|  |  |
| --- | --- |
|  | **Hello, dear colleagues!**  **My name is Denis Bakhvalov, I’m from RFSW department. Today we will talk about C++ Linker.**  **First of all I would like to thank you for your coming!**  **For C++ developers work of C++ linker is usually hidden. We only remind ourselves about C++ linker when we see strange errors like “unresolved symbols … smth smth”. I think it will be good to know about possible traps that we can get in, right?**  **We will cover some basic things about work of linker, but even though I’m sure that there will be a lot of interesting information also for advanced C++ developers.**  **I prepared quite a big amount of information for today so we will split it into two parts with a short break when you can ask some questions or go for a cup of coffee.**  **Please be ready for some C++ code and diagrams.** |
|  | So, what linker is doing? To answer this question let’s take a look at compilation steps. We have four main stages. And today we will focus on last step. But still lets quickly run through all steps just to remind ourselves about them. |
|  | First we have preprocessing stage. On this stage all preprocessor commands are executed. Like #include, #define and others. So, here is one example. You can see what is the output of this first stage.  **But if I will include for example #include <string> it will brings everything inside this header file to our cpp file.**  **And the first tip comes right away – don’t include everything in your \*.cpp files. It can drastically increase size of your binary files.**  **But okay, I’m pretty sure that you are aware of it, so let’s move on.** |
|  | Ok, second step is translation. So on this stage compiler is working. Lexical analyzer is checking syntax and your file is parsed and divided into tokens. As the output we have assembler listing of your code.  In it you can find the names of function, static variable and some numbers. |
|  | On assembling stage your assembler program is transformed into machine code (binary). Result is stored in binary object files \*.obj.  **Simply, you can open assembler listing using text editor, but you can’t open object file, because there are only bytes. But we have dedicated tools for it. Don’t worry we will see a lot of examples of their usage today.**  One of such tools is nm. Here is the example of it’s output. As you can see we have name of our function, this constant here… |
|  | So now, let me introduce one more tool that we will use for analysis. Its name is objdump. And here is one example of its output just to give you an intuition of what it can do for you.  -s option shows you the contents of the sections inside object file.  Full description can be found here…  **Later we will use this tool with some other options as this is quite a powerful tool.** |
|  | Ok, what’s next? I want to remind you that only cpp files are compiled in C++. In other words header files are not compiled standalone and always included in some file with implementation, known as cpp files. And as an output we have object file for each cpp file. But please note that on compilation stage cpp files don’t know anything about each other.  So, consider such example (write on the board):  void foo();  int main()  {  foo();  }  This code is absolutely valid C++, besides the fact it’s not linking. But compilation stage is successful. So, to have it working foo() should be somewhere defined. And if we will combine those 2 \*.cpp files together everything will be just perfect. |
|  | Like in this example.  So, this is exactly what linker is doing. It combines many object files together, building a bridge between them. This link here is this bridge. When those two object files are combined into one executable it’s pretty clear where to go if we want to call function A.  **Linker was designed for two mainly reasons. First to have interoperability with such languages like fortran and cobol. Second is because of memory limit while compiling the huge C or C++ program. As the compiler has to hold a lot of meta information about the code, it compiles.** |
|  | On this slide you can see the linker’s job from other point of view.  In main.o we have A() function unresolved (which upper case “U” stands for). But in A.o it is defined according to the upper case letter “T” and it is also defined in the final executable file.  So, we can say that linker is responsible for resolving all undefined symbols from all object files included into the linkage.  So now, I think we can understand why those error messages like “unresolved symbols … smth smth” come from linker. |
|  | As an extension I would like to show you this example.  Usually upper case letters are used for symbols with external linkage and lower case for symbols with internal linkage. B or b means that symbol is uninitialized. d means that symbol is defined. Symbols with no linkage (like localInt) are not considered by linker.  Full description can be found here… |
|  | **So far seems easy isn’t it? ☺**  So, let’s consider the following example. I choose the name rules without enforcement. Later you will understand why. So, We have almost the same two cpp files with struct S defined and functions fooA and fooB. Note, that definition of foo methods of structure A in those files are different.  In main we are simply invoking those two functions fooA and fooB.  What do you think will be the output of this program? |
|  | You can see in this example we have output A.cpp twice. And I can say that this situation is really might happen |
|  | Ok, now we are just observing the problem… Let’s change the order of linkage of A.o and B.o. Interesting what will be the output? |
|  | And here it is. Now we’ve got B.cpp twice. So, regarding this we can make the hypothesis that when linker found one definition of A it skips all other.  **For example I found this problem when I have been writing two different fake classes in UnitTests. I created two derived classes with the same name, from one base class. And overridden their behavior in a different way and put them in global namespace.**  **And really my first reaction to this was: “why linker is silent here? It won’t be silent if I will define two global functions with the same signatures.”**  **Of course we will know the answer to this a little bit later, but let’s proceed with the analysis first.** |
|  | Here you can see the objdump output with –d options. This option shows you the disassembled output for our binary. As you can see for both fooA and fooB we call the same implementation of method foo(). |
|  | And if we will look at foo() method we will see that the string A.cpp is used.  If we revert the order of object files linkage than we will see “B.cpp.A.cpp.” here  **I want to make a comment here regarding to the read only data section. If you will try to remove const qualifier using const\_cast <> of C-style casts, generally it will lead us to UB. Because in some platforms this section is protected from write operations. But if you create your variable on the stack, then you can do your cast operations “safely”.** |
|  | Ok. So now, let’s see what C++ standard says about it…  One of the reasons of such behavior is the possibility to put the definition of class templates and template functions inside header files. Although there might be some other reasons.  Although still I and I guess a lot of other developers are quite angry about such silent behavior.  If this example is clear I will move forward… |
|  | I’m almost sure that you heard about static libraries. In general it is just a collection (archive of object files).  On the picture you can see that many object files are combined in one static library. Usually they are named like…  **You can use static library as a building block for many of yours applications.** |
|  | As shown here. Executable statically link libA.a …  We will look on how to build and use static library just in a few slides.  **Not so long time ago I faced the question do linker throw away unused code? I didn’t knew the answer and I decided to clarify it. So, the question to you, how do you think do linker throws away unused code? And what is unused code?** |
|  | Ok. Let’s consider one example.  We have struct Counter… |
|  | Two cpp files with static instances and a main function that simply prints the counter value.  **What you think will be the output?** |
|  | **2? Do someone has another opinion? Well, very reasonable.** |
|  | But now, let’s ship A.o, B.o and Counter.o as a static library. Whoa. And the output now is 0. It is a little bit unexpected.  **The reason of such behavior is that linker drops unused object files only if they originated from a library to keep your program small.**  **Starting from the translation unit where main() is defined every referenced symbol causes object file to be included. And so on. The code that is not referenced is unused and will be dropped if it would be shipped as a library.** |
|  | If we will explicitly put A.o to the linkers input, than result will be one. This will not allow linker to throw unreferenced code away. |
|  | Also there is special option that forces linker to include all object files from static library. But be careful with it, as it can drastically increase the size of your binary. And do not forget to close it with -Wl,-no-whole-archive. |
|  | END OF PART 1. |

“Whenever I wondered: “Why it is so?”, I found good reasons for every single language decision, with efficiency and compatibility being the most frequent answer”

|  |  |
| --- | --- |
|  | So, now we will talk about shared libraries. On windows it is also called dynamic libraries or dll.  **Dynamic libraries also can be used like building blocks for different processes but they can be shared in runtime.**  First main advantage is memory saving. If we will consider this example we can see that the same version of printf function is included into two executable files. Idea of shared library was to share this function between different processes in runtime and keep only one copy of it.  Second is modularity. It is not so obvious but really important advantage. If we want to change printf we don’t need to recompile all executables that use this function. Once we modify it and rebuild the shared library everyone can use it after restart the application.  And third main advantage is the ability to load and unload library at any moment you want. For example your application can be relatively small and consume little amount of memory and load and unload the dynamic libraries at the time you want to use some functionality. |
|  | Static libraries allow you to call any methods from the libraries you want and C++ syntax allow you to do.  In Linux all symbols are exported automatically as it is done for static libraries.  So, to hide some symbol you can use compiler flag, visibility attribute or special pragma.  But on Windows shared libraries by default do not export all the symbols and you should do it manually.  We can consider Windows dlls as some API. You can share those functions that you want.  There are several ways to export symbols. We will consider only one. I’m pretty sure that you will find the rest on the web. |
|  | **Shared libraries also can be used in a way similar to static libraries. You can specify whether you want to link it on compile-time or in runtime. We will consider those examples, so don’t worry.**  **Static linkage means that you want to use shared library during the whole lifetime of your application. It will be loaded with your application and unloaded when your application will finish.**  As you can see here we building shared library first. And say to linker to create static dependency from the shared library.  **Please note, that symbols found in shared library are not included into executable. Instead of it name of the symbols are placed. Why? Simply because at compile time we don’t know the exact address of the function inside shared library. It is not even loaded into the memory.**  This additional stage is done when loading shared library into the memory. Now we can put the physical addresses from shared memory. In simple words there is additional stage for the linker. |
|  | Let’s consider example of static linking of shared library. **Please sorry for a little bit confusing, but from now I will use the terms shared library and dynamic library and will mean the same thing both for Linux and Windows.**  We have shared library that has function main\_DLL(). And we have executable with function main(). |
|  | To build shared library you don’t need to use ar tool, but you need to specify –shared parameter and link it to your executable in the way similar to static library.  This LD\_LIBRARY\_PATH environment variable specify the path where OS will search for your shared libraries. |
|  | I would like to introduce new tool that you can use for your analysis. It’s name is ldd. It shows all static dependencies of your application. If you such error… |
|  | Ok, so now let’s look on dynamic linkage.  **Dynamic linkage means that you want to use shared library only during some period of the lifetime of your application.**  Here you can see that there is no dependency between those components on the compilation stage. And all dependencies are visible only in the code.  We can manipulate loading, using and unloading of the library from the code. |
|  | Here we should add extern C qualifier for our function inside shared library. Because in C++ the symbol names are mangled and if we won’t do that we won’t be able to call this function. I will just say few words, that this qualifier force compiler not to decorate names of our symbols.  More information you can find on the web. |
|  | And actually here is the code for loading library and invoking main\_DLL() from it. |
|  | And here is the compilation steps that we need to make. As you can see we only loading ld which includes functions like dlopen() and dlsym(). |
|  | To give you an intuition how it is working let’s consider following example. **This example is from my own experience and yes I did spent quite some time on fixing this problem.**  So, we have struct that count it’s instances. This code should be very clear to you. |
|  | We have shared library that has the function that create one instance of this class and prints the address of the static counter and its value. |
|  | Ok. And we have an executable that call main\_DLL function and repeat those actions once again. |
|  | But wait a minute… Let’s think for a while. What address we should put in that place?  First, that we should keep in mind is that only text segment of dynamic library could be shared between processes. So, when two different processes will load our library they will have their own counters.  Look at the picture on the right. Here we can assume that this counter is one of this red symbols. If for example we will put in the dynamic library text segment the address of counter inside process1 then this is what will happen when process2 will come up in the game.  LTR: if several applications needed a dynamic library at the same time, it would mean that you would have exactly that many copies of the same dynamic library in memory.  Second drawback: if you have a lot of places where your counter is referenced – it means that you will need to patch as many places as you have.  Third drawback is that code (.text) segment is writable which means potential security problems. |
|  | So, taking into consideration all drawback new concept was invented.  The idea of PIC is to maintain a special table where we will keep the offsets (or addresses) of those red symbols. So, in the code we will place the offset to the this table. On linking time those offset are known already so, there is no problem for the linker to do it.  But then, after loader will load this library it will patch this GOT with the right address according to the each process memory layout. |
|  | Ok. Now let’s get back to our example. Let’s build and launch it on Linux. And we can see the same counter is addressed from both DLL and executable. |
|  | Let’s prove that output by doing a little analysis once again. We can see that both symbols have V class. According to … And we can see the address of the static member and we can conclude that this member is inside executable.  If it is clear than we will move forward. |
|  | Let’s now see what will be on Windows. We need to use those \_\_declspec qualifiers with export on DLL side and import on executable side. |
|  | If we will build it and launch on Win we will see this a little bit surprising result.  We see that those static members are different for DLL and executable. |
|  | Let’s do analysis once again. And we see that both in DLL and executable those static members have class D which stands for defined. And also we can spot the addresses of them. |
|  | So, this is the schematic view of what is going on in our situation. **We have two static counters, one came from DLL and another one is let’s say our own counter. I showed them both here, because the counter that comes from DLL is not shared between different processes. That’s why physically this counter will be placed inside address space of our process, not in shared memory. Here only main\_DLL is really shared.**  And on the picture below you can see desired situation when we have only one counter, like in Linux case. |
|  | Let me just draw the same picture in a little bit different manner. I want to show you the source of those static members. Note that those components are built separately, that’s why they creating their own copies of static members. |
|  | What you think how we can solve this problem?  **I can already hear how you want to say: “Don’t use Windows!”. But I want to say that this is not correct answer. Ok? ☺**  What we want to do is to break the link between static counter and executable and export this function from DLL. With this approach executable will not place one instance of static member but will imports the one from shared library. |
|  | So, first we split declaration and definition of Counter. And create special macros that we will use further. On dll side it will expand like export, on executable like import. This is an absolutely normal way to process export/import symbols. |
|  | And we export counter class and main\_DLL from DLL using this macros. |
|  | On executable side we are importing counter class and main\_DLL using the same macros. |
|  | And we received the result that we want. Simply, we hided this counter from executable. So now executable cannot create its own counter. But on the other way we export this counter from dll. |
|  | To make two different counters on Linux I haven’t found any other way but to make such a “hack”. You need to redefine the name of the class. Thus you will force compiler to create different counters, because now there are two different classes.  **But generally this problem could have more elegant solution. We can transform shared library into static and link them like that (draw the picture on the board).** |
|  | All examples as well as this presentation you can find under this link : <https://github.com/dendibakh/Linker>... |
|  | **Ok, I would to thank you very much for your attention. Hope, that now you know a little bit more about C++ Linker. Thank you!**  **Now, if you have questions feel free to ask.** |

“Whenever I wondered: “Why it is so?”, I found good reasons for every single language decision, with efficiency and compatibility being the most frequent answer”