

目录

环境.....	1
LLVM 编译 linux-5.4 内核并获取 CFG	1
以 sysfs_get_uname 函数为例.....	3

环境

```
ubuntu@ubuntu-VirtualBox:~$ uname -a
Linux ubuntu-VirtualBox 6.8.0-45-generic #45-Ubuntu SMP PREEMPT_DYNAMIC Fri Aug
30 12:02:04 UTC 2024 x86_64 x86_64 x86_64 GNU/Linux
ubuntu@ubuntu-VirtualBox:~$ cat /etc/os-release
PRETTY_NAME="Ubuntu 24.04.1 LTS"
NAME="Ubuntu"
VERSION_ID="24.04"
VERSION="24.04.1 LTS (Noble Numbat)"
VERSION_CODENAME=noble
ID=ubuntu
ID_LIKE=debian
HOME_URL="https://www.ubuntu.com/"
SUPPORT_URL="https://help.ubuntu.com/"
BUG_REPORT_URL="https://bugs.launchpad.net/ubuntu/"
PRIVACY_POLICY_URL="https://www.ubuntu.com/legal/terms-and-policies/privacy-poli
cy"
UBUNTU_CODENAME=noble
LOGO=ubuntu-logo
```

LLVM 编译 linux-5.4 内核并获取 CFG

安装 llvm clang graphviz(编译过程中会遇到缺其他工具，按照提示进行安装即可)

```
apt install llvm clang graphviz build-essential
```

下载 [Linux-5.4 tarball](#)

```
wget https://cdn.kernel.org/pub/linux/kernel/v5.x/linux-5.4.284.tar.xz
```

```
tar -xvf linux-5.4.284.tar.xz
```

安装 [wllvm](#)

```
python -m venv ~/venv00
```

```
source ~/venv00/bin/bash
```

```
pip install wllvm
```

```
(venv00) ubuntu@ubuntu-VirtualBox:~$ pip show wllvm
Name: wllvm
Version: 1.3.1
Summary: Whole Program LLVM
Home-page: https://github.com/SRI-CSL/whole-program-llvm
Author: Ian A. Mason, Tristan Ravitch, Dan Liew, Bruno Dutertre, Benjamin Schubert, Berkeley Churchill, Marko Dimjasevic, Will Dietz, Fabian Mager, Ben Liblit, Andrew Santosa, Tomas Kalibera, Loic Gelle, Joshua Cranmer, Alexander Bakst, Miguel Arroyo.
Author-email: iam@csl.sri.com
License: MIT
Location: /home/ubuntu/venv00/lib/python3.12/site-packages
Requires:
Required-by:
(venv00) ubuntu@ubuntu-VirtualBox:~$ python --version
Python 3.12.3
```

编译 linux-5.4

```
cd linux-5.4.284
export LLVM_COMPILER=clang
make CC=clang defconfig
make CC=wllvm LLVM=1
```

```
(venv00) ubuntu@ubuntu-VirtualBox:~/code/linux-5.4.284$ ls
arch          fs            LICENSES      net           usr
block         include       MAINTAINERS   README        virt
certs         init          Makefile      samples       vmlinux
COPYING       ipc           mm            scripts        vmlinux.bc
CREDITS       Kbuild       modules.builtin  security      vmlinux.o
crypto        Kconfig      modules.builtin.modinfo  sound
Documentation kernel        modules.order   System.map
drivers       lib          Module.symvers  tools
```

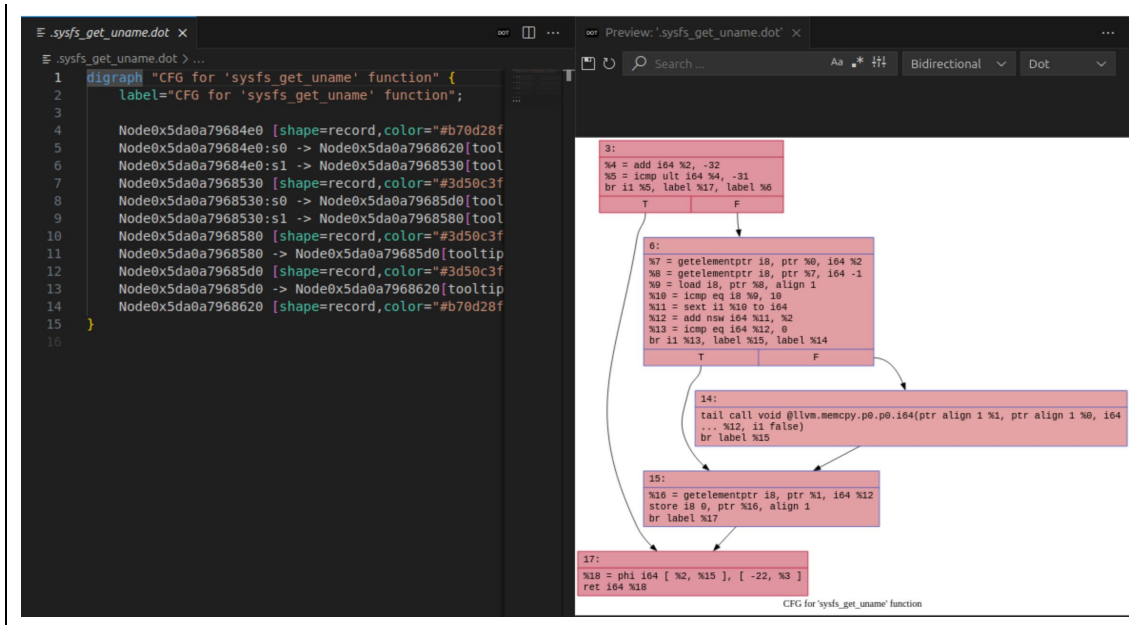
从 bitcode 提取 dot(dot 文件很多)

```
extract-bc vmlinux.bc
```

```
≡ .sysfs_format_mac.dot
≡ .sysfs_fs_context_free.dot
≡ .sysfs_get_tree.dot
≡ .sysfs_get_uname.dot
≡ .sysfs_init_fs_context.dot
≡ .sysfs_init.dot
≡ .sysfs_kf_bin_mmap.dot
```

dot 文件转换为 CFG 图

借助 vscode 中的 Graphviz Interactive Preview 插件



以 sysfs_get_uname 函数为例

- sysfs_get_uname 函数位于 kernel/time/clocksource.c 文件中

```
1095 ssize_t sysfs_get_uname(const char *buf, char *dst, size_t cnt)
1096 {
1097     size_t ret = cnt;
1098
1099     /* strings from sysfs write are not 0 terminated! */
1100     if (!cnt || cnt >= CS_NAME_LEN)
1101         return -EINVAL;
1102
1103     /* strip of \n: */
1104     if (buf[cnt-1] == '\n')
1105         cnt--;
1106     if (cnt > 0)
1107         memcpy(dst, buf, cnt);
1108     dst[cnt] = 0;
1109     return ret;
1110 }
```

- 将 vmlinux.bc 转换为可阅读的 vmlinux.ll, 全局搜索 sysfs_get_uname 找到对应的代码
llvm-dis vmlinux.bc

```
553545 ; Function Attrs: mustprogress norecurse noredzone nosync nounwind null pointer_is_valid sspstrong willreturn memory(argmem: readwrite)
553546 define dso local noundef i64 @sysfs_get_uname(ptr nocapture noundef readonly %0, ptr nocapture noundef writeonly %1, i64 noundef %2) local_unnamed_addr #37 {
553547     %4 = add i64 %2, -32
553548     %5 = icmp ult i64 %4, -31
553549     br i1 %5, label %17, label %6
553550
553551 6:                                     : preds = %3
553552     %7 = getelementptr i8, ptr %0, i64 %2
553553     %8 = getelementptr i8, ptr %7, i64 -1
553554     %9 = load i8, ptr %8, align 1
553555     %10 = icmp eq i8 %9, 10
553556     %11 = sext i1 %10 to i64
553557     %12 = add nsw i64 %11, %2
553558     %13 = icmp eq i64 %12, 0
553559     br i1 %13, label %15, label %14
553560
553561 14:                                     : preds = %6
553562     tail call void @llvm.memcpy.p0.p0.i64(ptr align 1 %1, ptr align 1 %0, i64 %12, i1 false)
553563     br label %15
553564
553565 15:                                     : preds = %14, %6
553566     %16 = getelementptr i8, ptr %1, i64 %12
553567     store i8 0, ptr %16, align 1
553568     br label %17
553569
553570 17:                                     : preds = %3, %15
553571     %18 = phi i64 [ %2, %15 ], [ -22, %3 ]
553572     ret i64 %18
553573 }
553574
```

- 通过 dot 命令获取 CFG 的 png 图片

dot -Tpng -o sysfs_get_uname.png .sysfs_get_uname.dot

```
(venv00) ubuntu@ubuntu-VirtualBox:~/code/linux-5.4.284$ dot -Tpng -o sysfs_get_uname.png .sysfs_get_uname.dot
(venv00) ubuntu@ubuntu-VirtualBox:~/code/linux-5.4.284$ ls
arch          init          modules.builtin      sysfs_get_uname.png
block         ipc          modules.builtin.modinfo System.map
certs         Kbuild       modules.order        tools
COPYING       Kconfig      Module.symvers       usr
CREDITS       kernel       net                  virt
crypto        lib          README               vmlinux
Documentation LICENSES     samples              vmlinux.bc
drivers        MAINTAINERS  scripts              vmlinux.ll
fs            Makefile     security             vmlinux.o
```

- 查看 CFG 图片

CFG for 'sysfs_get_uname' function

```

1095 ssize_t sysfs_get_uname(const char *buf, char *dst, size_t cnt)
1096 {
1097     size_t ret = cnt;
1098
1099     /* strings from sysfs write are not 0 terminated! */
1100     if (!cnt || cnt >= CS_NAME_LEN)
1101         return -EINVAL;
1102
1103     /* strip of '\n' */
1104     if (buf[cnt-1] == '\n')
1105         cnt--;
1106     if (cnt > 0)
1107         memcpy(dst, buf, cnt);
1108     dst[cnt] = 0;
1109     return ret;
1110 }

```

一个C语言语句一般对应多条llvm ir指令

比如

if(buf[cnt-1] == '\n')

对应

%7 = getelementptr i8, ptr %0, i64 %2; 获取 buf[cnt]

%8 = getelementptr i8, ptr %7, i64 -1; 获取 buf[cnt-1]

%9 = load i8, ptr %8, align 1; 加载 buf[cnt-1]

%10 = icmp eq i8 %9, 10; 比较是否等于 '\n'