info to static analysis.

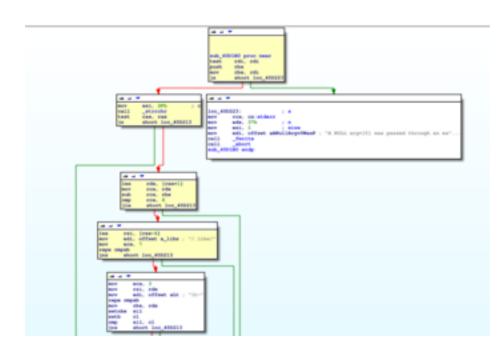
```
eax, [ebp+NewFileName]
1ea
                        ; dwFlags
push
                        ; lpNewFileName
push
        eax
                        ; lpExistingFileName
push
        esi
   arg 00: 0x00404314 --> 'C:\Documents and Settings\Administrator\Local Settings\RDSessMgr'
   arg 04: 0x0012fd8c --> 'C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\~da29.tmp'
   arg 08: 0x00000001 --> 'N/A'
        ds:MoveFileExA ; kernel32 MoveFileExA()
   EAX: 0x00000001 --> 'N/A'
   s arg 00: 0x00404314 --> 'C:\Documents and Settings\Administrator\Local Settings\RDSessMgr'
   s arg 04: 0x0012fd8c --> 'C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\~da29.tmp'
   s arg 08: 0x00000001 --> 'N/A'

    hEsill#Evicte
```

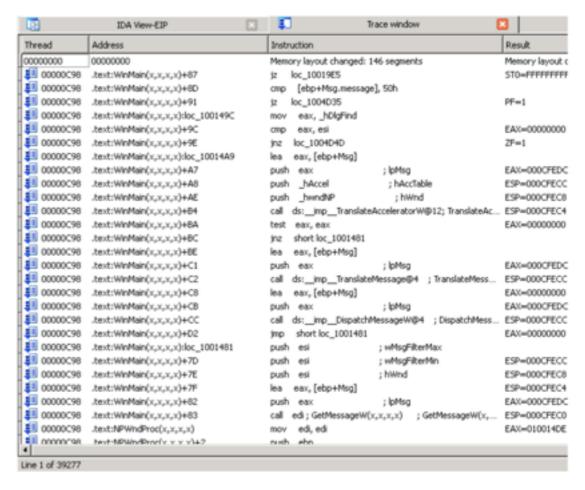


IDA Pro pintracer

- Maintained by Hex-Rays
- Highlights executed instructions
- Can also track registers







Joe Sandbox

- Commercial product from Joe Security
- Can produce execution graphs
- Claims to have similar plugin
 - Never used personally
 - Seeing that they were using API traces gave inspiration to look into doing similar with Cuckoo
 - Opted to not attempt to find code so my plugin would be "clean"









Malware Analysis Challenges

Packers / Crypters

- Compress or encrypt code, designed to make malware less detectable
- UPX most popular packer (also watch out for things that look like, but are not UPX)
- Lots of packers with various trial licenses
- TitaniumCore by ReversingLabs can help automate
- No known (to me) auto un-crypters
- PIN, Dynamo Rio have tools to facilitate
- IDA Pro as a "universal unpacker" that has been useful at times



Self Modifying Code

- Exhibited by numerous malware families
 - Shylock
 - Andromeda / Gamarue
- Modify code that already exists instead of allocating new memory to unpack
 - Usually will be stomped during execution
 - More problematic to do automated dumps



Process / DLL Injection

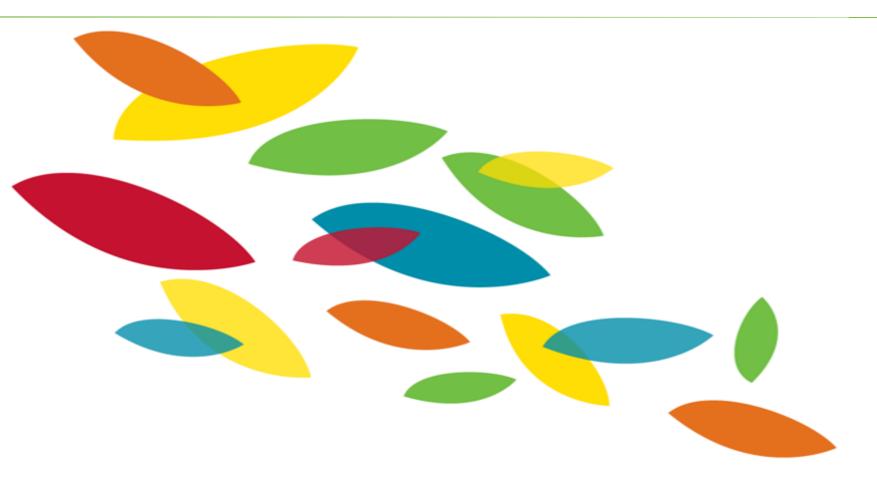
- Can be done via
 - CreateRemoteThread (Suspended)
 - QueueUserAPC
 - Process Hollowing
- Cuckoo uses injection to get monitor DLL into malicious processes



DLL Side Loading

- Popular technique with targeted malware
 - PlugX
 - HTTP Browser RAT
- Load malicious DLL into legit (signed) executable
 - Bypass (some) AV
 - Bypass requirements of running code in signed exe







Cuckoo Sandbox

Cuckoo Sandbox

- Likely most popular open-source / free sandbox available
- 2.0 Supports Android (via emulator), Linux, and x64 analysis
 - Switch to new monitor code
- Third-party kernel introspection support "zer0m0n"
- Popular fork "cuckoo-modified" by @spender of Optiv, Inc. (Accuvant)
 - https://github.com/brad-accuvant/cuckoo-modified
 - Contains bugfixes + additions to old cuckoomon not available in trunk
 - Cuckoo 2.0 solves many of the issues we relied on -modified fork for and adds new things

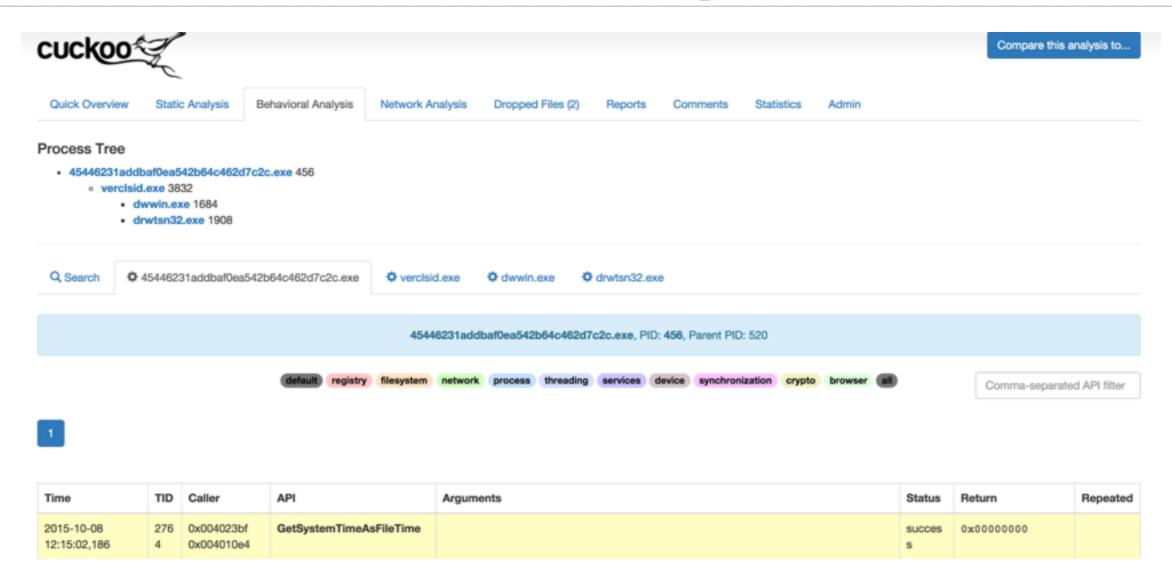


Cuckoo Sandbox

- Multiple analysis methods
- Cuckoo Monitor DLL injected into spawned process
 - Injects into any other spawned / injected processes
 - Hooks many common API calls
 - Nothing is immune to un-hooking, including Monitor
- Logs
 - Win32 API calls
 - Registry
 - Created / Modified Files
- Postprocessing Signatures



Cuckoo Behavior Report





Cuckoo Behavior - Calls

2015-10-08 12:15:07,015	276 4	0x02135130 0x02135936	Caller /	ProcessHandle: 0xfffffff	succes	0x0000000
				BaseAddress: 0x71ab1000		
2015-10-08 12:15:07,015	276 4	0x02135130 0x02135936	NtProtectVirtualMemory	OldAccessProtection: PAGE_READWRITE NumberOfBytesProtected: 0x00001000 NewAccessProtection: PAGE_EXECUTE_READ ProcessHandle: 0xffffffff BaseAddress: 0x71ab1000	succes	0x0000000
2015-10-08 12:15:07,015	276 4	0x02135130 0x02135936	NtOpenSection	DesiredAccess: 0x0000000e ObjectAttributes: WS2HELP.dll SectionHandle: 0x00000000	failed	OBJECT_NAME_NOT _FOUND
2015-10-08 12:15:07,015	276 4	0x02135130 0x02135936	NtQueryAttributesFile	FileName: C:\Documents and Settings\Administrator\Local Settings\Temp\WS2HELP.dll	failed	OBJECT_NAME_NOT _FOUND
2015-10-08 12:15:07,015	276 4	0x02135130 0x02135936	NtQueryAttributesFile	FileName: C:\WINDOWS\system32\ws2help.dll	succes	0x00000000
2015-10-08 12:15:07,015	276 4	0x02135130 0x02135936	NtOpenFile	ShareAccess: FILE_SHARE_READ FILE_SHARE_DELETE FileName: C:\WINDOWS\system32\ws2help.dll DesiredAccess: FILE_EXECUTE SYNCHRONIZE FileHandle: 0x00000094	succes s	0x0000000
2015-10-08 12:15:07,015	276 4	0x02135130 0x02135936	NtCreateSection	ObjectAttributes: DesiredAccess: SECTION_QUERY SECTION_MAP_READ SECTION_MAP_WRITE SECTION_MAP_EXECUT E SectionHandle: 0x00000098	succes s	0x0000000



Cuckoo Behavior JSON - Modified

```
"category": "network",
"parentcaller": "0x02d9bd76",
"return": "0x00cc00008",
"timestamp": "2015-09-29 00:18:04,734",
"caller": "0x02d9f5fd",
"thread_id": "1532",
"repeated": 0,
"api": "InternetConnectA",
"status": true,
"arguments": [
        "name": "Username",
        "value": ""
        "name": "Service",
       "value": "3"
        "name": "InternetHandle",
        "value": "0x00cc0004"
        "name": "ServerName",
       "value": "macsystem.jp.net"
```



Cuckoo Behavior JSON -2.0

```
"category": "network",
    "status": 1,
    "stacktrace": [
        "InternetSetOptionW+0x68 InternetCreateUrlA-0x68a6 wininet+0xbca9 @ 0x771bbca9",
        "8982410d05e1839148299ca96af9e4c8+0xb7f2 @ 0x100b7f2"
    "api": "InternetSetOptionA",
    "return_value": 1,
    "arguments": {
        "option": 31,
        "internet_handle": "0x00cc000c"
    "time": 1444464500.5,
    "tid": 1420,
    "flags": {
        "option": "INTERNET_OPTION_SECURITY_FLAGS"
Ъ,
    "category": "system",
    "status": 1,
    "stacktrace": [
        "GetProcAddress+0x3e IsProcessorFeaturePresent-0x4c kernel3Z+0xae6e ⊕ 0x7c80ae6e",
        "8982410d05e1839148299ca96af9e4c8+0x100b7 @ 0x10100b7"
    "api": "LdrGetProcedureAddress",
    "return_value": 0,
    "arguments": {
        "ordinal": 0,
        "module_address": "0x771b0000",
        "function_address": "0x77212ebc",
        "function_name": "HttpSendRequestW"
    "time": 1444464500.5,
    "tid": 1420.
    "flags": {}
```



ASERT's Sandbox Usage

- Treat Cuckoo (and other sandboxes) as a black-box
 - Malware in, report / memory dumps / files out
 - Tasks deleted upon completion
- Centralized malware processing system
 - Normalize + insert results
 - Post-processing of memory, network traffic, behavior
 - Custom post-processing of specific families to extract various sample properties



Cuckoo API Additions needed

- Cuckoo can produce a process dump
 - This is not loadable by IDA Pro (AFAIK)
 - Can be extremely large, especially in case of {explorer,svchost,iexplore,etc.}.exe
- Can also produce full RAM dump
- Volatility has plugins to dump processes, DLLs, VADs
 - Dumping process as a PE not supported natively by Cuckoo
 - Due to time needed to use volatility, decided that was not the right place
 - Don't always want dumps, sometimes we need to do "extra"
- Added new API call to allow for arbitrary volatility plugins to run "on-demand"



API Additions needed (cont)

- Run volatility against ramdump to get process dumps for all PIDs known
- Injection detected = run malfind and dump pages
 - Stitch dumped memory pages into process dumps for "complete" view
- Supports family specific behavior
 - DLL dump
 - Specific process / memdumps



Dumping Memory

- That said... malfind doesn't always find everything
 - Will not dump DLL injected with CreateRemoteThread by design
 - Permissions stomp = undetected
 - Walk the Cuckoo API Calls per process
 - Get list of memory ranges that contain executed code
 - Run vadwalk for the PID
 - Parse the output and find all the required VAD's to cover what got executed
 - Request those VADs and then order with malfind VAD's and stitch an executable together
- Using that dump, can now follow execution much better



Creating the Memory Dump

- Attempted to add as sections using http://git.n0p.cc/?p=SectionDoubleP.git
 - Works great for any case where section is above ImageBase
 - BUT many malwares like to inject below the ImageBase
 - Modify ImageBase
 - Modify each existing section's VirtualAddress
 - Modify AddressOfEntryPoint
 - Add Sections...
 - Fail.
 - Fallback to using IDA Pro segment create / put_many_bytes
 - Non-ideal, but IDA plugin requires IDA Pro...
- Non-trivial method of creating dumps, but worth it



Memory Dump Process Output

- python create_voldump.py --task 294832 --pid 3816
- [+] Base memory range: 01000000 -> 01005600
- [+] Interesting page: 0x000C0000
- [+] Interesting page: 0x00B40000
- [+] Interesting page: 0x00B50000
- [+] Interesting page: 0x00B60000
- [+] Interesting page 0x000C0000 is in VAD 0x000C0000 0x000DCFFF
- [+] Interesting page 0x00B40000 is in VAD 0x00B40000 0x00B70FFF
- [+] Interesting page 0x00B50000 is in VAD 0x00B40000 0x00B70FFF
- [+] Interesting page 0x00B60000 is in VAD 0x00B40000 0x00B70FFF
- [+] Retrieving VAD 0x000C0000
- [+] Retrieving VAD 0x00B40000
- [+] Generating IDB with new memory regions
- [+] IDB available at explorer.exe-3816.idb









TACO

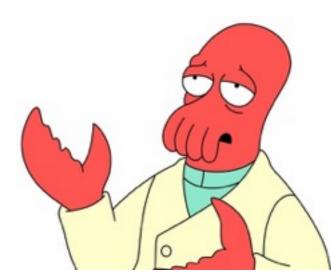
Overview

- Started out as dynamically generated Python scripts
 - Clunky, prevented from doing "cool" things
 - Dynamically generating "clean" IDAPython is hard
 - Some features incompatible with Cuckoo 1.2 due to lack of call metadata
 - Cuckoo-Modified and current Cuckoo 2.0-dev branch supported supported for markup
 - Cuckoo 2.0-dev is still a WIP as some oddities are encountered
- Idea sprung out of Joe Security's posts about execution graphs and seeing they imported analysis info into IDA
- Prior usage of tools like funcap and IDA's pintracer

TACO Overview

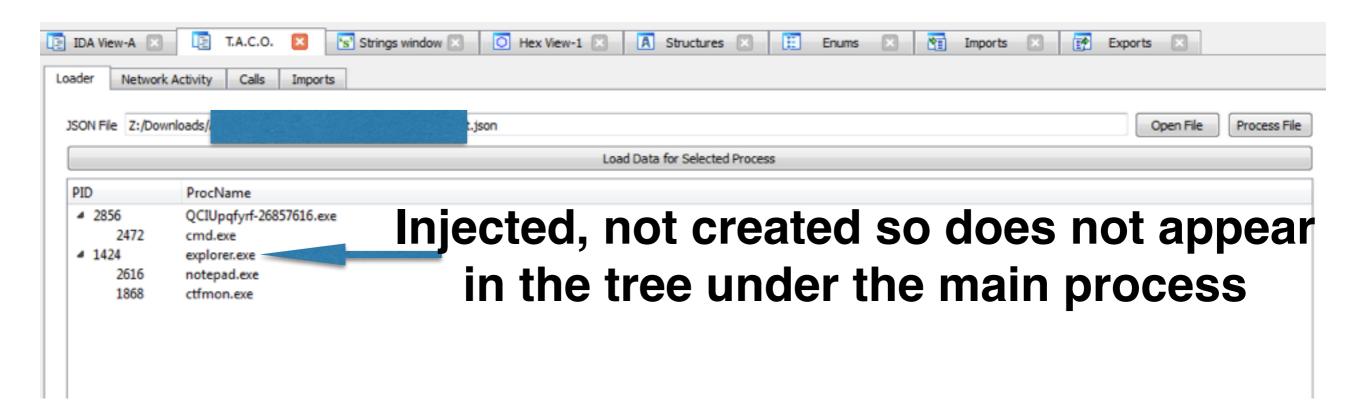
- What does TACO stand for?
 - It's fluid...
 - Considered naming TACOZ Tasty Analysis using Cuckoo Output and Zoidberg
 - Because why not Zoidberg?
- Consists of Cuckoo-based tabs for showing:
 - Processes
 - API Calls
 - Signatures
 - Imports
- Also includes other IDAPython scripts I have developed
 - Byte / Stack String viewer
 - "Interesting" XOR locator
 - Switch Jump / Case statement viewer





Loader Tab

 Main location to show a process tree and allow for specific processes to be inspected





API Call Tab

Reproduction of Cuckoo's Output

Filterable / Searchable / Clickable

Filterable by Category Cuckoo Loader Cuckoo Signatures Cuckoo Calls Cuckoo Imports Byte Strings Interesting XOR Switch Jumps Choose Color: Choose Function Color: Categories: V Browser V Crypto V Device V Filesystem V Misc V Network V Process V Registry V Services V Synch V System V Threading Category Parent Caller API Filterable by Call / Argument value 0x004039cb 0x004017cf InternetCloseHandle ShareAccess: 3 77 synchronization 0x00402c40 0x00401352 NtCreateNamedPipeFile DesiredAccess: 0x80100100NamedPipeHandle: 0x00000164 Buffer: jairocpejhkdol> FileHandle: 0x00000160 0x000000000 78 filesystem 0x004028ce 0x00401790 Length: 15 HandleName: \\Device\\NamedPipe\\Win32Pipes.00000b28.00000001 ApplicationName: 79 process CreateProcessInternalW 0x00000001 Processld: 3596 CommandLine: regedit /s C:\\DOCUME~1\\ADMINI~1\\LOCALS~1\\Temp\\kb71271.logThrea... Each row Color-coded and double-clickable



API Call Tab (cont.)

- Add / Remove Markup to IDB
 - All
 - Category
- Context menu
 - Markup per Instruction
 - Copy value

```
10 100 4007 10.
                                             , CODE AREF. SUD_4004C0TZOCIJ
                    push
                            400000h
                    push
                    push
                    push
                            ecx, [ebp+Dest]
                    lea
                    push
                    push
                    push
                            offset aGet
                                             ; "GET"
                    push
                            [esi+HttpOpenRequestA]
   api: HttpOpenRequestA
'32 Path: /AWS96.jsp?2gZ6Sg1/SaQ72a3xyBmMjnxfI/RBYO9Lhf9fj58MI5DmIZ3jnnRfjnhT
'32 InternetHandle: 0x00cc0008
```



Imports Tab

Tries to detect dynamic imports via direct / indirect calls

ucko	oo Loader Cuc	koo Network Activity	Cuckoo Calls	Cuckoo Imports	Byte Strings	Interesting XOR	Switch Jumps		
	Address	DLL		ProcName	ProcAddres	s Type			
1	0x00401E64	WS2_32.dll	WSASt	artup	0x71ab6a55	Dynamic			
2	0x00401EA6	WS2_32.dll	gethos	tname	0x71ab5449	Dynamic			
3	0x00401EE6	WS2_32.dll	gethos	tbyname	0x71ab5355	Dynamic	.Lext:00404/t		[eup+var_84], r
4	0x00401F16	WS2_32.dll	inet_nt	toa	0x71ab45c1	Dynamic	.text:0040470		[ebp+var_B3], 't [ebp+var_B2], 't
5	0x00401F5F	WS2_32.dll	WSAC	leanup	0x71ab3fed	Dynamic	.text:0040470		[ebp+var_B1], 'p [ebp+var_B0], 'S
6	0x00402010	mswsock.dll	NSPSta	artup	0x71a5bd98	Indirect	.text:0040470	DE nov	[ebp+var_AF], b]
7	0x00402010	mswsock.dll	NSPSta	artup	0x71a5bd98	Indirect	.text:004047E	EB nov	[ebp+var_AE], 'n [ebp+var_AD], 'd
8	0x00402010	winrnr.dll	NSPSta	artup	0x76fb1688	Indirect	.text:004047F		[ebp+var_AC], 'F [ebp+var_AB], b]
9	0x004045E2	wininet.dll	Interne	etOpenA	0x3d945828	Dynamic	.text:004047F		[ebp+var_AA], 'c [ebp+var_A9], 'c
10	0x00404674	wininet.dll	Interne	etSetOptionA	0x3d94c39a	Dynamic	.text:0040486	8D mov	[ebp+var_A8], b1
11	0x004046C3	wininet.dll	Interne	etConnectA	0x3d956f4e	Dynamic	.text:0040481		[ebp+var_A7], 's [ebp+var_A6], 't
12	0x00404712	wininet.dll	HttpO	penRequestA	0x3d9565a8	Dynamic	.text:0040482		[ebp+var_A5], 'E [ebp+var_A4], 'x
13	0x00404761	wininet.dll	HttpSe	endRequestA	0x3d947021	Dynamic	.text:0040482	2F nov	[ebp+var_A3], 'A
14	0x004047AA	wininet.dll	HttpQ	ueryInfoA	0x3d95182d	Dynamic	.text:0040483		[ebp+var_A2], <mark>0</mark> ds:GetProcAddres
15	0x0040483D	wininet.dll		endRequestExA	0x3d9baba6	Dynamic			
				'					



Cuckoo Signatures Tab

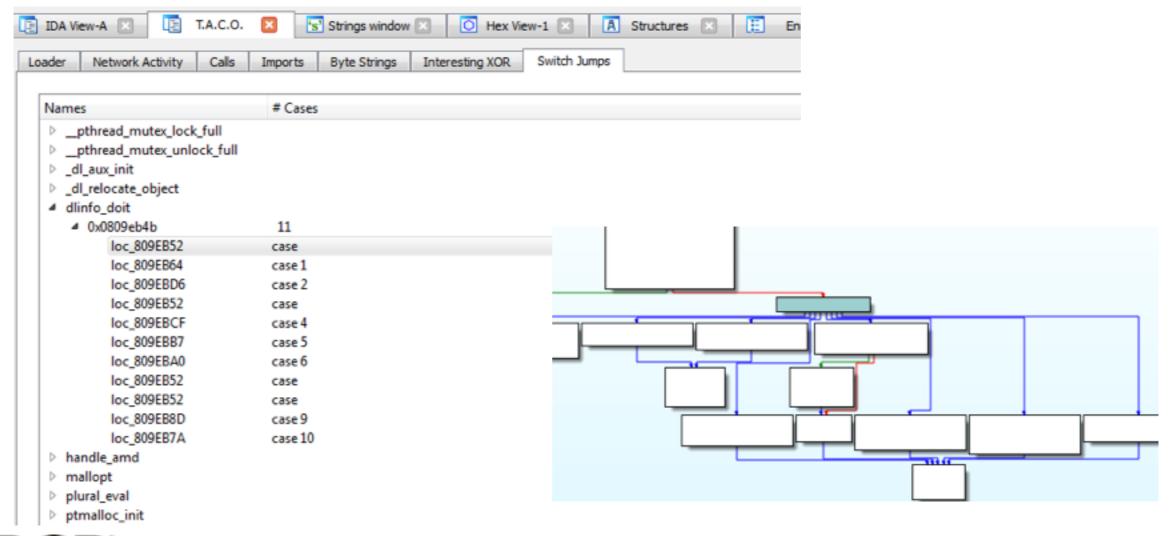
Simple Display of Cuckoo Triggered Signatures

uck	koo Loader	Cuckoo Signatures	Cuckoo Calls	Cuckoo Impor	ts Byte Strings	Interesting XOR	Switch Jumps		
		Signatu	re			Informati	on		Severity
1 Keads data out of its own hinary image					read: process: Gad 006600	Severity: 2 Confidence: 30 Weight: 1			
2	Creates a	hidden or system file		file:	C:\Documents and	Severity: 3 Confidence: 100 Weight: 1			
3	Deletes its	original binary from	disk						Severity: 3 Confidence: 100 Weight: 1
4	The binary	/ likely contains encry	/pted or compres	ssed data. IMA	GE_SCN_CNT_INIT	ntropy: 7.22, characi IALIZED_DATA IMA(ITE, raw_size: 0x000)	GE_SCN_MEM_RE		Severity: 2 Confidence: 100 Weight: 1
5	Drops a bi	nary and executes it		bina	ry: C:\Documents	and Settings\All Use	ers\AVck\Gadget.e	exe	Severity: 2 Confidence: 50 Weight: 1
6	Creates RWX memory								Severity: 2 Confidence: 50 Weight: 1
7	Generates	some ICMP traffic							Severity: 3 Confidence: 100 Weight: 1

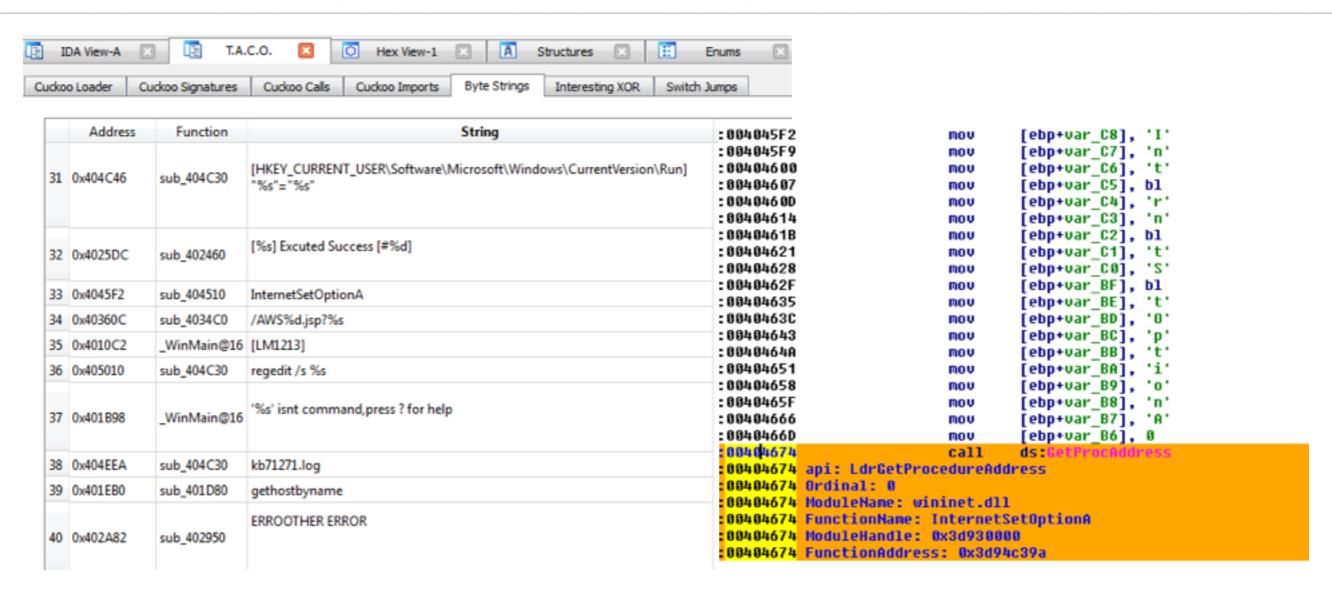


Switch Viewer

Switch jumps in malware can indicate config or cmd parsing



Byte String / Stack String Finder

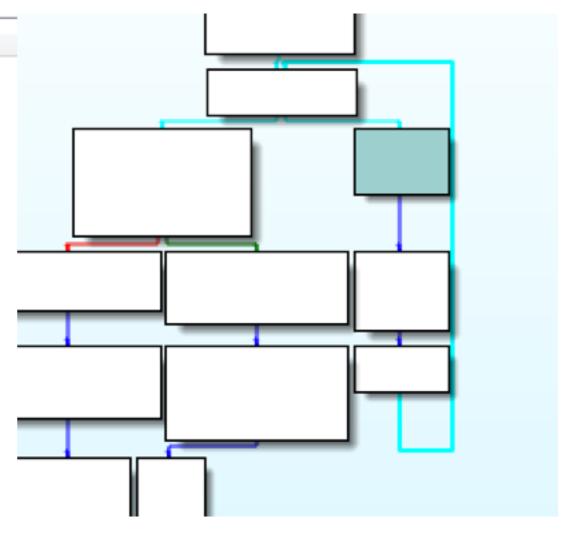




XOR Locator

Loader Ne	etwork Activity	Calls	Imports	Byte Strings	Interesting XOR	
-----------	-----------------	-------	---------	--------------	-----------------	--

	Function	Address	Loop	Disa	assembly
1	sub_401130	0x401349L	True	xor a	al, 1Ah
2	sub_401130	0x40132dL	True	xor a	al, 0CDh
3	sub_4266C4	0x4268c1L	True	xor a	al, 0CDh
4	sub_4266C4	0x4268ddL	True	xor a	al, 1Ah

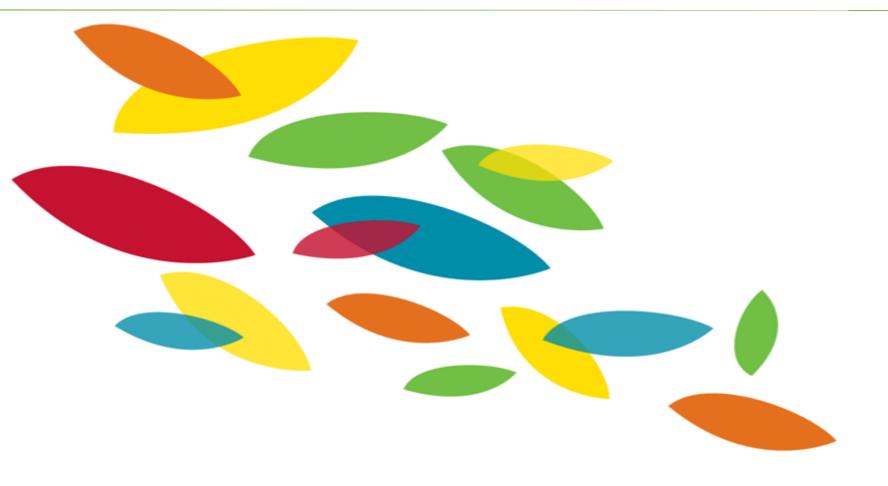




DEMO

- TACO Time!
 - Shifu (banker)
 - Andromeda (loader / stealer)
 - PlugX (targeted)
 - Etumbot (targeted)
 - Fobber (banker, Cuckoo 2.0-dev)
 - HttpBrowserRAT (targeted, Cuckoo 1.2)







Wrap-Up

Wrap-Up

- Hopefully you agree that a TACO is both a tasty treat and is a useful tool to bring run-time info into IDA Pro
- All code is / will be freely available on GitHub
 - https://github.com/arbor-jjones/idataco
 - https://github.com/arbor-jjones/cuckoo_idadump/malware/ create_voldump.py
 - https://github.com/arbor-jjones/cuckoo_idadump/malware/ ida_load_mem.py
 - API addition: https://gist.github.com/arbor-jjones/18dd572e6b3e391e8418



Future Work

- Add path-finding capabilities
- Direct comments to API call arguments with values
- Clean up filter code to allow for arg- or API call-specific filtering
- Rename vars / dwords used to store GetProcAddress result
- Rename unknown calls
- Determine way to achieve 'persistence' for names / ops (allow more 'undo')
 - SQLite?
 - Marks?
- Batch mode to markup / rename things in IDB
- Support other sandboxes where possible

Questions/Comments/Feedback





