MOSDEF

Dave Aitel
Immunity, Inc
http://www.immunitysec.com









Who am I?

- Founder, Immunity, Inc. NYC based consulting and products company
 - CANVAS: Exploitation Demonstration toolkit
 - BodyGuard: Solaris Kernel Forensics
 - SPIKE, SPIKE Proxy: Application and Protocol Assessment
- Vulns found in:
 - RealServer, IIS, CDE, SQL Server 2000,
 WebSphere, Solaris, Windows

Definitions

- MOSDEF (mose-def) is short for "Most Definately"
- MOSDEF is a retargetable, position independent code, C compiler that supports dynamic remote code linking
- In short, after you've overflowed a process you can compile programs to run inside that process and report back to you

Why?

- _ To Support Immunity CANVAS
 - A sophisticated exploit development and demonstration tool
 - Supports every platform (potentially)
 - 100% pure Python

What's Wrong with Current Shellcode Techniques

- Current Techniques
 - _ Standard execve("/bin/sh")
 - Or Windows CreateProcess("cmd.exe")
 - LSD-Style assembly components
 - Stack-transfer or "syscall-redirection"

Unix: execve("/bin/sh")

- Does not work against chrooted() hosts
 sometimes you cannot unchroot with a simple shellcode
- Annoying to transfer files with echo, printf, and uuencode
- Cannot easily do portforwarding or other advanced requirements

_Windows (cmd.exe redir)

- Loses all current authentication tokens, handles to other processes/files, or other priviledged access
- VERY annoying to transfer files
- Cannot easily do portforwarding or other advanced requirments

_Additionally

- Blobs of "shellcode" inside exploits are impossible to adapt and debug
 - Going to GCC every time you want to modify an exploit's shellcode is a pain
 - Testing and debugging shellcode can waste valuable hours that should be spent coding SPIKE scripts

LSD-style Assembly Components

- Only semi-flexible
 - Not well oriented torwards complex interactions, such as calling CreateProcessAsUser(), fixing a heap, or other advanced techniques while staying in-process to maintain tokens

Little actual connectivity to back-end

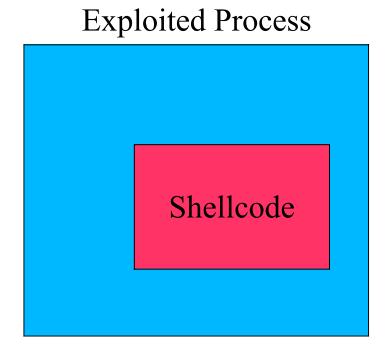
- Choice is to "choose a component" rather than implement any intelligence into your exploits
 - i.e. I want to exploit a process, then if there is an administrative token in it, I want to offer the user the chance to switch to that, and perhaps to switch to any other tokens in the process

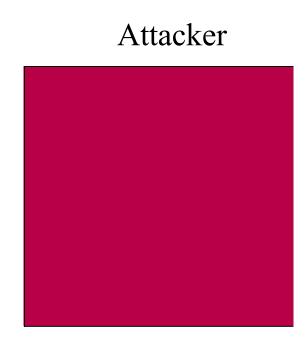
_Not Extensible

- Writing in assembly is a big pain Each component must be written by hand
 - Interacting with the components is done via C – a poor language for large scale projects

Shellcode Missions

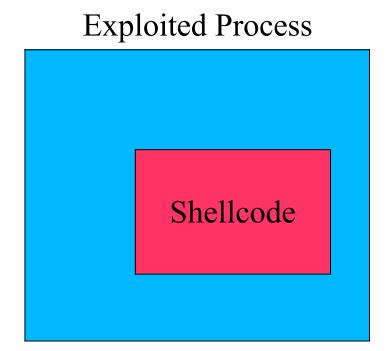
Shellcode can be thought of as two processes

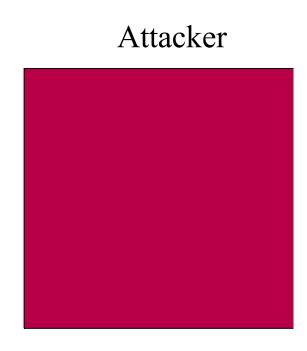




_Shellcode Missions

- _ Step 1 is to establish back-connectivity
- _ Step 2 is to run a mission



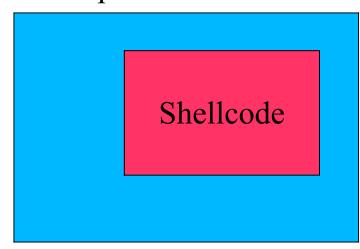


_Establishing Back-Connectivity Step 1 is to establish back-connectivity

- Connect-back
- Steal Socket
- Listen on a TCP/UDP port
- Don't establish any back-connectivity (if mission does not require/cannot get any)

 Attacker

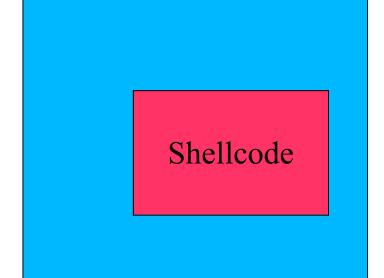
Exploited Process



Running a Mission

- _ Step 2 is to run a mission
 - Recon
 - Trojan Install
 - Etc

Exploited Process

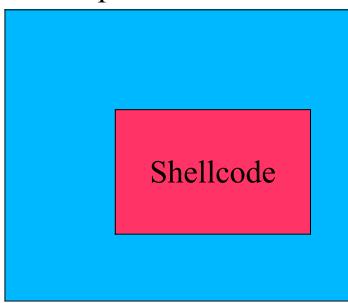


Attacker

_Running a Mission

- _ Missions are supported by various services from the shellcode
 - Shell access
 - File transfer
 - Priviledge manipulation

Exploited Process

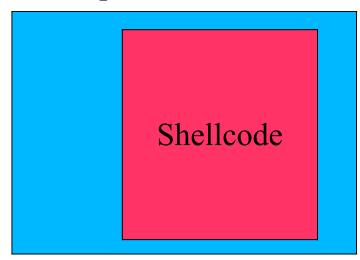


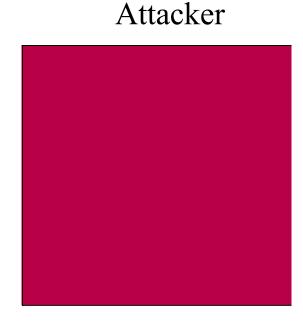
Attacker

_Mission Support

- Missions are poorly supported by traditional execve() shellcode
 - Confuses "pop a shell" with the true mission
 - Moving the mission and the connectivity code into the same shellcode makes for big unwieldy shellcode

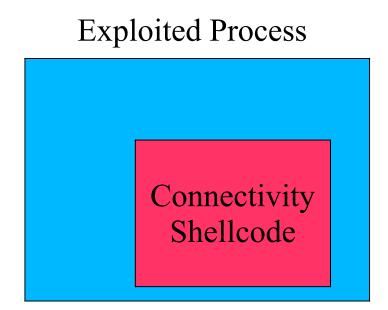
Exploited Process

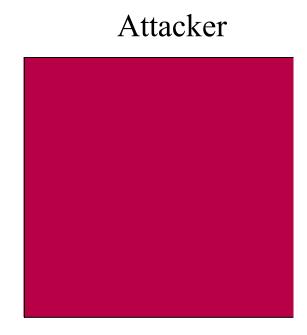




Mission Split

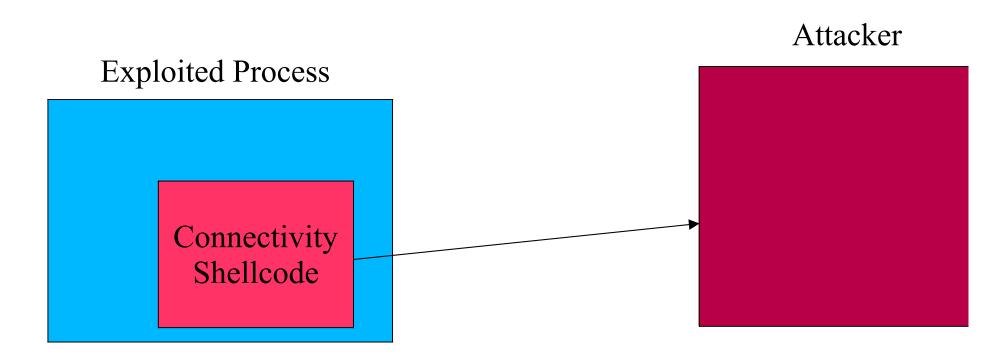
- Solution: split the mission from the stage1 shellcode
 - Smaller, more flexible shellcode





Mission Split

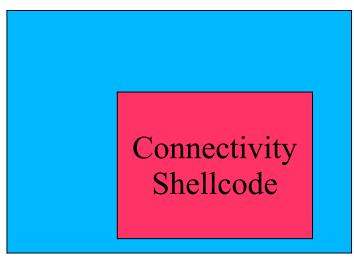
- Solution: split the mission from the stage1 shellcode
 - Smaller, more flexible shellcode
 - Simple paradigm: download more shellcode and execute it



_Stage 2

- _ Options:
 - Send traditional execve() shellcode
 - Or similar 1-shot mission shellcode
 - Establish remote stack-swapping service ("syscall redirection")
 - Establish remote MOSDEF service

Exploited Process

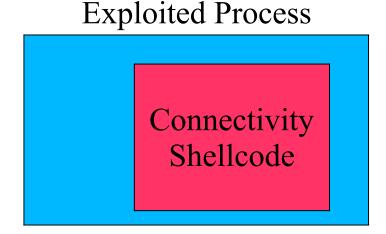


Attacker

Mission belongs here

_Stack Swapping

- _ Aka "Syscall redirection":
 - 3 steps:
 - Send a stack and a function pointer/system call number
 - Remote shellcode stub executes function pointer/system call using stack sent over
 - Entire stack is sent back



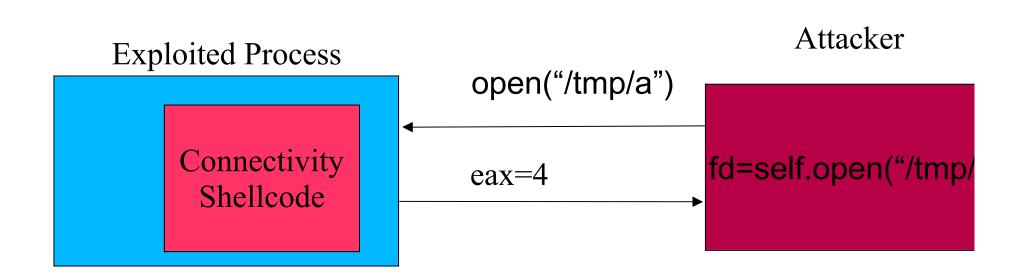
Attacker

Mission belongs here

_Stack Swapping

Benefits

- Interactive with remote machine:
 - Allows for interactive mission support on top of fairly simple shellcode



_Stack Swapping - Benefits

Most function arguments on Unix are easy to marshall and demarshall

```
def unlink(self,path):
                                                  def setuid(self,uid):
                                                        self.setreg("call",posixsyscalls["setuid"])
 Deletes a file - returns -1 on error
                                                        self.setreg("arg1",uid)
 self.setreg("call",posixsyscalls["unlink"])
                                                       request=""
 self.setreg("arg1",self.ESP)
                                                        self.sendrequest(request)
 request=""
                                                        result=self.readresult()
                                                        ret=self.unorder(result[0:4])
 request+=sunstring(path)
 self.sendrequest(request)
                                                        return ret
 result=self.readresult()
 ret=self.unorder(result[0:4])
 return ret
```

_Stack Swapping - Benefits

- Most missions can be supported with relatively few remotely executed functions
 - Execute a command
 - Transfer a File
 - Chdir()
 - Chroot()
 - popen()

- You cannot share a socket with stack swapping shellcode
 - Fork() becomes a real problem
 - Solution: set a fake syscall number for "exec the stack buffer"
 - _ Have to write fork()+anything in assembly
 - Not a nicely portable solution
 - _ Makes our shellcode more complex
 - Still cannot return a different error message for when the fork() fails versus when the execve() fails

- You cannot share a socket with stack swapping shellcode
 - You are going to do one function call at a time
 - China's pingtime is 1 second from my network
 - Those who do not use TCP are doomed to repeat it

Basic stack swapping download code for Solaris

```
def download(self,source,dest):
 downloads a file from the remote server
 infile=self.open(source,O NOMODE)
 if infile==-1:
   return "Couldn't open remote file %s, sorry."%source
 if os.path.isdir(dest):
   dest=os.path.join(dest,source)
 outfile=open(dest,"wb")
 if outfile==None:
   return "Couldn't open local file %s"%dest
 self.log( "infile = %8.8x"%infile)
 data="A"
 size=0
 while data!="":
    data=self.read(infile)
   size+=len(data)
   outfile.write(data)
 self.close(infile)
 outfile.close()
 ----- "D --- 0/ d b. d-- -- -- d--- :--- 0/ -"0/ /-:-- d---t\
```

File download protocol from randomhost.cn

Exploited Process Attacker open(/etc/shadow) 01 = 4read(4,1000 bytes) while data!="": %O1=1000 1000 bytes data=self.read(infi Stack size+=len(data) Swapping outfile.write(data) read(4,1000 bytes) Shellcode self.close(infile) %O1=1000 1000 bytes close(4) %O1=0

time=1second * (sizeof(file)/1000)+2

Exploited Process Attacker open(/etc/shadow) 01 = 4read(4,1000 bytes) while data!="": %O1=1000 1000 bytes data=self.read(infi Stack size+=len(data) Swapping outfile.write(data) read(4,1000 bytes) Shellcode self.close(infile) %O1=1000 1000 bytes close(4) %O1=0

- All iterative operations take 1second * n in China
 - Finding valid thread tokens
 - Downloading and uploading files
 - Executing commands with large output
 - Things I haven't thought of but may want to do in the future
- "But usually you have a fast network!"
- "You can always hand-code these things as a special case to make it faster!"

- Although stack swapping does give us needed dynamic mission support:
 - Inefficient network protocol
 - Inability to do more than one thing at a time
 - Complex functions require painful hand marshalling and demarshalling – or the creation of IDL files and an automatic IDL marshaller, which is just as bad
 - Common requirements, such as fexec() and GetLastError() require special casing – a bad sign
 - Cannot port from one architecture to the other

MOSDEF design requirments

- Efficient network protocol
- The ability to do more than one thing at a time
 - I want cross-platform job control in my shellcode!
- No hand marshalling/demarshalling
- No need to special case fork() or GetLastError()
- Port from one architecture to the other nicely

MOSDEF sample

Compare and Contrast

```
creat(self,filename):
inputs: the filename to open
outpts: returns -1 on failure, otherwise a file handle
truncates the file if possible and it exists
addr=self.getprocaddress("kernel32.dll"," lcreat")
if addr==0:
  print "Failed to find Icreat function!"
  return -1
#ok, now we know the address of Icreat
request=intel order(addr)
request+=intel order(self.ESP+0xc)
request+=intel order(0)
request+=filename+chr(0)
self.sendrequest(request)
result=self.readresult()
fd=istr2int(result[:4])
return fd
```

```
def lcreat(self,filename):
 inputs: the filename to open
 outputs: returns -1 on failure, otherwise a file handle
 truncates the file if possible and it exists
 request=self.compile("""
 #import "remote", "Kernel32. Icreat" as " Icreat"
 #import "local", "sendint" as "sendint"
 #import "string", "filename" as "filename"
 //start of code
 void main()
   int i:
   i= lcreat(filename);
   sendint(i);
 self.sendrequest(request)
 fd=self.readint()
 return fd
```

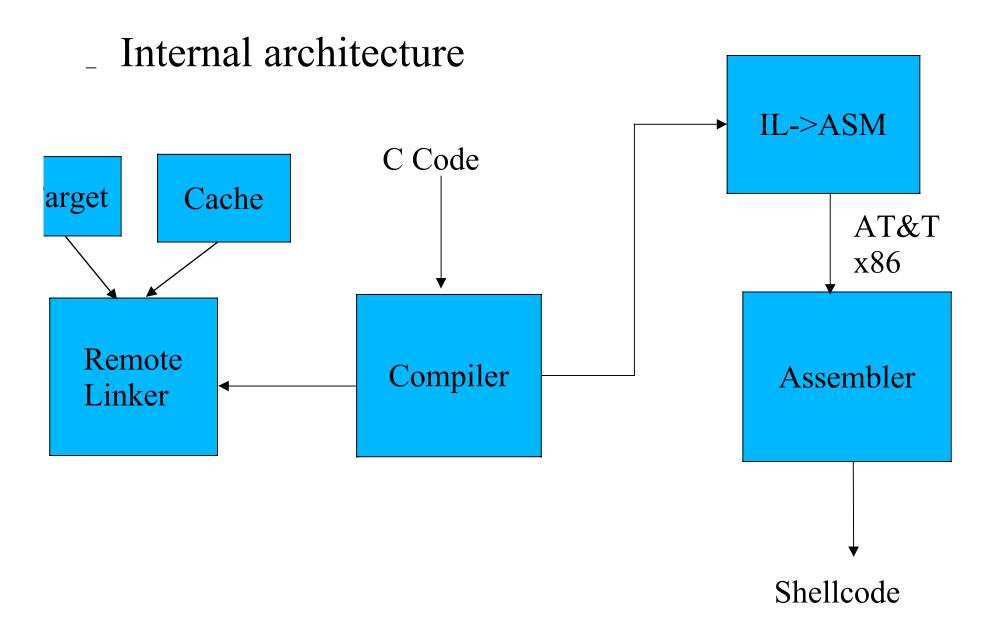
MOSDEF sample

What does this take?

```
ef lcreat(self,filename):
inputs: the filename to open
outputs: returns -1 on failure, otherwise a file handle
truncates the file if possible and it exists
request=self.compile("""
#import "remote", "Kernel32._lcreat" as "_lcreat"
#import "local", "sendint" as "sendint"
#import "string", "filename" as "filename"
//start of code
void main()
  int i;
  i= lcreat(filename);
  sendint(i);
111111
self.sendrequest(request)
fd=self.readint()
```

_A C compiler _An x86 assembler A remote linker

MOSDEF portability



MOSDEF network efficiencies

- While loops are moved to remote side and executed inside hacked process
- Only the information that is needed is sent back – write() only sends 4 bytes back
- Multiple paths can be executed
 - on error, you can send back an error message
 - On success you can send back a data structure

_MOSDEF marshalling

- [Un]Marshalling is done in C
 - Easy to read, understand, modify
 - Easy to port
 - integers don't need re-endianing
 - _ Types can be re-used

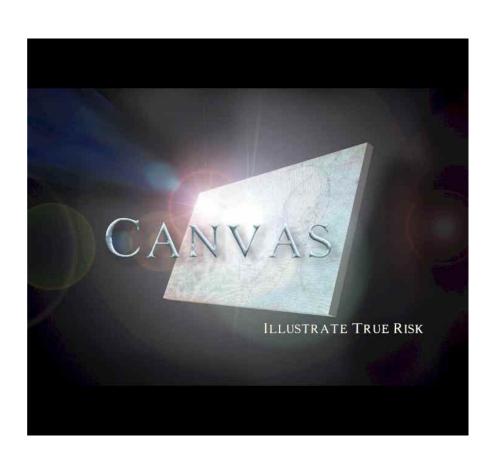
_Cross-platform job control

- The main problem is how to share the outbound TCP socket
 - What we really need is cross-platform locking
 - _ Unix (processes) flock()
 - Windows (threads) EnterCriticalSection()
 - Now we can spin off a "process", and have it report back!
 - The only things that change are sendint(), sendstring() and sendbuffer()
 - These change globally our code does not need to be "thread aware"

_Other benefits

- No special cases
- Having an assembler in pure python gives you the ability to finally get rid of giant blocks of "\xeb\x15\x44\x55\x11" in your exploits. You can just self.assemble() whatever you need
- Future work around finding smaller shellcode, writing shellcode without bad characters, polymorphic shellcode

Conclusion



- MOSDEF is a new way to build attack infrastructures, avoiding many of the problems of earlier infrastructures
- Prevent hacker starvation buy CANVAS for \$995 today
- More information on this and other fun things at http://www.immunitysec.com/