Information Security 5 Secure System Engineering Professor Chester Rebeiro Indian Institute of Technology Madras Demonstration of Position Independent Code Mod03_Lec19

Hello and welcome to this demonstration in the course for secure system engineering, in this demo will be actually looking at an example of how position independent code works, so in the previous demonstration we looked at runtime relocatable code and we seen that it requires that the loader have a lot of thing to do and it requires the loader to go and modify the code and fix the address in each of the locations and a lot of this becomes much more simplified with a PIC, a position independent code, so we will use exactly the same code as we have done the previous demonstration, this code can be downloaded in the virtual machine that comes along with this course.

(Refer Slide Time: 1:05)





And this particular example is nptel_codes module 5 relocgot, so we use exactly the same library as before, the library is called mylib.c and which is as shown over here it has mylib_int, setmylib_int and getmylib_int at to set and get functions and the driver for this as before is here which has an invocation to setmylib_int and getmylib_int, in order to generate position independent code the command or the compiler options are slightly different, so in order to do this in this example how a makefile would require makelib_pic which would generate the library as a position independent code library or a pic library.

So as you see here we have in additional option minus F pic which would generate, which would cost the compiler to generate a position independent code library and as before we also dump this assembly of this library in library in library.

the open this disassembly is that the code looks very different compared to the one we seen previously that is with the load time relocation.

(Refer Slide Time: 2:44)





Okay, so let us, so what you see over here and if we actually compare this with the previous library code that we obtained in, you see that this and this are completely different, so will go into what exactly happens over here, now the setmylib_int essentially would take an input parameter X and said the global variable mylib_int with that particular value, so this is done very differently a with pic code, so essentially this two instructions creates the stack frame, the next instruction is a call to this function call_X86.get_pc_thunk.ax, so essentially what this function does? This function is something which the compiler adds in.

(Refer Slide Time: 4:15)



This function just as two statements, it moves the stack pointer or the contents of the stack pointer into the EAX register and then returns, so note that when a call is done the return address for this call that is corresponding to, the address corresponding to this instruction is pushed onto the stack and at that time the stack pointer is pointing to that return address, therefore what is moved into the EAX register is the address of the following instruction that is this instruction.

The reason why we do this is that you want to get the address of this particular instruction and this instruction can be relocatable, so therefore we cannot hard core the address but rather find an indirect way to actually get the address of this instruction, next we add a value of 1A B8 to EAX register, so this value gives the address of a table known as a GOT table, so this is the GOT table, therefore at the end of this instruction the address of the GOT table is move into the EAX register and then at an offset of -14 in the GOT table is where the actual address of mylib_int is stored.

So this address is then moved into the EAX register and the value of X present in the stack at a location 8 bytes from the frame pointer is then moved into the EDX register, then X is stored into mylib_int, so let us see GDB working with this, so we will open this and run GDB, executable for the driver is d pic, so we run d pic as it is, ops as before we need to export the LD library path okay, so as expected we would see in mylib is 100 and the value of log is 5555.

(Refer Slide Time: 7:01)





So let us see how this is working internally with the GDB, gdb./dpic, we put a breakpoint at setmylib_int and run the program and then we disassemble the function, so you see this okay, so there is a call to this getpcthunk.ax, so as we know what happens over here is the address of the next instruction that is this instruction gets moved into the EAX register right and let us open G edit and just make note of it that the EAX register comprises of the next instruction, to this EAX register we are adding the offset of 1AB8.

Let see what we get over here, take the calculator put it in hexadecimal mode and plus 1AB8, so what we see is that we get an address of X7FD3000, address of our GOT table is F7FD3000 so we can just store it over here and we can also verify this by looking at GDB and we can see this by looking at contents of the EAX register as follows, register EAX and we see that indeed it has F7FD3000.

Now within the GOT table at an offset of 0X-14 is where we have the address of the mylib_int present, so let us subtract from this, the value of 0X14 and we see this entry, so F7FD3000-14 is this entry which is the entry in GOT table which contains the address of the variable mylib_int.

(Refer Slide Time: 10:30)



So let us actually look at the contents of this location, so you can dump that memory location using the GDB command x/x and paste and put a 0X here and we see this F7, so this address is the location of mylib_int, so we can check this by also printing the address of mylib_int, where you see that it is identical, so essentially what is happening here is the GOT table stores the offset for each and every variable present in the program, this GOT table is fixed in the program space but the variables move into any region in the program space.

The loader will ensure that the contents of the GOT table will be appropriately modified, so as to reflect the correct address of these variables, so we can actually play a small trick over here, if we modify this particular location which contains the address of the global variable mylib_int, we can actually cost the program to behave spuriously.

(Refer Slide Time: 12:10)



So see the driver.c and let us look at the address of this global variable glob, so that is done by p/x glob and we see that this has an address 804A024, now what we can do is change the contents of the GOT table to point to glob, so this is done as follows, we can set int to be equal to ampersand glob, so what we have done over here is that the GOT entry for mylib_int now points to this global variable glob.

You can check this, so like p/x okay, so this is the address of glob and now if you continue executing what you see is that we have got a invalid a malicious output, so the value of glob according to this program should have been 5555 but since we have modify the GOT table therefore glob actually has a value of 100, so in this way even though ASLR actually makes attacks much more difficult but still there are potential flaws which an attacker could use to

manipulate the working of the program, it could actually change the behaviour of the program and so on. Thank you.