zynamics VxClass

VxClass for Incident Response

zynamics

info@zynamics.com

Introduction

zynamics VxClass

- Binary code is often left behind by attackers
  - Running processes
  - Dropped executables
  - Kernel memory snapshots
  - Network traffic
  - Crash dumps

Introduction

zynamics VxClass

- Useful evidence but difficult to analyze
- Current methods:
  - Use AV scanner
  - Run executable to provoke/observe behavior
  - Remove packer/obfuscator code
  - Manual analysis using IDA Pro

Current Methods

zynamics VxClass

- Error-prone and time-consuming
- AV signatures are brittle, out of date
- Behavior can be difficult to provoke
- Removal of protection code is difficult
- Manual analysis
  - Does not scale
  - No easy correlation of results

VxClass

zynamics VxClass

- Structural malware classification tool
- Categorizes malware samples into families
- Groups malware that shares code
- Allows correlation between samples
  - Regardless of how they were obtained

VxClass

zynamics VxClass

- Upload of samples through a web server
- Generic unpacking through emulation
- Extraction of structural information
- Comparison with known samples
- Storage of the results in a SQL database
- Visualization of the results in the browser

Uploading

zynamics VxClass

- Upload samples
  - Through a web interface in your browser
  - Through XML-RPC
  - User-based access control to samples:
    - Public: All users can see and download
    - Limited: All users can see, but not download
    - Private: Only original uploader can see and download

Logged in as: root (logout)

Administration | System Configuration | Account Settings |

zynamics VxClass

Upload | Unpacking | Classification | Statistics

Executable Image Selection

Executable: Browse...

Description

Executable Image Options

Max Ticks (?): 1000 Million(s)

Access Rights (?): Private / Limited / Public

Keep in the database: ☑

Unpack executable (?): ☑

Command line

Filename of the dropped file

Classify executable (?): ☑

IDB / IDA Database (?): ☐

Symbian SIS package (?): ☐

Upload

Upload Item: Submit

top

Unpacking

zynamics VxClass

- Generic unpacking is difficult
  - Anti-debugging tricks
  - Attempts to foil emulators
  - Creation of and interaction between multiple processes
  - Code obfuscation

Unpacking

zynamics VxClass

- Our approach: Full system emulation
- Emulated Windows XP SP2 in Bochs
- Run the executable until it looks unpacked
- Aquire memory of all processes and dirty kernel pages
- Use code in aquired memory for classification

Unpacking

zynamics VxClass

- Solved problems
  - Anti-Debugging tricks
  - Legacy API calls
  - Multiple processes
  - Interprocess communication
  - Kernel memory analysis
- Result: Most packers can be unpacked automatically

Comparison

zynamics VxClass

- Problem: Meaningful comparison of binary code
- Byte-by-byte comparison is useless
- Our approach: Structural comparison
  - Award-winning (German IT-Security Award 2006)
  - Uses industry-standard BinDiff engine
  - Uses patent-pending MD-Index (more later)

Structural Comparison

zynamics VxClass

- Extract call graph and flow graph information from samples
- Compare the structure of these graphs instead of byte sequences
- Compares code derived from same source
  - Regardless of compiler settings
  - Regardless of compiler

zynamics VxClass

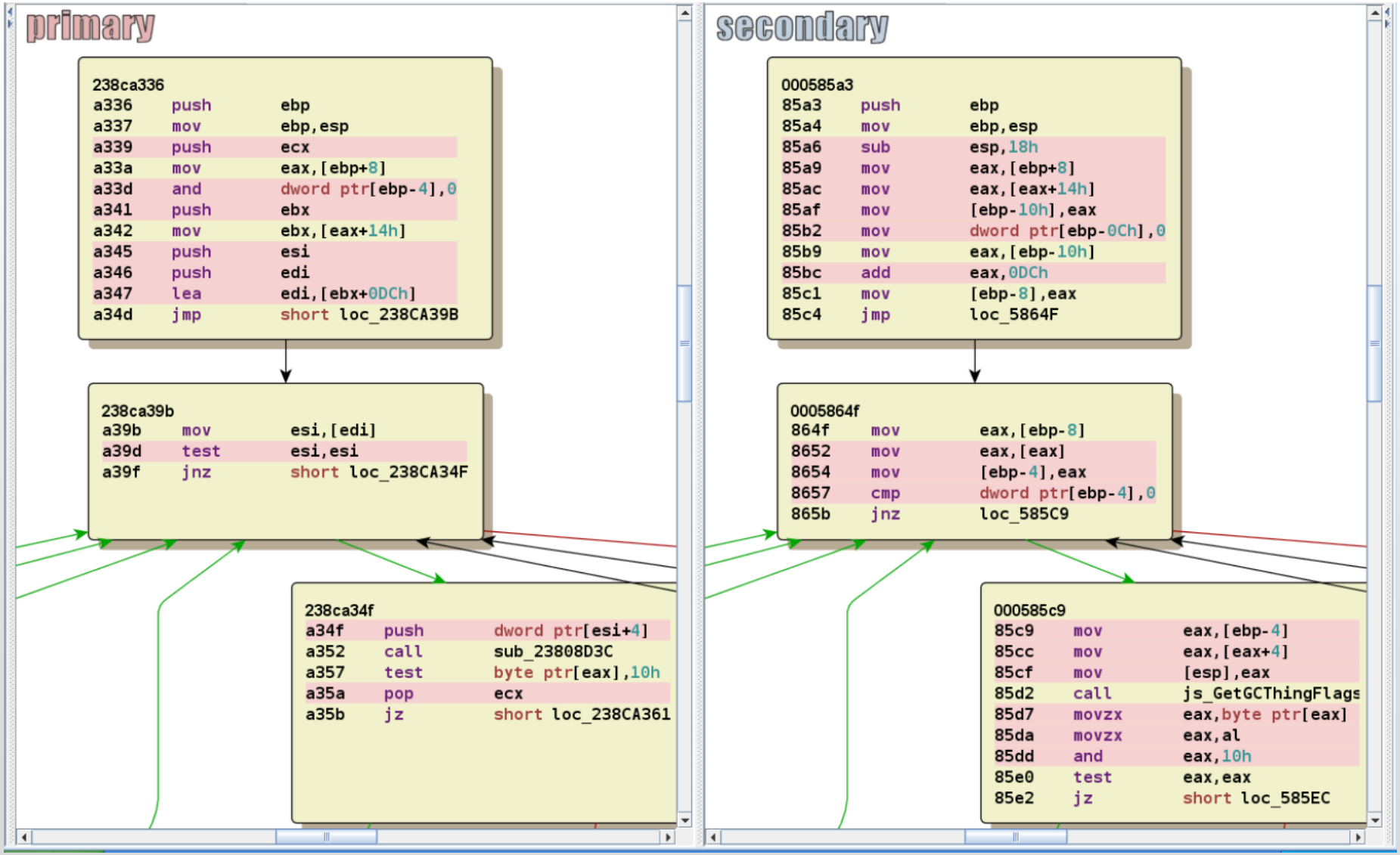
Address			Basic Block		Address
238ca336	push	ebp	push	ebp	**000585a3**
238ca337	mov	ebp,esp	mov	ebp,esp	000585a4
238ca339	push	ecx	sub	esp,18h	000585a6
238ca33a	mov	eax,[ebp+8]	mov	eax,[ebp+8]	000585a9
238ca33d	and	dword ptr[ebp-4],0	mov	eax,[eax+14h]	000585ac
238ca341	push	ebx	mov	[ebp-10h],eax	000585af
238ca342	mov	ebx,[eax+14h]	mov	dword ptr[ebp-0Ch],0	000585b2
238ca345	push	esi	mov	eax,[ebp-10h]	000585b9
238ca346	push	edi	add	eax,0DCh	000585bc
238ca347	lea	edi,[ebx+0DCh]	mov	[ebp-8],eax	000585c1
238ca34d	jmp	short loc_238CA39B	jmp	loc_5864F	000585c4
238ca34f	push	dword ptr[esi+4]	mov	eax,[ebp-4]	**000585c9**
238ca352	call	sub_23808D3C	mov	eax,[eax+4]	000585cc
238ca357	test	byte ptr[eax],10h	mov	[esp],eax	000585cf
238ca35a	pop	ecx	call	js_GetGCThingFlags	000585d2
238ca35b	jz	short loc_238CA361	movzx	eax,byte ptr[eax]	000585d7
			movzx	eax,al	000585da
			and	eax,10h	000585dd
			test	eax,eax	000585e0
			jz	short loc_585EC	000585e2
238ca35d	mov	edi,esi	mov	eax,[ebp-4]	**000585e4**
238ca35f	jmp	short loc_238CA39B	mov	[ebp-8],eax	000585e7
			jmp	short loc_5864F	000585ea
238ca361	mov	eax,[esi]	mov	eax,[ebp-4]	**000585ec**
238ca363	mov	[edi],eax	mov	edx,[eax]	000585ef
238ca365	cmp	dword ptr[esi+8],1	mov	eax,[ebp-8]	000585f1
238ca369	jnz	short loc_238CA39B	mov	[eax],edx	000585f4
			mov	eax,[ebp-4]	000585f6
			mov	eax,[eax+8]	000585f9
			cmp	eax,1	000585fc
			jnz	short loc_5864F	000585ff
238ca36b	push	dword ptr[esi+44h]	mov	eax,[ebp-4]	**00058601**
238ca36e	push	dword ptr[esi+18h]	mov	edx,[eax+44h]	00058604
238ca371	call	sub_238BC1CD	mov	eax,[ebp-4]	00058607
238ca376	test	eax,eax	mov	eax,[eax+18h]	0005860a
238ca378	pop	ecx	mov	[esp+4],edx	0005860d
238ca379	pop	ecx	mov	[esp],eax	00058611
238ca37a	jz	short loc_238CA39B	call	js_FindFinallyHandler	00058614
			test	eax,eax	00058619
			jz	short loc_5864F	0005861b
238ca37c	mov	eax,esi	mov	eax,[ebp-4]	**0005861d**
238ca37e	call	sub_238C9E35	mov	[esp],eax	00058620
238ca383	test	eax,eax	call	CanScheduleCloseHook	00058623
238ca385	jz	short loc_238CA39B	test	eax,eax	00058628
			jz	short loc_5864F	0005862a
238ca387	and	dword ptr[esi],0	mov	eax,[ebp-4]	**0005862c**
238ca38a	cmp	dword ptr[ebp-4],0	mov	dword ptr[eax],0	0005862f
238ca38e	mov	eax,[ebp+10h]	mov	edx,[ebp+10h]	00058635
238ca391	mov	[eax],esi	mov	eax,[ebp-4]	00058638
238ca393	mov	[ebp+10h],esi	mov	[edx],eax	0005863b
238ca396	jnz	short loc_238CA39B	mov	eax,[ebp-4]	0005863d
			mov	[ebp+10h],eax	00058640
			cmp	dword ptr[ebp-0Ch],0	

Structural Comparison

zynamics VxClass

Structural Comparison

zynamics VxClass



MD-Index

zynamics VxClass

- Patent-pending
- Clever hash function for directed graphs
- Assigns 80-bit value to a directed graph
  - Allows keeping a database of flow graphs
  - Allows efficient queries into the database
- Is used within VxClass for several purposes:
  - Very fast approximate comparison
  - Code search

Results

zynamics VxClass

- Memory dumps and recovered strings
- IDA files (IDB) of the resulting disassemblies
- Pairwise similarity scores
- Visualisation:
  - Family trees
  - Top-10-most-similar list

Results

zynamics VxClass

Logged in as: root (logout)

Administration | System Configuration | Account Settings |

zynamics VxClass

Upload | Unpacking | Classification | Statistics

Tags | Files | Tree

Select a field to sort by... | Sort direction: Ascending | Show 200 items per page

Enter your filter expression here (?) Refresh

Download EXE | Download Dump | Examine Dump | Download IDB | Delete selected

		Item Name	Item Description	State	PE Dump	PE State	Time Added
	Edit	vxtest_00a54be8aaf5472f7c7f6dafa...(more)	bulk-upload	Classification successful	Info	Tainted	2009-09-29 10:58:41
	Edit	vxtest_00a9e0ac14395ecdd51ce37df...(more)	bulk-upload	Classification successful	Info	Valid	2009-09-29 10:58:41
	Edit	vxtest_00aa09a0ba645755699cd0f83...(more)	bulk-upload	Classification successful	Info	Tainted	2009-09-29 10:58:41
	Edit	vxtest_00b4b09088d7018ee0d360923...(more)	bulk-upload	Classification successful	Info	Valid	2009-09-29 10:58:41
	Edit	vxtest_00b69f877659659937e2554e5...(more)	bulk-upload	Classification successful	Info	Tainted	2009-09-29 10:58:41
	Edit	vxtest_00b7325455a9df3ff49c93137...(more)	bulk-upload	Classification successful	Info	Tainted	2009-09-29 10:58:42
	Edit	vxtest_00ba8db27e3ab64763ee0d98d...(more)	bulk-upload	Classification successful	Info	Valid	2009-09-29 10:58:43
	Edit	vxtest_00bd5b42e91ced7f6f1b8c992...(more)	bulk-upload	Classification successful	Info	Valid	2009-09-29 10:58:43

Results

zynamics VxClass

Logged in as: root (logout)

Administration | System Configuration | Account Settings |

zynamics VxClass

Upload | Unpacking | Classification | Statistics

Tags | Files | Tree

Search | Regular expression | Case sensitive

Overview

Tags

Win32/Hupigon

Win32/Allaple

50%

Results

zynamics VxClass

Logged in as: root (logout)

Administration | System Configuration | Account Settings |

zynamics VxClass

Upload | Unpacking | Classification | Statistics

Tags | Files | Tree

Search | Regular expression | Case sensitive

Overview

Tags

Win32/Hupigon

Win32/Allaple

0.989394

vxtest_0a40c5ab4e2c4bce110c2fbc2572d512.exe.58

0.99022

0.983815

vxtest_0a40c5ab4e2c4bce110c2fbc2572d512.exe->'urdvxc.exe_592

0.987876

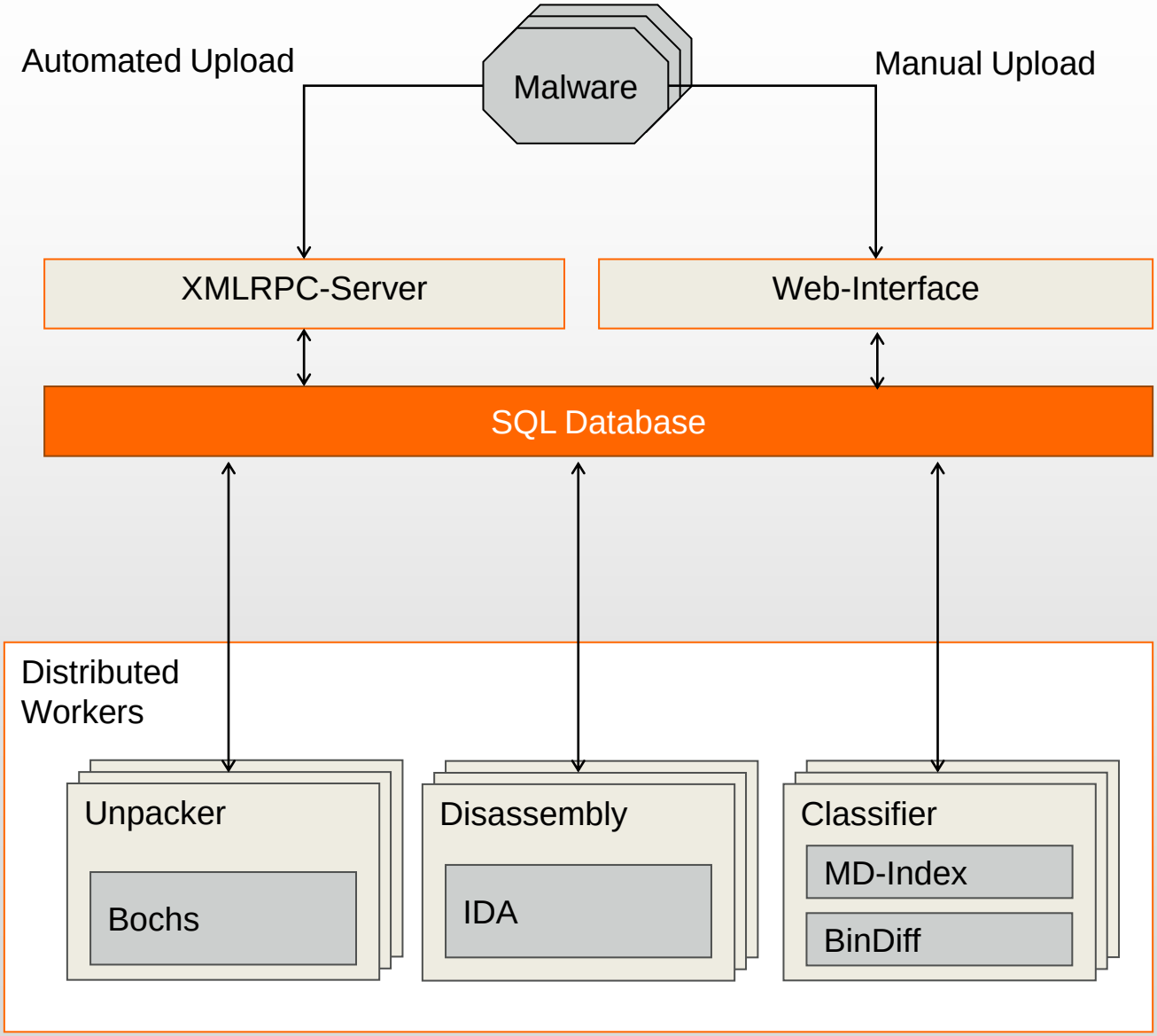
0.982545

vxtest_0a5ca2969df374f84411cec8693802ad.exe->'KERNEL_MEMORY'.167

50%

Architecture

zynamics VxClass



Case Studies

zynamics VxClass

- Noise reduction
  - Automatically filter uninteresting samples
- Knowledge management
  - Share information between analysts
- Attacker Correlation
  - Is a set of attacks performed with one toolset
- Code searching
  - Find certain functions in known samples

Noise Reduction

zynamics VxClass

- Upload new files to the system
- How similar are they to interesting samples ?
  - Comparison to database of known samples
- Prioritize accordingly

Knowledge management

zynamics VxClass

- Each analyst uploads samples he knows to VxClass
- New malware comes in, gets uploaded
- VxClass determines which known samples this is similar to
- The expert for similar samples can be found

Attacker Correlation

zynamics VxClass

- A series of incidents is investigated
- On a large number of machines, code is found
- Classify the code using VxClass to find out:
  - Is this one group of attackers ?
  - Is this similar to attacks seen in the past ?

Code Searching

zynamics VxClass

- A particularly strange piece of code (just one function) is identified
  - Perhaps a strange encryption function
- Does this particular piece of code appear in other samples in the database ?
- Search is not byte-based, but flow graph based (MD-Index)
- The answer is one click away

Performance

zynamics VxClass

- One VxClass machine
  - 800-1600 samples per day
- Performance depends on
  - Obfuscation complexity
  - Size of the malware
  - Size of the database
- Can be fully parallelized
  - The only bottleneck is the central database

Behavioral Analysis

zynamics VxClass

- VxClass is **not** a behavioral-analysis tool
- VxClass is **complementary** to such tools
- We recommend combining VxClass with behavior-monitoring tools such as
  - CWSandbox (http://www.cwsandbox.org)
  - Anubis (Free) (http://anubis.iseclabs.org)

VxClass Options

zynamics VxClass

- VxClass on a single machine
  - Run it inside your organisation
- VxClass distributed
  - Scale it to your needs
- VxClass as service
  - We host a machine for you
- VxClass as shared service
  - We host a machine for you
  - Multiple clients use a shared database

Existing Customers

zynamics VxClass

- The German BSI
  - Agency for security in information systems
- Vodafone Germany
  - Pre-filters Symbian/ARM executables
- Other government entities and private companies
- Mostly used for attacker correlation and noise filtering

Limitations

zynamics VxClass

- Heavy obfuscation of control flow
- Virtualizing packers
- Unpacking only works on 32-bit Windows
  - No Linux / OSX / Mobile unpacking
  - 64 bit support is in the works

- Upload of IDBs allows heavy manual intervention beforehand

FAQ

zynamics VxClass

- What OS does it run on ?
  - It runs on a 64-bit Debian Lenny install
- Does it have any network dependencies ?
  - No
- How can we extend the system ?
  - All generated data is accessible through XML-RPC
  - If needed, direct access to the SQL can be used
  - The SQL schema is available on request

Other questions ?

zynamics VxClass