

我国基础软件现状，以编译器为例

编译原理第一次作业

卢雨轩 19071125

2022 年 2 月 21 日

基础软件包括什么？回答这个问题，只需看计算机科学与技术专业的本科生学的四大基础课程：《计算机组成原理》、《数据库原理》、《操作系统》、《编译原理》，也就是说，EDA 软件、数据库、操作系统以及编译器是四种最基础的基础软件。

国际上，知名的编译器有 GCC、LLVM 等支持众多语言的编译器组合，也有 ICC、MSVC 等 C / C++ 编译器，也有 Go、RustC、TSC 等各个语言的编译器。现有的『国产』编译器中，比较知名的有提供更为容易上手的 GPU 编程能力的『太极 (Taichi)』^[1]，也有『木兰』等面向嵌入式平台但是由于宣发问题广受批评的编译器。由此可见，对于基础语言（如 C++、C）的编译器，我国还是显著依赖于开源编译器，并没有实现真正的『自主』。

那么，使用这些开源编译器『安全』吗？答案当然是不安全。如果不经代码审计而直接使用，无论是编译产物中注入还是在标准库中注入，攻击者都可以轻易构造攻击向量。如，密码学算法大多依赖安全且具有足够熵的随机数，OpenSSL 等库均使用自己编写的随机数生成算法而不是语言标准库。

既然如此，那么，是否有必要投入编译器等基础软件的开发呢？我的答案是『有』。对于已有成熟开源产品的应用场景，如编译器或操作系统，我们可以采取审计源码的方式，维护一个可控的分支。对于没有成熟开源产品的领域，如 EDA 软件，我们则要投入精力开发。

参考文献

- [1] HU Y. The taichi programming language[C/OL]//SIGGRAPH '20: ACM SIGGRAPH 2020 Courses. Association for Computing Machinery, 2020: 1-50. <https://doi.org/10.1145/3388769.3407493>.