

# 华为方舟编译器开发文档

书栈(BookStack.CN)

# 目 录

致谢

架构设计

环境配置

开发指南

    源码编译

    编程规范

        原则

        命名

        格式

        注释

        头文件

        作用域

        类

        函数

        C++其他特性

        现代C++特性

如何贡献

    文档贡献指南

    代码贡献指南

    邮件列表指南

FAQ

# 致谢

当前文档 《华为方舟编译器开发文档》 由 进击的皇虫 使用 书栈(BookStack.CN) 进行构建，生成于 2019-08-31。

书栈(BookStack.CN) 仅提供文档编写、整理、归类等功能，以及对文档内容的生成和导出工具。

文档内容由网友们编写和整理，书栈(BookStack.CN) 难以确认文档内容知识点是否错漏。如果您在阅读文档获取知识的时候，发现文档内容有不恰当的地方，请向我们反馈，让我们共同携手，将知识准确、高效且有效地传递给每一个人。

同时，如果您在日常工作、生活和学习中遇到有价值有营养的知识文档，欢迎分享到 书栈(BookStack.CN) ，为知识的传承献上您的一份力量！

如果当前文档生成时间太久，请到 书栈(BookStack.CN) 获取最新的文档，以跟上知识更新换代的步伐。

内容来源：华为 <https://www.openarkcompiler.cn/document/frameworkDesgin>

文档地址：<http://www.bookstack.cn/books/openarkcompiler>

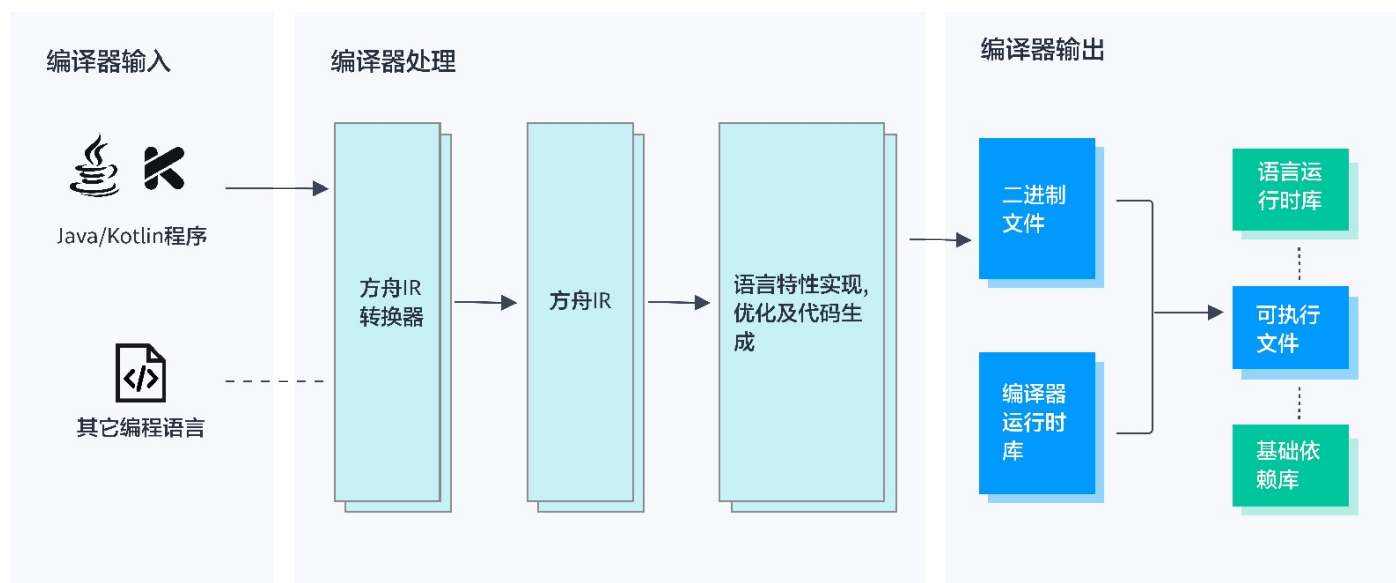
书栈官网：<http://www.bookstack.cn>

书栈开源：<https://github.com/TruthHun>

分享，让知识传承更久远！ 感谢知识的创造者，感谢知识的分享者，也感谢每一位阅读到此处的读者，因为我们都将成为知识的传承者。

# 架构设计

## 方舟编译器架构示意图



方舟编译器架构示意图

当前方舟编译器支持Java/Kotlin程序字节码的前端输入，其它编程语言的支持（如 C/C++/JS 等）还在规划中，方舟编译器的中间表示（IR）转换器将前端输入转换成方舟IR，并输送给后端的优化器，最终生成二进制文件，二进制文件与编译器运行时库文件链接生成可执行文件，在方舟的运行环境中就可执行该文件

方舟编译器IR是支持程序编译和运行的中间程序表示。程序源代码中的任何信息对于程序分析和优化都是有幫助的，所以方舟IR的目标是尽可能完整详细地提供源程序的信息。关于方舟编译器IR的详细信息，请参考文档：[方舟IR设计](#)

方舟编译器开源范围请参考 [这里](#)

# 环境配置

## 硬件推荐配置

- 2 GHz 双核处理器或者更高等级CPU
- 2 GB 系统内存及以上
- 200GB 可用磁盘空间

## 开发环境推荐

您需要安装一个64位版本的Ubuntu（推荐Ubuntu 16.04）

```
sudo apt-get -y install openjdk-8-jdk git-core gnupg flex bison gperf build-essential zip curl zlib1g-dev  
libc6-dev-i386 lib32ncurses5-dev x11proto-core-dev libx11-dev lib32z-dev ccache libgl1-mesa-dev libxml2-utils  
xsltproc unzip lib32z1-dev qemu g++-multilib gcc-multilib python3-paramiko python-paramiko python-jenkins  
1. python-requests python-xlwt libgl2.0-dev libpixmap-1-dev linux-libc-dev:i386  
2. sudo apt-get -y install gcc-5-aarch64-linux-gnu g++-5-aarch64-linux-gnu
```

## 安装Clang编译器并完成配置（用于编译方舟编译器代码）

下载clang+llvm-8.0.0-x86\_64-linux-gnu-ubuntu-16.04

LLVM下载地址：<http://releases.llvm.org/download.html#8.0.0>

放置到 `openarkcompiler/tools` 目录，打开 `openarkcompiler/build/config/BUILDCONFIG.gn` 文件，  
将 `GN_C_COMPILER`、`GN_CXX_COMPILER` 和 `GN_AR` 三个变量配置为Clang编译器所在路径，例如：

```
1. GN_C_COMPILER = "${MAPLE_ROOT}/tools/clang_llvm-8.0.0-x86_64-linux-gnu-ubuntu-16.04/bin/clang"  
2. GN_CXX_COMPILER = "${MAPLE_ROOT}/tools/clang_llvm-8.0.0-x86_64-linux-gnu-ubuntu-16.04/bin/clang++"  
3. GN_AR = "${MAPLE_ROOT}/tools/clang_llvm-8.0.0-x86_64-linux-gnu-ubuntu-16.04/bin/llvm-ar"
```

其中`${MAPLE_ROOT}`为openarkcompiler源码根目录。

## 安装Ninja、GN并完成配置

下载Ninja(v1.9.0)及GN(Linux Version)Ninja下载地址：<https://github.com/ninja-build/ninja/releases>

GN下载地

址：[https://archive.softwareheritage.org/browse/content/sha1\\_git:2dc0d5b26caef44f467de8120b26f8aad8b878be/raw/?filename=gn](https://archive.softwareheritage.org/browse/content/sha1_git:2dc0d5b26caef44f467de8120b26f8aad8b878be/raw/?filename=gn)

将GN和Ninja可执行程序放置到openarkcompiler/tools目录，打开openarkcompiler/Makefile文件，将GN和NINJA两个变量配置为GN和Ninja可执行程序所在路径。例如：

1. GN := `${MAPLE_ROOT}/tools/gn/gn`
2. NINJA := `${MAPLE_ROOT}/tools/ninja_1.9.0/ninja`

# 开发者指南

---

通过参考本文档，您可以下载编译器源码编译出OpenArkCompiler。

请先按照 [环境配置](#) 文档完成您的开发环境准备。

# 源码编译

## 源码下载

下载地址：<https://code.opensource.huaweicloud.com/HarmonyOS/OpenArkCompiler/home>，可以通过 Clone or Download 的方式下载openarkcompiler源码

注：默认源码下载目录为openarkcompiler

## 源码编译

在openarkcompiler目录下执行以下命令，编译出OpenArkCompiler，默认输出路径openarkcompiler/out/bin。

```
1. source build/envsetup.sh
2. make
```

命令说明：

- `source build/envsetup.sh` 初始化环境，将OpenArkCompiler工具链路径openarkcompiler/src/bin设置到环境变量中
- `make` 编译OpenArkCompiler的Release版本
- `make BUILD_TYPE=DEBUG` 编译OpenArkCompiler的Debug版本



# 编程规范

---

## 目的

规则并不是完美的，通过禁止在特定情况下有用的特性，可能会对代码实现造成影响。但是我们制定规则的目的“为了大多数程序员可以得到更多的好处”，如果在团队运作中认为某个规则无法遵循，希望可以共同改进该规则。

参考该规范之前，希望您具有相应的C++基础能力，而不是通过该文档来学习C++。

- 了解C++的ISO标准；
- 熟知C++的基本语言特性，包括C++ 03/11/14/17相关特性；
- 了解C++的标准库；

## 重点关注

- 约定C++的编程风格，比如命名，排版等。
- C++的模块化设计，如何设计头文件，类，接口和函数。
- C++相关特性的优秀实践，比如常量，类型转换，资源管理，模板等。
- 现代C++的优秀实践，包括C++11/14/17中可以提高代码可维护性，提高代码可靠性的相关约定。

## 约定

规则：编程时必须遵守的约定(must)

建议：编程时应该遵守的约定(should)

本规范适用通用C++标准，如果没有特定的标准版本，适用所有的版本(C++03/11/14/17)。

## 例外

无论是'规则'还是'建议'，都必须理解该条目这么规定的原因，并努力遵守。但是，有些规则和建议可能会有例外。

在不违背总体原则，经过充分考虑，有充足的理由的前提下，可以适当违背规范中约定。例外破坏了代码的一致性，请尽量避免。'规则'的例外应该是极少的。

下列情况，应风格一致性原则优先：修改外部开源代码、第三方代码时，应该遵守开源代码、第三方代码已有规范，保持风格统一。**\*\*某些特定领域，优先参考其行业规范。\*\***

# 原则

## 好代码的原则

我们参考Kent Beck的简单设计四原则来指导我们的如何写出优秀的代码，如何有效地判断我们的代码是优秀的。

- 通过所有测试 (Passes its tests)
- 尽可能消除重复 (Minimizes duplication)
- 尽可能清晰表达 (Maximizes clarity)
- 更少代码元素 (Has fewer elements)
- 以上四个原则的重要程度依次降低。这组定义被称做简单设计原则。第一条强调的是外部需求，这是代码实现最重要的；第二点就是代码的模块架构设计，保证代码的正交性，保证代码更容易修改；第三点是代码的可阅读性，保证代码是容易阅读的；最后一点才是保证代码是简洁的，在简洁和表达力之间，我们更看重表达力。

## 类和函数设计指导原则

C++是典型的面向对象编程语言，软件工程界已经有很多OOP原则来指导我们编写大规模的，高可扩展的，可维护性的代码：

- 高内聚，低耦合的基本原则
- SOLID原则
- 迪米特法则
- “Tell, Don’t ask”原则
- 组合/聚合复用原则

## 保证静态类型安全

我们希望C++应该是静态类型安全的，这样可以减少运行时的错误，提高代码的健壮性。但是由于C++的下面的特性存在，会破坏C++静态类型安全，我们针对这部分特性要仔细处理。

- unions联合体
- 类型转换cast
- 缩窄转换narrowing conversions
- 类型退化type decay
- 范围错误range errors
- `void*` 类型指针我们可以通过约束这些特性的使用，或者使用C++的新特性，比如variant(C++17)，GSL的span，narrow\_cast等来解决这些问题，提高C++代码的健壮性。

## 遵循C++ISO标准

希望通过使用ISO C++标准的特性来编写C++代码，对于ISO标准中未定义的或者编译器实现的特性要谨慎使用，对于GCC等编译器的提供的扩展特性也需要谨慎使用，这些特性会导致代码的可移植性比较差。

注意：如果模块中需要使用相关的扩展特性来，那么尽可能将这些特性封装成独立的接口，并且可以通过编译选项关闭或者编译这些特性。对于这些扩展特性的使用，请模块制定特性编程指南来指导这些特性的使用。

## 优先编译时检查错误

通过编译器来优先保证代码健壮性，而不是通过编写错误处理代码来处理编译就可以发现的异常，比如：

- 通过`const`来保证数据的不变性，防止数据被无意修改。
- 通过`gsl::span`等来保证`char`数组不越界，而不是通过运行时的`length`检查。
- 通过`static_assert`来进行编译时检查。

## 使用命名空间来限定作用域

全局变量，全局常量和全局类型定义由于都属于全局作用域，在项目中，使用第三方库中容易出现冲突。

命名空间将作用域细分为独立的，具名的作用域，可有效地防止全局作用域的命名冲突。

- `class`，`struct`等都具有自己的类作用域。
- 具名的`namespace`可以实现类作用域更上层的作用域。
- 匿名`namespace`和`static`可以实现文件作用域。对于没有作用域的宏变量，宏函数强烈建议不使用。

作用域的一些缺点：

- 虽然可以通过作用域来区分两个命名相同的类型，但是还是具有迷惑性。
- 内联命名空间会让命名空间内部的成员摆脱限制，让人迷惑。
- 通过多重嵌套来定义`namespace`，会让完整的命名空间比较冗长。所以，我们使用命名空间的建议如下：
- 对于变量，常量和类型定义尽可能使用`namespace`，减少全局作用域的冲突
- 不要在头文件中使用`using namespace`
- 不要使用内联命名空间
- 鼓励在`.cpp`文件中通过匿名`namespace`或者`static`来封装，防止不必要的定义通过API暴露出去。

## 优先使用C++特性而不是C特性

C++比起C语言更加类型安全，更加抽象。我们更推荐使用C++的语言特性来编程，比如使用`string`而不是 `char*`，使用`vector`而不是原生数组，使用`namespace`而不是`static`。

# 命名

## 通用命名

常见命名风格有：驼峰风格(**CamelCase**)大小写字母混用，单词连在一起，不同单词间通过单词首字母大写来分开。按连接后的首字母是否大写，又分：大驼峰(UperCamelCase)和小驼峰(lowerCamelCase)

内核风格(**unix\_like**)单词全小写，用下划线分割。如：'test\_result'

匈牙利风格在‘大驼峰’的基础上，加上前缀；前缀用于表达类型或用途。如：'uiSavedCount', 'bTested'

### 规则2.1.1 标识符命名使用驼峰风格

不考虑匈牙利命名，在内核风格与驼峰风格之间，根据存量代码的情况，我们选择驼峰风格。

类型	命名风格
类类型，结构体类型，枚举类型，联合体类型等类型定义	大驼峰
函数(包括全局函数，作用域函数，成员函数)	大驼峰（接口部分可加前缀，如XXX_函数名）
全局变量(包括全局和命名空间域下的变量，类静态变量)，局部变量，函数参数，类、结构体和联合体中的成员变量	小驼峰
常量(const)，枚举值	k+大小写混合
宏	大写+下划线
命名空间	全小写

注意：上表中常量是指全局作用域、namespace域、类的静态成员域下，以 const或constexpr 修饰的基本数据类型、枚举、字符串类型的变量。上表中变量是指除常量定义以外的其他变量，均使用小驼峰风格。

## 文件命名

### 建议2.2.1 C++文件以.cpp结尾，头文件以.h结尾

我们推荐使用.h作为头文件的后缀，这样头文件可以直接兼容C和C++。我们推荐使用.cpp作为实现文件的后缀，这样可以直接区分C++代码，而不是C代码。

目前业界还有一些其他的后缀的表示方法：

- 头文件：.hh, .hpp, .hxx
- cpp文件：.cc, .cxx, .C对于本文档，我们默认使用.h和.cpp作为后缀。

### 建议2.2.2 C++文件名和类名保持一致

C++的头文件和cpp文件名和类名保持一致，使用下划线小写风格。

如下：

- database\_connection.h
- database\_connection.cpp 结构体，命名空间，枚举等定义的文件名类似。

## 函数命名

函数命名统一使用大驼峰风格，一般采用动词或者动宾结构。接口部分可加前缀，如XXX\_函数名。

```
1. class List {
2.     public:
3.         void AddElement(const Element& element);
4.         Element GetElement(const unsigned int index) const;
5.         bool IsEmpty() const;
6.         bool MCC_GetClass();
7. };
8.
9. namespace utils {
10. void DeleteUser();
11. }
```

## 类型命名

类型命名采用大驼峰命名风格。所有类型命名——类、结构体、联合体、类型定义（typedef）、枚举——使用相同约定，例如：

```
1. // classes, structs and unions
2. class UrlTable { ...
3. class UrlTableTester { ...
4. struct UrlTableProperties { ...
5. union Packet { ...
6.
7. // typedefs
8. typedef std::map<std::string, UrlTableProperties*> PropertiesMap;
9.
10. // enums
11. enum UrlTableErrors { ...
```

对于命名空间的命名，建议全小写：

```
1. // namespace
2. namespace osutils {
3.
4. namespace fileutils {
5.
6. }
7.
8. }
```

### 建议2.4.1 避免滥用 typedef或者#define 对基本类型起别名

除有明确的必要性，否则不要用 `typedef/#define` 对基本数据类型进行重定义。优先使用 `<stdint>` 头文件中的基本类型：

有符号类型	无符号类型	描述
<code>int8_t</code>	<code>uint8_t</code>	宽度恰为8的有/无符号整数类型
<code>int16_t</code>	<code>uint16_t</code>	宽度恰为16的有/无符号整数类型
<code>int32_t</code>	<code>uint32_t</code>	宽度恰为32的有/无符号整数类型
<code>int64_t</code>	<code>uint64_t</code>	宽度恰为64的有/无符号整数类型
<code>intptr_t</code>	<code>uintptr_t</code>	足以保存指针的有/无符号整数类型

如果模块有自己的定义，请使用统一的typedef来定义类型：

```
1. typedef signed char VOS_INT8;
2. typedef unsigned char VOS_UINT8;
3.
4. #if __WORDSIZE == 64
5. typedef unsigned long int VOS_UINTPTR;
6. #else
7. typedef unsigned int VOS_UINTPTR;
8. #endif
```

如果模块为了封装某个类型的信息，方便后续的扩展，可以使用typedef来重新定义。

```
1. typedef uint8_t DeviceID;
2. // ...
3. // 若干版本后扩展成 16-bit
4. typedef uint16_t DeviceID;
```

有特殊作用的类型 `typedef void Handle;` 注意：\*不要使用 `#define` 进行别名定义，并且在C++11以后推荐使用 `using`来定义类型。

除上述理由外，应避免给其本数值类型别名定义。因为类型别名可读性并不好，隐藏了基本数值类型信息，如位宽，是否带符号。滥用举例：

```
1. typedef uint16_t MyCounter;
2. // ...
3. int Foo(...) {
4.     MyCounter c;
5.     // ...
6.     while (c >= 0) {
7.         printf("counter = %d\n", c);
8.         // ...
9.     }
10.    // ...
11. }
```

对 'MyCounter' 是否可能小于0，打印时用 '%d' 还是 '%u' 都不是很直观，极容易引入上述类似缺陷。

## 变量命名

通用变量命名采用小驼峰，包括全局变量，函数形参，局部变量，成员变量。

```
1. std::string tableName; // Good: 推荐此风格
2. std::string tablename; // Bad: 禁止此风格
3. std::string path;      // Good: 只有一个单词时，小驼峰为全小写
```

### 规则2.5.1 类的成员变量命名使用小驼峰。

```
1. class Foo {
2.     private:
3.         std::string fileName; // 不添加任何作用域前缀或者后缀
4. };
```

当构造函数参数和成员变量重名时，可通过this->来引用成员变量。

```
1. class MyClass {
2.     public:
3.         MyClass(int myVar) : myVar(myVar) { // OK, 初始化列表允许同名入参初始化同名成员
4.             if (NeedNewVar()) {
5.                 this->myVar = GetValue(); // 注意不要漏掉this->, 否则就成了给入参赋值
6.             }
7.         }
8.
9.     private:
10.         int myVar;
11. };
```

## 宏、常量、枚举命名

宏采用全大写，下划线连接的格式。常量、枚举值使用k+大小写混合。函数局部 const 常量和类的普通const成员变量，使用小驼峰命名风格。

```
1. #define MAX(a, b) (((a) < (b)) ? (b) : (a)) // 仅对宏命名举例，并不推荐用宏实现此类功能
2.
3. enum TintColor { // 注意，枚举类型名用大驼峰，其下面的取值是k+大小写混合
4.     kRed,
5.     kDarkRed,
6.     kGreen,
7.     kLightGreen
8. };
9.
10. int Func(...) {
11.     const unsigned int bufferSize = 100; // 函数局部常量
12.     char *p = new char[bufferSize];
13.     ...
14. }
```

```
15.  
16. namespace utils {  
17.     const unsigned int kFileSize = 200;    // 全局常量  
18. }
```



# 格式

尽管有些编程的排版风格因人而异，但是我们强烈建议和要求使用统一的编码风格，以便所有人都能够轻松的阅读和理解代码，增强代码的可维护性。

## 行宽

### 建议3.1.1 行宽不超过 120 个字符

建议每行字符数不要超过 120 个。如果超过120个字符，请选择合理的方式进行换行。

例外：

- 如果一行注释包含了超过120 个字符的命令或URL，则可以保持一行，以方便复制、粘贴和通过grep查找；
- 包含长路径的 `#include` 语句可以超出120 个字符，但是也需要尽量避免；
- 编译预处理中的error信息可以超出一行。预处理的 `error` 信息在一行便于阅读和理解，即使超过 120 个字符。

```
1. #ifndef XXX_YYY_ZZZ
   #error Header aaaa/bbbb/cccc/abc.h must only be included after xxxx/yyyy/zzzz/xyz.h, because
2. xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
3. #endif
```

## 缩进

### 规则3.2.1 使用空格进行缩进，每次缩进2个空格

只允许使用空格(space)进行缩进，每次缩进为 2 个空格。

## 大括号

### 规则3.3.1 除函数外，使用 K&R 缩进风格

**K&R**风格函数左大括号跟随语句放行末。右大括号独占一行，除非后面跟着同一语句的剩余部分，如 `do` 语句中的 `while`，或者 `if` 语句的 `else/else if`，或者逗号、分号。

如：

```
1. struct MyType {      // 跟随语句放行末，前置1空格
2.     ...
3. };
4.
5. int Foo(int a) {      // 函数左大括号跟随语句放行末
6.     if (...) {
7.         ...
8.     } else {
```

```

9.      ...
10.    }
11. }

```

推荐这种风格的理由：

- 代码更紧凑；
- 相比另起一行，放行末使代码阅读节奏感上更连续；
- 符合后来语言的习惯，符合业界主流习惯；
- 现代集成开发环境（IDE）都具有代码缩进对齐显示的辅助功能，大括号放在行尾并不会对缩进和范围产生理解上的影响。对于空函数体，可以将大括号放在同一行：

```

1. class MyClass {
2.     public:
3.         MyClass() : value(0) {}
4.
5.     private:
6.         int value;
7. };

```

## 函数声明和定义

### 规则3.4.1 函数声明和定义的返回类型和函数名在同一行；函数参数列表超出行宽时要换行并合理对齐

在声明和定义函数的时候，函数的返回值类型应该和函数名在同一行；如果行宽度允许，函数参数也应该放在一行；否则，函数参数应该换行，并进行合理对齐。参数列表的左圆括号总是和函数名在同一行，不要单独一行；右圆括号总是跟随最后一个参数。

换行举例：

```

1. ReturnType FunctionName(ArgType paramName1, ArgType paramName2) { // Good : 全在同一行
2.     ...
3. }
4.
5. ReturnType VeryVeryVeryLongFunctionName(ArgType paramName1,      // 行宽不满足所有参数，进行换行
6.                                           ArgType paramName2,      // Good : 和上一行参数对齐
7.                                           ArgType paramName3) {
8.     ...
9. }
10.
11. ReturnType LongFunctionName(ArgType paramName1, ArgType paramName2, // 行宽限制，进行换行
12.                              ArgType paramName3, ArgType paramName4, ArgType paramName5) { // Good : 换行后 4 空格缩进
13.     ...
14. }
15.
16. ReturnType ReallyReallyReallyReallyLongFunctionName(                // 行宽不满足第1个参数，直接换行
17.     ArgType paramName1, ArgType paramName2, ArgType paramName3) { // Good : 换行后 4 空格缩进
18.     ...

```

```
19. }
```

## 函数调用

### 规则3.5.1 函数调用入参列表应放在一行，超出行宽换行时，保持参数进行合理对齐

函数调用时，函数参数列表放在一行。参数列表如果超过行宽，需要换行并进行合理的参数对齐。左圆括号总是跟函数名，右圆括号总是跟最后一个参数。

换行举例：

```
1. ReturnType result = FunctionName(paramName1, paramName2); // Good : 函数参数放在一行
2.
3. ReturnType result = FunctionName(paramName1,
4.                                paramName2, // Good : 保持与上方参数对齐
5.                                paramName3);
6.
7. ReturnType result = FunctionName(paramName1, paramName2,
8.    paramName3, paramName4, paramName5); // Good : 参数换行, 4 空格缩进
9.
10. ReturnType result = VeryVeryVeryLongFunctionName( // 行宽不满足第1个参数, 直接换行
11.    paramName1, paramName2, paramName3); // 换行后, 4 空格缩进
```

如果函数调用的参数存在内在关联性，按照可理解性优先于格式排版要求，对参数进行合理分组换行。

```
1. // Good : 每行的参数代表一组相关性较强的数据结构，放在一行便于理解
2. int result = DealWithStructureLikeParams(left.x, left.y, // 表示一组相关参数
3.    right.x, right.y); // 表示另外一组相关参数
```

## if语句

### 规则3.6.1 if语句必须要使用大括号

我们要求if语句都需要使用大括号，即便只有一条语句。

理由：

- 代码逻辑直观，易读；
- 在已有条件语句代码上增加新代码时不容易出错；
- 对于在if语句中使用函数式宏时，有大括号保护不易出错（如果宏定义时遗漏了大括号）。

```
1. if (objectIsNotExist) { // Good : 单行条件语句也加大括号
2.     return CreateNewObject();
3. }
```

### 规则3.6.2 禁止 if/else/else if 写在同一行

条件语句中，若有多个分支，应该写在不同行。

如下是正确的写法：

```
1. if (someConditions) {
2.     DoSomething();
3.     ...
4. } else { // Good: else 与 if 在不同行
5.     ...
6. }
```

下面是不符合规范的案例：

```
1. if (someConditions) { ... } else { ... } // Bad: else 与 if 在同一行
```

## 循环语句

### 规则3.7.1 循环语句要求使用大括号

和if语句类似，我们要求for/while循环语句必须加上的大括号，即使循环体是空的，或者循环语句只有一条。

```
1. for (int i = 0; i < someRange; i++) {
2.     DoSomething();
3. }
```

如果循环体是空的，应该使用空的大括号，而不是使用单个分号。 单个分号容易被遗漏，也容易被误认为是循环语句中的一部分。

```
1. for (int i = 0; i < someRange; i++) { } // Good: for循环体是空，使用大括号，而不是使用分号
2.
3. while (someCondition) { } // Good: while循环体是空，使用大括号，而不是使用分号
4.
5. while (someCondition) {
6.     continue; // Good: continue表示空逻辑，可以使用大括号也可以不使用
7. }
```

坏的例子：

```
1. for (int i = 0; i < someRange; i++) ; // Bad: for循环体是空，也不要只使用分号，要使用大括号
2.
3. while (someCondition) ; // Bad: 使用分号容易让人误解是while语句中的一部分
```

## switch语句

### 规则3.8.1 switch 语句的 case/default 要缩进一层

switch 语句的缩进风格如下：

```

1. switch (var) {
2.     case 0:                // Good: 缩进
3.         DoSomething1();    // Good: 缩进
4.         break;
5.     case 1: {              // Good: 带大括号格式
6.         DoSomething2();
7.         break;
8.     }
9.     default:
10.        break;
11. }
```

```

1. switch (var) {
2. case 0:                    // Bad: case 未缩进
3.     DoSomething();
4.     break;
5. default:                  // Bad: default 未缩进
6.     break;
7. }
```

## 表达式

### 建议3.9.1 表达式换行要保持换行的一致性，运算符放行末

较长的表达式，不满足行宽要求的时候，需要在适当的地方换行。一般在较低优先级运算符或连接符后面截断，运算符或连接符放在行末。运算符、连接符放在行末，表示“未结束，后续还有”。例：

// 假设下面第一行已经不满足行宽要求

```

1. if (currentValue > threshold && // Good: 换行后，逻辑操作符放在行尾
2.     someConditionion) {
3.     DoSomething();
4.     ...
5. }
6.
7. int result = reallyReallyLongVariableName1 + // Good
8.     reallyReallyLongVariableName2;
```

表达式换行后，注意保持合理对齐，或者4空格缩进。参考下面例子

```

1. int sum = longVariableName1 + longVariableName2 + longVariableName3 +
2.     longVariableName4 + longVariableName5 + longVariableName6; // Good: 4空格缩进
3.
4. int sum = longVariableName1 + longVariableName2 + longVariableName3 +
5.     longVariableName4 + longVariableName5 + longVariableName6; // Good: 保持对齐
```

## 变量赋值

### 规则3.10.1 多个变量定义和赋值语句不允许写在一行

每行只有一个变量初始化的语句，更容易阅读和理解。

```
1. int maxCount = 10;
2. bool isCompleted = false;
```

下面是不符合规范的示例：

```
1. int maxCount = 10; bool isCompleted = false; // Bad : 多个变量初始化需要分开放在多行，每行一个变量初始化
2. int x, y = 0; // Bad : 多个变量定义需要分行，每行一个
3.
4. int pointX;
5. int pointY;
6. ...
7. pointX = 1; pointY = 2; // Bad : 多个变量赋值语句放同一行
```

例外：for 循环头、if 初始化语句（C++17）、结构化绑定语句（C++17）中可以声明和初始化多个变量。这些语句中的多个变量声明有较强关联，如果强行分成多行会带来作用域不一致，声明和初始化割裂等问题。

## 初始化

初始化包括结构体、联合体、及数组的初始化

### 规则3.11.1 初始化换行时要有缩进，并进行合理对齐

结构体或数组初始化时，如果换行应保持4空格缩进。从可读性角度出发，选择换行点和对齐位置。

```
1. const int rank[] = {
2.     16, 16, 16, 16, 32, 32, 32, 32,
3.     64, 64, 64, 64, 32, 32, 32, 32
4. };
```

## 指针与引用

### 建议3.12.1 指针类型"\*"跟随变量名或者类型，不要两边都留有或者都没有空格

指针命名： \* 靠左靠右都可以，但是不要两边都有或者都没有空格。

```
1. int* p = NULL; // Good
2. int *p = NULL; // Good
3.
4. int*p = NULL; // Bad
5. int * p = NULL; // Bad
```

例外：当变量被 `const` 修饰时，"`*`" 无法跟随变量，此时也不要跟随类型。

```
1. char * const VERSION = "V100";
```

## 建议3.12.2 引用类型"&"跟随变量名或者类型，不要两边都留有或者都没有空格

引用命名：`&` 靠左靠右都可以，但是不要两边都有或者都没有空格。

```
1. int i = 8;
2.
3. int& p = i;      // Good
4. int &p = i;      // Good
5.
6. int & p = i;     // Bad
7. int&p = i;      // Bad
```

## 编译预处理

### 规则3.13.1 编译预处理的"#"统一放在行首，嵌套编译预处理语句时，"#"不缩进

编译预处理的"#"统一放在行首，即使编译预处理的代码是嵌入在函数体中的，"#"也应该放在行首。

```
1. #if defined(__x86_64__) && defined(__GCC_HAVE_SYNC_COMPARE_AND_SWAP_16) // Good : "#"放在行首
2. #define ATOMIC_X86_HAS_CMPXCHG16B 1 // Good : "#"放在行首
3. #else
4. #define ATOMIC_X86_HAS_CMPXCHG16B 0
5. #endif
6.
7.
8. int FunctionName() {
9.     if (somethingError) {
10.         ...
11. #ifdef HAS_SYSLOG // Good : 即便在函数内部，"#"也放在行首
12.     WriteToSysLog();
13. #else
14.     WriteToFileLog();
15. #endif
16. }
17. }
```

### 内嵌的预处理语句"#"不缩进

```
1. #if defined(__x86_64__) && defined(__GCC_HAVE_SYNC_COMPARE_AND_SWAP_16)
2. #define ATOMIC_X86_HAS_CMPXCHG16B 1 // Good : 区分层次，便于阅读
3. #else
4. #define ATOMIC_X86_HAS_CMPXCHG16B 0
5. #endif
```

## 空格和空行

### 建议3.14.1 水平空格应该突出关键字和重要信息，避免不必要的留白

水平空格应该突出关键字和重要信息，每行代码尾部不要加空格。总体规则如下：

- if, switch, case, do, while, for等关键字之后加空格；
- 小括号内部的两侧，不要加空格；
- 大括号内部两侧有无空格，左右必须保持一致；
- 一元操作符(& \* + - ~ ! )之后不要加空格；
- 二元操作符(= + - < > \* / % | & ^ <= >= == != )左右两侧加空格
- 三目运算符( ? : )符号两侧均需要空格
- 前置和后置的自增、自减( ++ -- )和变量之间不加空格
- 结构体成员操作符( . -> )前后不加空格
- 逗号( , )前面不加空格，后面增加空格
- 对于模板和类型转换(<>)和类型之间不要添加空格
- 域操作符(: :)前后不要添加空格
- 冒号(:)前后根据情况来判断是否要添加空格常规情况：

```
1. void Foo(int b) { // Good : 大括号前应该留空格
2.
3. int i = 0; // Good : 变量初始化时, =前后应该有空格, 分号前面不要留空格
4.
5. int buf[kBufSize] = {0}; // Good : 大括号内两侧都无空格
```

函数定义和函数调用：

```
1. int result = Foo(arg1,arg2);
2.           ^ // Bad: 逗号后面需要增加空格
3.
4. int result = Foo( arg1, arg2 );
5.           ^ ^ // Bad: 函数参数列表的左括号后面不应该有空格, 右括号前面不应该有空格
```

指针和取地址

```
1. x = *p; // Good : *操作符和指针p之间不加空格
2. p = &x; // Good : &操作符和变量x之间不加空格
3. x = r.y; // Good : 通过.访问成员变量时不加空格
4. x = r->y; // Good : 通过->访问成员变量时不加空格
```

操作符：

```
1. x = 0; // Good : 赋值操作的=前后都要加空格
2. x = -5; // Good : 负数的符号和数值之前不要加空格
3. ++x; // Good : 前置和后置的++/--和变量之间不要加空格
4. x--;
5.
6. if (x && !y) // Good : 布尔操作符前后要加上空格, !操作和变量之间不要空格
```



## 格式

```
7. v = w * x + y / z; // Good : 二元操作符前后要加空格
8. v = w * (x + z);    // Good : 括号内的表达式前后不需要加空格
9.
10. int a = (x < y) ? x : y; // Good : 三目运算符, ? 和 : 前后需要添加空格
```

## 循环和条件语句:

```
1. if (condition) { // Good : if关键字和括号之间加空格, 括号内条件语句前后不加空格
2.     ...
3. } else {          // Good : else关键字和大括号之间加空格
4.     ...
5. }
6.
7. while (condition) {} // Good : while关键字和括号之间加空格, 括号内条件语句前后不加空格
8.
9. for (int i = 0; i < someRange; ++i) { // Good : for关键字和括号之间加空格, 分号之后加空格
10.     ...
11. }
12.
13. switch (condition) { // Good: switch 关键字后面有1空格
14.     case 0:           // Good : case语句条件和冒号之间不加空格
15.         ...
16.         break;
17.         ...
18.     default:
19.         ...
20.         break;
21. }
```

## 模板和转换

```
1. // 尖括号(< and >) 不与空格紧邻, < 前没有空格, > 和 ( 之间也没有.
2. vector<string> x;
3. y = static_cast<char*>(x);
4.
5. // 在类型与指针操作符之间留空格也可以, 但要保持一致.
6. vector<char *> x;
```

## 域操作符

```
1. std::cout; // Good: 命名空间访问, 不要留空格
2.
3. int MyClass::GetValue() const {} // Good: 对于成员函数定义, 不要留空格
```

## 冒号

```
1. // 添加空格的场景
2.
3. // Good: 类的派生需要留有空格
4. class Sub : public Base {
```

```
5.
6. };
7.
8. // 构造函数初始化列表需要留有空格
9. MyClass::MyClass(int var) : someVar(var) {
10.     DoSomething();
11. }
12.
13. // 位域表示也留有空格
14. struct XX {
15.     char a : 4;
16.     char b : 5;
17.     char c : 4;
18. };
```

```
1. // 不添加空格的场景
2.
3. // Good: 对于public:, private:这种类访问权限的冒号不用添加空格
4. class MyClass {
5.     public:
6.         MyClass(int var);
7.     private:
8.         int someVar;
9. };
10.
11. // 对于switch-case的case和default后面的冒号不用添加空格
12. switch (value) {
13.     case 1:
14.         DoSomething();
15.         break;
16.     default:
17.         break;
18. }
```

注意：当前的集成开发环境（IDE）可以设置删除行尾的空格，请正确配置。

## 建议3.14.2 合理安排空行，保持代码紧凑

减少不必要的空行，可以显示更多的代码，方便代码阅读。下面有一些建议遵守的规则：

- 根据上下内容的相关程度，合理安排空行；
- 函数内部、类型定义内部、宏内部、初始化表达式内部，不使用连续空行
- 不使用连续 **3** 个空行，或更多
- 大括号内的代码块行首之前和行尾之后不要加空行。

```
1. int Foo() {
2.     ...
3. }
4.
5.
6. // Bad：两个函数定义间超过了一个空行
```

```

7.  int Bar() {
8.     ...
9. }
10.
11.
12. if (...) {
13.     // Bad : 大括号内的代码块行首不要加入空行
14.     ...
15.     // Bad : 大括号内的代码块行尾不要加入空行
16. }
17.
18. int Foo(...) {
19.     // Bad : 函数体内行首不要加空行
20.     ...
21. }

```

## 类

规则3.15.1 类访问控制块的声明依次序是 `public:`, `protected:`, `private:`, 每个都缩进 1 个空格

```

1. class MyClass : public BaseClass {
2.     public:          // 注意没有缩进
3.         MyClass(); // 标准的4空格缩进
4.         explicit MyClass(int var);
5.         ~MyClass() {}
6.
7.         void SomeFunction();
8.         void SomeFunctionThatDoesNothing() {
9.             }
10.
11.        void SetVar(int var) { someVar = var; }
12.        int GetVar() const { return someVar; }
13.
14.    private:
15.        bool SomeInternalFunction();
16.
17.        int someVar;
18.        int someOtherVar;
19. };

```

在各个部分中，建议将类似的声明放在一起，并且建议以如下的顺序：类型（包括 `typedef`, `using` 和嵌套的结构体与类），常量，工厂函数，构造函数，赋值运算符，析构函数，其它成员函数，数据成员。

规则3.15.2 构造函数初始化列表放在同一行或按四格缩进并排多行

```

1. // 如果所有变量能放在同一行：
2. MyClass::MyClass(int var) : someVar(var) {
3.     DoSomething();

```

```
4.  }
5.
6.  // 如果不能放在同一行,
7.  // 必须置于冒号后, 并缩进4个空格
8.  MyClass::MyClass(int var)
9.      : someVar(var), someOtherVar(var + 1) { // Good: 逗号后面留有空格
10.     DoSomething();
11. }
12.
13. // 如果初始化列表需要置于多行, 需要逐行对齐
14. MyClass::MyClass(int var)
15.     : someVar(var),           // 缩进4个空格
16.       someOtherVar(var + 1) {
17.     DoSomething();
18. }
```

# 注释

一般的，尽量通过清晰的架构逻辑，好的符号命名来提高代码可读性；需要的时候，才辅以注释说明。注释是为了帮助阅读者快速读懂代码，所以要从读者的角度出发，按需注释。

注释内容要简洁、明了、无二义性，信息全面且不冗余。

注释跟代码一样重要。写注释时要换位思考，用注释去表达此时读者真正需要的信息。在代码的功能、意图层次上进行注释，即注释解释代码难以表达的意图，不要重复代码信息。修改代码时，也要保证其相关注释的一致性。只改代码，不改注释是一种不文明行为，破坏了代码与注释的一致性，让阅读者迷惑、费解，甚至误解。

## 注释风格

在 C++ 代码中，使用 `//` 和 `/* */` 都是可以的。按注释的目的和位置，注释可分为不同的类型，如文件头注释、函数头注释、代码注释等等；同一类型的注释应该保持统一的风格。

注意：本文示例代码中，大量使用 `'//'` 后置注释只是为了更精确的描述问题，并不代表这种注释风格更好。

## 文件头注释

### 规则4.2.1 文件头注释必须包含版权许可

```
1. /*
2.  * Copyright (c) [2019] [name of copyright holder]
3.  * [Software Name] is licensed under the Mulan PSL v1.
4.  * You can use this software according to the terms and conditions of the Mulan PSL v1.
5.  * You may obtain a copy of Mulan PSL v1 at:
6.  *   http://license.coscl.org.cn/MulanPSL
7.  * THIS SOFTWARE IS PROVIDED ON AN "AS IS" BASIS, WITHOUT WARRANTIES OF ANY KIND, EITHER EXPRESS OR
8.  * IMPLIED, INCLUDING BUT NOT LIMITED TO NON-INFRINGEMENT, MERCHANTABILITY OR FIT FOR A PARTICULAR
9.  * PURPOSE.
10. * See the Mulan PSL v1 for more details.
11. */
```

## 函数头注释

### 规则4.3.1 禁止空有格式的函数头注释

并不是所有的函数都需要函数头注释；函数签名无法表达的信息，加函数头注释辅助说明；

函数头注释统一放在函数声明或定义上方，使用如下风格之一：使用 `//` 写函数头

```
1. // 单行函数头
2. int Func1(void);
3.
4. // 多行函数头
```

```
5. // 第二行
6. int Func2(void);
```

## 使用 `//` 写函数头

```
1. /* 单行函数头 */
2. int Func1(void);
3.
4. /*
5.  * 另一种单行函数头
6.  */
7. int Func2(void);
8.
9. /*
10.  * 多行函数头
11.  * 第二行
12.  */
13. int Func3(void);
```

函数尽量通过函数名自注释，按需写函数头注释。不要写无用、信息冗余的函数头；不要写空有格式的函数头。

函数头注释内容可选，但不限于：功能说明、返回值，性能约束、用法、内存约定、算法实现、可重入的要求等等。模块对外头文件中的函数接口声明，其函数头注释，应当将重要、有用的信息表达清楚。

例：

```
1. /*
2.  * 返回实际写入的字节数，-1表示写入失败
3.  * 注意，内存 buf 由调用者负责释放
4.  */
5. int WriteString(const char *buf, int len);
```

坏的例子：

```
1. /*
2.  * 函数名:WriteString
3.  * 功能：写入字符串
4.  * 参数：
5.  * 返回值：
6.  */
7. int WriteString(const char *buf, int len);
```

上面例子中的问题：

- 参数、返回值，空有格式没内容
- 函数名信息冗余
- 关键的 buf 由谁释放没有说清楚

## 代码注释

## 规则4.4.1 代码注释放于对应代码的上方或右边

## 规则4.4.2 注释符与注释内容间要有1空格；右置注释与前面代码至少1空格

代码上方的注释，应该保持对应代码一样的缩进。选择并统一使用如下风格之一：使用 `//`

```
1.
2. // 这是单行注释
3. DoSomething();
4.
5. // 这是多行注释
6. // 第二行
7. DoSomething();
```

使用 `/* */`

```
1. /* 这是单行注释 */
2. DoSomething();
3.
4. /*
5.  * 另一种方式的多行注释
6.  * 第二行
7.  */
8. DoSomething();
```

代码右边的注释，与代码之间，至少留1空格，建议不超过4空格。通常使用扩展后的 `TAB` 键即可实现 1-4 空格的缩进。

选择并统一使用如下风格之一：

```
1. int foo = 100; // 放右边的注释
2. int bar = 200; /* 放右边的注释 */
```

右置格式在适当的时候，上下对齐会更美观。对齐后的注释，离左边代码最近的那一行，保证1-4空格的间隔。例：

```
1. const int kConst = 100;          /* 相关的同类注释，可以考虑上下对齐 */
2. const int kAnotherConst = 200;   /* 上下对齐时，与左侧代码保持间隔*/
```

当右置的注释超过行宽时，请考虑将注释置于代码上方。

## 规则4.4.3 不用的代码段直接删除，不要注释掉

被注释掉的代码，无法被正常维护；当企图恢复使用这段代码时，极有可能引入易被忽略的缺陷。正确的做法是，不需要的代码直接删除掉。若再需要时，考虑移植或重写这段代码。

这里说的注释掉代码，包括用 `//` 和 `/**/`，还包括 `#if 0`，`#ifndef NEVER_DEFINED` 等等。

## 建议4.4.1 正式交付给客户的代码不能包含 `TODO/TBD/FIXME` 注释

TODO/TBD 注释一般用来描述已知待改进、待补充的修改点FIXME 注释一般用来描述已知缺陷它们都应该有统一风格，方便文本搜索统一处理。如：

1. // TODO(<author-name>)：补充XX处理
2. // FIXME：XX缺陷



# 头文件

## 头文件职责

头文件是模块或文件的对外接口，头文件的设计体现了大部分的系统设计。头文件中适合放置接口的声明，不适合放置实现（内联函数除外）。对于cpp文件中内部才需要使用的函数、宏、枚举、结构定义等不要放在头文件中。头文件应当职责单一。头文件过于复杂，依赖过于复杂还是导致编译时间过长的主要原因。

### 建议5.1.1 每一个.cpp文件应有一个对应的.h文件，用于声明需要对外公开的类与接口

通常情况下，每个.cpp文件都有一个相应的.h，用于放置对外提供的函数声明、宏定义、类型定义等。如果一个.cpp文件不需要对外公布任何接口，则其就不应当存在。例外：程序的入口（如main函数所在的文件），单元测试代码，动态库代码。

示例：

```
1. // Foo.h
2.
3. #ifndef FOO_H
4. #define FOO_H
5.
6. class Foo {
7. public:
8.     Foo();
9.     void Fun();
10.
11. private:
12.     int value;
13. };
14.
15. #endif
```

```
1. // Foo.cpp
2. #include "Foo.h"
3.
4. namespace { // Good: 对内函数的声明放在.cpp文件的头部，并声明为匿名namespace或者static限制其作用域
5. void Bar()
6. {
7. }
8. }
9.
10. ...
11.
12. void Foo::Fun() {
13.     Bar();
14. }
```

## 头文件依赖

### 规则5.2.1 禁止头文件循环依赖

头文件循环依赖，指 a.h 包含 b.h, b.h 包含 c.h, c.h 包含 a.h, 导致任何一个头文件修改，都导致所有包含了a.h/b.h/c.h的代码全部重新编译一遍。而如果是单向依赖，如a.h包含b.h, b.h包含c.h, 而c.h不包含任何头文件，则修改a.h不会导致包含了b.h/c.h的源代码重新编译。

头文件循环依赖直接体现了架构设计上的不合理，可通过优化架构去避免。

### 规则5.2.2 禁止包含用不到的头文件

用不到的头文件被包含的同时引入了不必要的依赖，增加了模块或单元之间的耦合度，只要该头文件被修改，代码就要重新编译。

很多系统中头文件包含关系复杂，开发人员为了省事起见，直接包含一切想到的头文件，甚至发布了一个god.h，其中包含了所有头文件，然后发布给各个项目组使用，这种只图一时省事的做法，导致整个系统的编译时间进一步恶化，并对后来人的维护造成了巨大的麻烦。

### 规则5.2.3 头文件应当自包含

简单的说，自包含就是任意一个头文件均可独立编译。如果一个文件包含某个头文件，还要包含另外一个头文件才能工作的话，给这个头文件的用户增添不必要的负担。

示例：如果a.h不是自包含的，需要包含b.h才能编译，会带来的危害：每个使用a.h头文件的.cpp文件，为了让引入的a.h的内容编译通过，都要包含额外的头文件b.h。额外的头文件b.h必须在a.h之前进行包含，这在包含顺序上产生了依赖。

### 规则5.2.4 头文件必须编写#define保护，防止重复包含

为防止头文件被重复包含，所有头文件都应当使用 `#define` 保护；不要使用 `#pragma once`

定义包含保护符时，应该遵守如下规则：1) 保护符使用唯一名称；2) 不要在受保护部分的前后放置代码或者注释，文件头注释除外。

示例：假定VOS工程的timer模块的timer.h，其目录为VOS/include/timer/Timer.h, 应按如下方式保护：

```
1. #ifndef VOS_INCLUDE_TIMER_TIMER_H
2. #define VOS_INCLUDE_TIMER_TIMER_H
3. ...
4. #endif
```

也可以不用像上面添加路径，但是要保证当前工程内宏是唯一的。

```
1. #ifndef TIMER_H
2. #define TIMER_H
3. ...
4. #endif
```

## 建议5.2.1 禁止通过声明的方式引用外部函数接口、变量

只能通过包含头文件的方式使用其他模块或文件提供的接口。通过 `extern` 声明的方式使用外部函数接口、变量，容易在外部接口改变时可能导致声明和定义不一致。同时这种隐式依赖，容易导致架构腐化。

不符合规范的案例：

// a.cpp内容

```
1. extern int Fun();    // Bad: 通过extern的方式使用外部函数
2.
3. void Bar() {
4.     int i = Fun();
5.     ...
6. }
```

// b.cpp内容

```
1. int Fun() {
2.     // Do something
3. }
```

应该改为：

// a.cpp内容

```
1. #include "b.h"    // Good: 通过包含头文件的方式使用其他.cpp提供的接口
2.
3. void Bar() {
4.     int i = Fun();
5.     ...
6. }
```

// b.h内容

```
1. int Fun();
```

// b.cpp内容

```
1. int Fun() {
2.     // Do something
3. }
```

例外，有些场景需要引用其内部函数，但并不想侵入代码时，可以 `extern` 声明方式引用。如：针对某一内部函数进行单元测试时，可以通过 `extern` 声明来引用被测函数；当需要对某一函数进行打桩、打补丁处理时，允许 `extern` 声明该函数。

## 规则5.2.5 禁止在extern "C"中包含头文件

在 `extern "C"` 中包含头文件，有可能会导导致 `extern "C"` 嵌套，部分编译器对 `extern "C"` 嵌套层次有限制，嵌套层次太多会编译错误。

在C, C++混合编程的情况下，在`extern "C"`中包含头文件，可能会导致被包含头文件的原有意图遭到破坏，比如链接规范被不正确地更改。

示例，存在`a.h`和`b.h`两个头文件：

// a.h内容

```
1. ...
2. #ifdef __cplusplus
3. void Foo(int);
4. #define A(value) Foo(value)
5. #else
6. void A(int)
7. #endif
```

// b.h内容

```
1. ...
2. #ifdef __cplusplus
3. extern "C" {
4. #endif
5.
6. #include "a.h"
7. void B();
8.
9. #ifdef __cplusplus
10. }
11. #endif
```

使用C++预处理器展开`b.h`，将会得到

```
1. extern "C" {
2.     void Foo(int);
3.     void B();
4. }
```

按照 `a.h` 作者的本意，函数 `Foo` 是一个 C++ 自由函数，其链接规范为 `"C++"`。但在 `b.h` 中，由于 `#include "a.h"` 被放到了 `extern "C"` 的内部，函数 `Foo` 的链接规范被不正确地更改了。

例外：如果在 C++ 编译环境中，想引用纯C的头文件，这些C头文件并没有 `extern "C"` 修饰。非侵入式的做法是，在 `extern "C"` 中去包含C头文件。

## 建议5.2.2尽量避免使用前置声明，而是通过`#include`来包含头文件

前置声明 (forward declaration) 是类、函数和模板的纯粹声明，没伴随着其定义。

- 优点：

- 前置声明能够节省编译时间，多余的 `#include` 会迫使编译器展开更多的文件，处理更多的输入。
- 前置声明能够节省不必要的重新编译的时间。 `#include` 使代码因为头文件中无关的改动而被重新编译多次。
- 缺点：
  - 前置声明隐藏了依赖关系，头文件改动时，用户的代码会跳过必要的重新编译过程。
  - 前置声明可能会被库的后续更改所破坏。前置声明函数或模板有时会妨碍头文件开发者变动其 API。例如扩大形参类型，加个自带默认参数的模板形参等等。
  - 前置声明来自命名空间 `std::` 的 symbol 时，其行为未定义（在C++11标准规范中明确说明）。
  - 前置声明了不少来自头文件的 symbol 时，就会比单单一行的 `include` 冗长。
  - 仅仅为了能前置声明而重构代码（比如用指针成员代替对象成员）会使代码变得更慢更复杂。
  - 很难判断什么时候该用前置声明，什么时候该用 `#include`，某些场景下面前置声明和 `#include` 互换以后会导致意想不到的结果。所以我们尽可能避免使用前置声明，而是使用 `#include` 头文件来保证依赖关系。

### 建议5.2.3 头文件包含顺序：首先是.cpp相应的.h文件，其它头文件按照稳定度排序

使用标准的头文件包含顺序可增强可读性，避免隐藏依赖，建议按照稳定度排序：cpp对应的头文件，C/C++标准库，系统库的.h，其他库的.h，本项目内其他的.h。

举例，Foo.cpp中包含头文件的次序如下：

```
1. #include "Foo/Foo.h"
2.
3. #include <cstdlib>
4. #include <string>
5.
6. #include <linux/list.h>
7. #include <linux/time.h>
8.
9. #include "platform/Base.h"
10. #include "platform/Framework.h"
11.
12. #include "project/public/Log.h"
```

将Foo.h放在最前面可以保证当Foo.h遗漏某些必要的库，或者有错误时，Foo.cpp的构建会立刻中止，减少编译时间。对于头文件中包含顺序也参照此建议。

例外：平台特定代码需要条件编译，这些代码可以放到其它 `includes` 之后。

```
1. #include "foo/public/FooServer.h"
2.
3. #include "base/Port.h" // For LANG_CXX11.
4.
5. #ifdef LANG_CXX11
6. #include <initializer_list>
7. #endif // LANG_CXX11
```

# 作用域

## 命名空间

命名空间里的内容不缩进。

### 建议 6.1.1 对于 cpp 文件中不需要导出的变量，常量或者函数，请使用匿名 namespace 封装或者用 static 修饰

在 C++ 2003 标准规范中，使用 static 修饰文件作用域的变量，函数等被标记为 deprecated 特性，所以更推荐使用匿名 namespace。

主要原因如下：

- static 在 C++中已经赋予了太多的含义，静态函数成员变量，静态成员函数，静态全局变量，静态函数局部变量，每一种都有特殊的处理。
- static 只能保证变量，常量和函数的文件作用域，但是 namespace 还可以封装类型等。
- 统一 namespace 来处理 C++的作用域，而不需要同时使用 static 和 namespace 来管理。
- static 修饰的函数不能用来实例化模板，而匿名 namespace 可以。但是不要在 .h 中使用中使用匿名 namespace 或者 static。

```
1. // Foo.cpp
2.
3. namespace {
4.     const int kMaxCount = 20;
5.     void InternalFun(){};
6. }
7.
8. void Foo::Fun() {
9.     int i = kMaxCount;
10.
11.     InternalFun();
12. }
```

### 规则 6.1.1 不要在头文件中或者#include 之前使用 using 导入命名空间

说明：使用 using 导入命名空间会影响后续代码，易造成符号冲突，所以不要在头文件以及源文件中的#include 之前使用 using 导入命名空间。示例：

```
1. // 头文件a.h
2. namespace namespacea {
3.     int Fun(int);
4. }
```

```
1. // 头文件b.h
2. namespace namespaceb {
```

```

3. int Fun(int);
4. }
5.
6. using namespace namespaceb;
7.
8. void G() {
9.     Fun(1);
10. }

```

```

1. // 源代码a.cpp
2. #include "a.h"
3. using namespace namespacea;
4. #include "b.h"
5.
6. void main() {
7.     G(); // using namespace namespacea在#include "b.h"之前, 引发歧义: namespacea::Fun, namespaceb::Fun调用不明确
8. }

```

对于在头文件中使用 `using` 导入单个符号或定义别名, 允许在模块自定义名字空间中使用, 但禁止在全局名字空间中使用。

```

1. // foo.h
2.
3. #include <fancy/string>
4. using fancy::string; // Bad, 禁止向全局名字空间导入符号
5.
6. namespace foo {
7. using fancy::string; // Good, 可以在模块自定义名字空间中导入符号
8. using MyVector = fancy::vector<int>; // Good, C++11可在自定义名字空间中定义别名
9. }

```

## 全局函数和静态成员函数

**建议 6.2.1** 优先使用命名空间来管理全局函数, 如果和某个 `class` 有直接关系的, 可以使用静态成员函数

说明: 非成员函数放在名字空间内可避免污染全局作用域, 也不要类+静态成员方法来简单管理全局函数。如果某个全局函数和某个类有紧密联系, 那么可以作为类的静态成员函数。

如果你需要定义一些全局函数, 给某个 `cpp` 文件使用, 那么请使用匿名 `namespace` 来管理。

```

1. namespace mynamespace {
2. int Add(int a, int b);
3. }
4.
5. class File {
6. public:
7.     static File CreateTempFile(const std::string& fileName);
8. };

```

## 全局常量和静态成员常量

**建议 6.3.1** 优先使用命名空间来管理全局常量，如果和某个 `class` 有直接关系的，可以使用静态成员常量

说明：全局常量放在命名空间内可避免污染全局作用域， 也不要再用类+静态成员常量来简单管理全局常量。 如果某个全局常量和某个类有紧密联系， 那么可以作为类的静态成员常量。

如果你需要定义一些全局常量，只给某个 `cpp` 文件使用，那么请使用匿名 `namespace` 来管理。

```
1. namespace mynamespace {
2.     const int kMaxSize = 100;
3. }
4.
5. class File {
6.     public:
7.         static const std::string kName;
8. };
```

## 全局变量

**建议 6.4.1** 尽量避免使用全局变量，考虑使用单例模式

说明：全局变量是可以修改和读取的，那么这样会导致业务代码和这个全局变量产生数据耦合。

```
1. int counter = 0;
2.
3. // a.cpp
4. counter++;
5.
6. // b.cpp
7. counter++;
8.
9. // c.cpp
10. cout << counter << endl;
```

使用单实例模式

```
1. class Counter {
2.     public:
3.         static Counter& GetInstance() {
4.             static Counter counter;
5.             return counter;
6.         } // 单实例实现简单举例
7.
8.         void Increase() {
9.             value++;
```



```
10.     }
11.
12.     void Print() const {
13.         std::cout << value << std::endl;
14.     }
15.
16. private:
17.     Counter() : value(0) {}
18.
19. private:
20.     int value;
21. };
22.
23. // a.cpp
24. Counter::GetInstance().Increase();
25.
26. // b.cpp
27. Counter::GetInstance().Increase();
28.
29. // c.cpp
30. Counter::GetInstance().Print();
```

实现单例模式以后，实现了全局唯一一个实例，和全局变量同样的效果，并且单实例提供了更好的封装性。

例外：有的时候全局变量的作用域仅仅是模块内部，这样进程空间里面就会有多个全局变量实例，每个模块持有一份，这种场景下是无法使用单例模式解决的。

# 类

如果仅有数据成员，使用结构体，其他使用类

## 构造，拷贝构造，赋值和析构函数

构造，拷贝，移动和析构函数提供了对象的生命周期管理方法：

- 构造函数 (constructor) : `X()`
- 拷贝构造函数 (copy constructor) : `X(const X&)`
- 拷贝赋值操作符 (copy assignment) : `operator=(const X&)`
- 移动构造函数 (move constructor) : `X(X&&)` C++11以后提供
- 移动赋值操作符 (move assignment) : `operator=(X&&)` C++11以后提供
- 析构函数 (destructor) : `~X()`

### 规则7.1.1 类的成员变量必须显式初始化

说明：如果类有成员变量，没有定义构造函数，又没有定义默认构造函数，编译器将自动生成一个构造函数，但编译器生成的构造函数并不会对成员变量进行初始化，对象状态处于一种不确定性。

例外：

- 如果类的成员变量具有默认构造函数，那么可以不需要显式初始化。示例：如下代码没有构造函数，私有数据成员无法初始化：

```
1. class Message {
2.     public:
3.         void ProcessOutMsg() {
4.             //...
5.         }
6.     private:
7.         unsigned int msgID;
8.         unsigned int msgLength;
9.         unsigned char* msgBuffer;
10.        std::string someIdentifier;
11. };
12.
13. Message message;    // message成员变量没有初始化
14. message.ProcessOutMsg();    // 后续使用存在隐患
15.
16. // 因此，有必要定义默认构造函数，如下：
17. class Message {
18.     public:
19.         Message() : msgID(0), msgLength(0) {
20.         }
21.
22.         void ProcessOutMsg() {
23.             // ...
24.         }
```

```

25.
26. private:
27.     unsigned int msgID;
28.     unsigned int msgLength;
29.     unsigned char* msgBuffer;
30.     std::string someIdentifier; //具有默认构造函数，不需要显式初始化
31. };

```

## 建议7.1.1 成员变量优先使用声明时初始化（C++11）和构造函数初始化列表初始化

说明：C++11的声明时初始化可以一目了然的看出成员初始值，应当优先使用。如果成员初始化值和构造函数相关，或者不支持C++11，则应当优先使用构造函数初始化列表来初始化成员。相比起在构造函数体中对成员赋值，初始化列表的代码更简洁，执行性能更好，而且可以对const成员和引用成员初始化。

```

1. class Message {
2. public:
3.     Message() : msgLength(0) { // Good, 优先使用初始化列表
4.         msgBuffer = NULL;     // Bad, 不推荐在构造函数中赋值
5.     }
6.
7. private:
8.     unsigned int msgID{0}; // Good, C++11中使用
9.     unsigned int msgLength;
10.    unsigned char* msgBuffer;
11. };

```

## 规则7.1.2 为避免隐式转换，将单参数构造函数声明为explicit

说明：单参数构造函数如果没有用explicit声明，则会成为隐式转换函数。示例：

```

1. class Foo {
2. public:
3.     explicit Foo(const string& name): name(name) {
4.     }
5. private:
6.     string name;
7. };
8.
9.
10. void ProcessFoo(const Foo& foo){}
11.
12. int main(void) {
13.     std::string test = "test";
14.     ProcessFoo(test); // 编译不通过
15.     return 0;
16. }

```

上面的代码编译不通过，因为 `ProcessFoo` 需要的参数是Foo类型，传入的string类型不匹配。

如果将Foo构造函数的explicit关键字移除，那么调用 `ProcessFoo` 传入的string就会触发隐式转换，生成一个临时的Foo对象。往往这种隐式转换是让人迷惑的，并且容易隐藏Bug，得到了一个不期望的类型转换。所以对于单参数的构造函数是要求explicit声明。

### 规则7.1.3 如果不需要拷贝构造函数、赋值操作符 / 移动构造函数、赋值操作符，请明确禁止

说明：如果用户不定义，编译器默认会生成拷贝构造函数和拷贝赋值操作符，移动构造和移动赋值操作符（移动语义的函数C++11以后才有）。如果我们不要使用拷贝构造函数，或者赋值操作符，请明确拒绝：

- 将拷贝构造函数或者赋值操作符设置为private，并且不实现：

```
1. class Foo {
2.     private:
3.         Foo(const Foo&);
4.         Foo& operator=(const Foo&);
5. };
```

- 使用C++11提供的delete，请参见后面现代C++的相关章节。

### 规则7.1.4 拷贝构造和拷贝赋值操作符应该是成对出现或者禁止

拷贝构造函数和拷贝赋值操作符都是具有拷贝语义的，应该同时出现或者禁止。

```
1. // 同时出现
2. class Foo {
3.     public:
4.         ...
5.         Foo(const Foo&);
6.         Foo& operator=(const Foo&);
7.         ...
8. };
9.
10. // 同时default, C++11支持
11. class Foo {
12.     public:
13.         Foo(const Foo&) = default;
14.         Foo& operator=(const Foo&) = default;
15. };
16.
17. // 同时禁止, C++11可以使用delete
18. class Foo {
19.     private:
20.         Foo(const Foo&);
21.         Foo& operator=(const Foo&);
22. };

```

### 规则7.1.5 移动构造和移动赋值操作符应该是成对出现或者禁止

在C++11中增加了move操作，如果需要某个类支持移动操作，那么需要实现移动构造和移动赋值操作符。

移动构造函数和移动赋值操作符都是具有移动语义的，应该同时出现或者禁止。

```

1. // 同时出现
2. class Foo {
3.     public:
4.         ...
5.         Foo(Foo&&);
6.         Foo& operator=(Foo&&);
7.         ...
8. };
9.
10. // 同时default, C++11支持
11. class Foo {
12.     public:
13.         Foo(Foo&&) = default;
14.         Foo& operator=(Foo&&) = default;
15. };
16.
17. // 同时禁止, 使用C++11的delete
18. class Foo {
19.     public:
20.         Foo(Foo&&) = delete;
21.         Foo& operator=(Foo&&) = delete;
22. };

```

## 规则7.1.6 禁止在构造函数和析构函数中调用虚函数

说明：在构造函数和析构函数中调用当前对象的虚函数，会导致未实现多态的行为。在C++中，一个基类一次只构造一个完整的对象。

示例：类Base是基类，Sub是派生类

```

1. class Base {
2.     public:
3.         Base();
4.         virtual void Log() = 0;    // 不同的派生类调用不同的日志文件
5. };
6.
7. Base::Base() {                    // 基类构造函数
8.     Log();                        // 调用虚函数Log
9. }
10.
11. class Sub : public Base {
12.     public:
13.         virtual void Log();
14. };

```

当执行如下语句：`Sub sub;` 会先执行Sub的构造函数，但首先调用Base的构造函数，由于Base的构造函数调用虚函数Log，此时Log还是基类的版本，只有基类构造完成后，才会完成派生类的构造，从而导致未实现多态的行为。同样

的道理也适用于析构函数。

## 继承

### 规则7.2.1 基类的析构函数应该声明为virtual

说明：只有基类析构函数是virtual，通过多态调用的时候才能保证派生类的析构函数被调用。

示例：基类的析构函数没有声明为virtual导致了内存泄漏。

```
1. class Base {
2.     public:
3.         virtual std::string getVersion() = 0;
4.
5.     ~Base() {
6.         std::cout << "~Base" << std::endl;
7.     }
8. };
```

```
1. class Sub : public Base {
2.     public:
3.         Sub() : numbers(NULL) {
4.         }
5.
6.         ~Sub() {
7.             delete[] numbers;
8.             std::cout << "~Sub" << std::endl;
9.         }
10.
11.         int Init() {
12.             const size_t numberCount = 100;
13.             numbers = new (std::nothrow) int[numberCount];
14.             if (numbers == NULL) {
15.                 return -1;
16.             }
17.
18.             ...
19.         }
20.
21.         std::string getVersion() {
22.             return std::string("hello!");
23.         }
24.     private:
25.         int* numbers;
26. };
```

```
1. int main(int argc, char* args[]) {
2.     Base* b = new Sub();
3.
4.     delete b;
```

```
5.     return 0;
6. }
```

由于基类Base的析构函数没有声明为virtual，当对象被销毁时，只会调用基类的析构函数，不会调用派生类Sub的析构函数，导致内存泄漏。

## 规则7.2.2 禁止虚函数使用缺省参数值

说明：在C++中，虚函数是动态绑定的，但函数的缺省参数却是在编译时就静态绑定的。这意味着你最终执行的函数是一个定义在派生类，但使用了基类中的缺省参数值的虚函数。为了避免虚函数重载时，因参数声明不一致给使用者带来的困惑和由此导致的问题，规定所有虚函数均不允许声明缺省参数值。示例：虚函数display缺省参数值text是由编译时刻决定的，而非运行时刻，没有达到多态的目的：

```
1. class Base {
2. public:
3.     virtual void Display(const std::string& text = "Base!") {
4.         std::cout << text << std::endl;
5.     }
6.
7.     virtual ~Base(){}
8. };
9.
10. class Sub : public Base {
11. public:
12.     virtual void Display(const std::string& text = "Sub!") {
13.         std::cout << text << std::endl;
14.     }
15.
16.     virtual ~Sub(){}
17. };
18.
19. int main() {
20.     Base* base = new Sub();
21.     Sub* sub = new Sub();
22.
23.     ...
24.
25.     base->Display(); // 程序输出结果：Base! 而期望输出：Sub!
26.     sub->Display(); // 程序输出结果：Sub!
27.
28.     delete base;
29.     delete sub;
30.     return 0;
31. };
```

## 规则7.2.3 禁止重新定义继承而来的非虚函数

说明：因为非虚函数无法实现动态绑定，只有虚函数才能实现动态绑定：只要操作基类的指针，即可获得正确的结果。

示例：

```

1. class Base {
2.     public:
3.         void Fun();
4. };
5.
6. class Sub : public Base {
7.     public:
8.         void Fun();
9. };
10.
11. Sub* sub = new Sub();
12. Base* base = sub;
13.
14. sub->Fun();    // 调用子类的Fun
15. base->Fun();  // 调用父类的Fun
16. //...
```

## 多重继承

在实际开发过程中使用多重继承的场景是比较少的，因为多重继承使用过程中有下面的典型问题：

- 菱形继承所带来的数据重复，以及名字二义性。因此，C++引入了virtual继承来解决这类问题；
- 即便不是菱形继承，多个父类之间的名字也可能存在冲突，从而导致的二义性；
- 如果子类需要扩展或改写多个父类的方法时，造成子类的职责不明，语义混乱；
- 相对于委托，继承是一种白盒复用，即子类可以访问父类的protected成员，这会导致更强的耦合。而多重继承，由于耦合了多个父类，相对于单根继承，这会产生更强的耦合关系。多重继承具有下面的优点：多重继承提供了一种更简单的组合来实现多种接口或者类的组装与复用。

所以，对于多重继承的只有下面几种情况下面才允许使用多重继承。

### 建议7.3.1 使用多重继承来实现接口分离与多角色组合

如果某个类需要实现多重接口，可以通过多重继承把多个分离的接口组合起来，类似 scala 语言的 traits 混入。

```

1. class Role1 {};
2. class Role2 {};
3. class Role3 {};
4.
5. class Object1 : public Role1, public Role2 {
6.     // ...
7. };
8.
9. class Object2 : public Role2, public Role3 {
10.    // ...
11.};
```



在C++标准库中也有类似的实现样例：

```
1. class basic_istream {};  
2. class basic_ostream {};  
3.  
4. class basic_iostream : public basic_istream, public basic_ostream {  
5.  
6. };
```

## 重载

重载操作符要有充分理由, 而且不要改变操作符原有语义, 例如不要使用 ‘+’ 操作符来做减运算。操作符重载令代码更加直观, 但也有一些不足:

- 混淆直觉, 误以为该操作和内建类型一样是高性能的, 忽略了性能降低的可能;
- 问题定位时不够直观, 按函数名查找比按操作符显然更方便。
- 重载操作符如果行为定义不直观 (例如将 ‘+’ 操作符来做减运算), 会让代码产生混淆。
- 赋值操作符的重载引入的隐式转换会隐藏很深的bug。可以定义类似Equals()、CopyFrom()等函数来替代 =, == 操作符。

# 函数

## 函数设计

### 建议8.1.1 避免函数过长，函数不超过50行（非空非注释）

函数应该可以一屏显示完（50行以内），只做一件事情，而且把它做好。

过长的函数往往意味着函数功能不单一，过于复杂，或过分呈现细节，未进行进一步抽象。

例外：某些实现算法的函数，由于算法的聚合性与功能的全面性，可能会超过50行。

即使一个长函数现在工作的非常好，一旦有人对其修改，有可能出现新的问题，甚至导致难以发现的bug。建议将其拆分为更加简短并易于管理的若干函数，以便于他人阅读和修改代码。

## 内联函数

### 建议8.2.1 内联函数不超过10行（非空非注释）

说明：内联函数具有一般函数的特性，它与一般函数不同之处只在于函数调用的处理。一般函数进行调用时，要将程序执行权转到被调用函数中，然后再返回到调用它的函数中；而内联函数在调用时，是将调用表达式用内联函数体来替换。

内联函数只适合于只有 1~10 行的小函数。对一个含有许多语句的大函数，函数调用和返回的开销相对来说微不足道，也没有必要用内联函数实现，一般的编译器会放弃内联方式，而采用普通的方式调用函数。

如果内联函数包含复杂的控制结构，如循环、分支(switch)、try-catch 等语句，一般编译器将该函数视同普通函数。虚函数、递归函数不能被用来做内联函数。

## 函数参数

### 建议8.3.1 函数参数使用引用取代指针

说明：引用比指针更安全，因为它一定非空，且一定不会再指向其他目标；引用不需要检查非法的NULL指针。

选择 const 避免参数被修改，让代码阅读者清晰地知道该参数不被修改，可大大增强代码可读性。

### 建议8.3.2 使用强类型参数，避免使用void\*

尽管不同的语言对待强类型和弱类型有自己的观点，但是一般认为c/c++是强类型语言，既然我们使用的语言是强类型的，就应该保持这样的风格。好处是尽量让编译器在编译阶段就检查出类型不匹配的问题。

使用强类型便于编译器帮我们发现错误，如下代码中注意函数 FooListAddNode 的使用：

```
1. struct FooNode {  
2.     struct List link;
```

```
3.  int foo;
4.  };
5.
6.  struct BarNode {
7.      struct List link;
8.      int bar;
9.  }
10.
11. void FooListAddNode(void *node) { // Bad: 这里用 void * 类型传递参数
12.     FooNode *foo = (FooNode *)node;
13.     ListAppend(&fooList, &foo->link);
14. }
15.
16. void MakeTheList() {
17.     FooNode *foo = nullptr;
18.     BarNode *bar = nullptr;
19.     ...
20.
21.     FooListAddNode(bar);           // Wrong: 这里本意是想传递参数 foo, 但错传了 bar, 却没有报错
22. }
```

- 可以使用模板函数来实现参数类型的变化。
- 可以使用基类指针来实现多态。

### 建议8.3.3 函数的参数个数不超过5个

函数的参数过多，会使得该函数易于受外部变化的影响，从而影响维护工作。函数的参数过多同时也会增大测试的工作量。

如果超过可以考虑：

- 看能否拆分函数
- 看能否将相关参数合在一起，定义结构体

## C++ 其他特性

### 常量与初始化

不变的值更易于理解、跟踪和分析，所以应该尽可能地使用常量代替变量，定义值的时候，应该把const作为默认的选项。

#### 建议9.1.1 不允许使用宏来表示常量

说明：宏是简单的文本替换，在预处理阶段时完成，运行报错时直接报相应的值；跟踪调试时也是显示值，而不是宏名；宏没有类型检查，不安全；宏没有作用域。

```
1. #define MAX_MSISDN_LEN 20    // 不好
2.
3. // C++请使用const常量
4. const int kMaxMsisdnLen = 20; // 好
5.
6. // 对于C++11以上版本，可以使用constexpr
7. constexpr int kMaxMsisdnLen = 20;
```

#### 建议9.1.2 一组相关的整型常量应定义为枚举

说明：枚举比 `#define` 或 `const int` 更安全。编译器会检查参数值是否位于枚举取值范围内，避免错误发生。

```
1. // 好的例子：
2. enum Week {
3.     kSunday,
4.     kMonday,
5.     kTuesday,
6.     kWednesday,
7.     kThursday,
8.     kFriday,
9.     kSaturday
10. };
11.
12. enum Color {
13.     kRed,
14.     kBlack,
15.     kBlue
16. };
17.
18. void ColorizeCalendar(Week today, Color color);
19.
20. ColorizeCalendar(kBlue, kSunday); // 编译报错，参数类型错误
21.
22. // 不好的例子：
23. const int kSunday = 0;
24. const int kMonday = 1;
```

```

25.
26. const int kRed = 0;
27. const int kBlack = 1;
28.
29. bool ColorizeCalendar(int today, int color);
30. ColorizeCalendar(kBlue, kSunday); // 不会报错

```

当枚举值需要对应到具体数值时，须在声明时显式赋值。否则不需要显式赋值，以避免重复赋值，降低维护(增加、删除成员)工作量。

```

1. // 好的例子：S协议里定义的设备ID值，用于标识设备类型
2. enum DeviceType {
3.     kUnknown = -1,
4.     kDsmP = 0,
5.     kIsmg = 1,
6.     kWapportal = 2
7. };

```

### 建议9.1.3 不允许使用魔鬼数字

所谓魔鬼数字即看不懂、难以理解的数字。

魔鬼数字并非一个非黑即白的概念，看不懂也有程度，需要自行判断。例如数字 12，在不同的上下文中情况是不一样的：`type = 12;` 就看不懂，但 `month = year * 12` ；就能看懂。数字 0 有时候也是魔鬼数字，比如 `status = 0` ；并不能表达是什么状态。

解决途径：对于局部使用的数字，可以增加注释说明对于多处使用的数字，必须定义 `const` 常量，并通过符号命名自注释。

禁止出现下列情况：没有通过符号来解释数字含义，如 `const int kZero = 0` 符号命名限制了其取值，如 `const int kXxTimerInterval = 300` ，直接使用 `kXxTimerInterval` 来表示该常量是定时器的时间间隔。

### 规则9.1.1 常量应该保证单一职责

说明：一个常量只用来表示一个特定功能，即一个常量不能有多种用途。

```

1. // 好的例子：协议A和协议B，手机号(MSISDN)的长度都是20。
2. const unsigned int kAMaxMsisdnLen = 20;
3. const unsigned int kBMaxMsisdnLen = 20;
4.
5. // 或者使用不同的名字空间：
6. namespace namespace1 {
7.     const unsigned int kMaxMsisdnLen = 20;
8. }
9.
10. namespace namespace2 {
11.     const unsigned int kMaxMsisdnLen = 20;
12. }

```

## 建议9.1.4 禁止用memcpy\_s、memset\_s初始化非POD对象

说明：`POD` 全称是 `Plain Old Data`，是C++ 98标准(ISO/IEC 14882, first edition, 1998-09-01)中引入的一个概念，`POD` 类型主要包括 `int`，`char`，`float`，`double`，`enumeration`，`void`，指针等原始类型以及聚合类型，不能使用封装和面向对象特性（如用户定义的构造/赋值/析构函数、基类、虚函数等）。

由于非POD类型比如非聚合类型的class对象，可能存在虚函数，内存布局不确定，跟编译器有关，滥用内存拷贝可能会导致严重的问题。

即使对聚合类型的class，使用直接的内存拷贝和比较，破坏了信息隐蔽和数据保护的作用，也不提倡 `memcpy_s`、`memset_s` 操作。

对于POD类型的详细说明请参见附录。

## 建议9.1.5 变量使用时才声明并初始化

说明：变量在使用前未赋初值，是常见的低级编程错误。使用前才声明变量并同时初始化，非常方便地避免了此类低级错误。

在函数开始位置声明所有变量，后面才使用变量，作用域覆盖整个函数实现，容易导致如下问题：

- 程序难以理解和维护：变量的定义与使用分离。
- 变量难以合理初始化：在函数开始时，经常没有足够的信息进行变量初始化，往往用某个默认的空值(比如零)来初始化，这通常是一种浪费，如果变量在被赋予有效值以前使用，还会导致错误。遵循变量作用域最小化原则与就近声明原则，使得代码更容易阅读，方便了解变量的类型和初始值。特别是，应使用初始化的方式替代声明再赋值。

```
1. // 不好的例子：声明与初始化分离
2. string name;           // 声明时未初始化：调用缺省构造函数
3. name = "zhangsan";    // 再次调用赋值操作符函数；声明与定义在不同的地方，理解相对困难
4.
5. // 好的例子：声明与初始化一体，理解相对容易
6. string name("zhangsan"); // 调用构造函数
```

## 表达式

### 规则9.2.1 含有变量自增或自减运算的表达式中禁止再次引用该变量

含有变量自增或自减运算的表达式中，如果再引用该变量，其结果在C++标准中未明确定义。各个编译器或者同一个编译器不同版本实现可能会不一致。为了更好的可移植性，不应该对标准未定义的运算次序做任何假设。

注意，运算次序的问题不能使用括号来解决，因为这不是优先级的問題。

示例：

```
1. x = b[i] + i++; // Bad: b[i]运算跟 i++, 先后顺序并不明确。
```

正确的写法是将自增或自减运算单独放一行：

```
1. x = b[i] + i;
2. i++;           // Good: 单独一行
```

## 函数参数

```
1. Func(i++, i);    // Bad: 传递第2个参数时, 不确定自增运算有没有发生
```

## 正确的写法

```
1. i++;           // Good: 单独一行
2. x = Func(i, i);
```

## 规则9.2.2 switch语句要有default分支

大部分情况下，switch语句中要有default分支，保证在遗漏case标签处理时能够有一个缺省的处理行为。

特例：如果switch条件变量是枚举类型，并且 case 分支覆盖了所有取值，则加上default分支处理有些多余。现代编译器都具备检查是否在switch语句中遗漏了某些枚举值的case分支的能力，会有相应的warning提示。

```
1. enum Color {
2.     kRed = 0,
3.     kBlue
4. };
5.
6. // 因为switch条件变量是枚举值, 这里可以不用加default处理分支
7. switch (color) {
8.     case kRed:
9.         DoRedThing();
10.        break;
11.    case kBlue:
12.        DoBlueThing();
13.        ...
14.        break;
15. }
```

## 建议9.2.1 表达式的比较，应当遵循左侧倾向于变化、右侧倾向于不变的原则

当变量与常量比较时，如果常量放左边，如 `if (MAX == v)` 不符合阅读习惯，而 `if (MAX > v)` 更是难于理解。应当按人的正常阅读、表达习惯，将常量放右边。写成如下方式：

```
1. if (value == MAX) {
2.
3. }
4.
5. if (value < MAX) {
6.
7. }
```

也有特殊情况，如：`if (MIN < value && value < MAX)` 用来描述区间时，前半段是常量在左的。

不用担心将 `'=='` 误写成 `'=`'，因为 `if (value = MAX)` 会有编译告警，其他静态检查工具也会报错。让工具去解决笔误问题，代码要符合可读性第一。

## 建议9.2.2 使用括号明确操作符的优先级

使用括号明确操作符的优先级，防止因默认的优先级与设计思想不符而导致程序出错；同时使得代码更为清晰可读，然而过多的括号会分散代码使其降低了可读性。下面是如何使用括号的建议。

- 二元及以上操作符，如果涉及多种操作符，则应该使用括号

```
1. x = a + b + c;          /* 操作符相同，可以不加括号 */
2. x = Foo(a + b, c);      /* 逗号两边的表达式，不需要括号 */
3. x = 1 << (2 + 3);       /* 操作符不同，需要括号 */
4. x = a + (b / 5);        /* 操作符不同，需要括号 */
5. x = (a == b) ? a : (a - b); /* 操作符不同，需要括号 */
```

## 类型转换

避免使用类型分支来定制行为：类型分支来定制行为容易出错，是企图用C++编写C代码的明显标志。这是一种很不灵活的技术，要添加新类型时，如果忘记修改所有分支，编译器也不会告知。使用模板和虚函数，让类型自己而不是调用它们的代码来决定行为。

建议避免类型转换，我们在代码的类型设计上应该考虑到每种数据的数据类型是什么，而不是应该过度使用类型转换来解决问题。在设计某个基本类型的时候，请考虑：

- 是无符号还是有符号的
- 是适合float还是double
- 是使用int8, int16, int32还是int64，确定整形的长度但是我们无法禁止使用类型转换，因为C++语言是一门面向机器编程的语言，涉及到指针地址，并且我们会与各种第三方或者底层API交互，他们的类型设计不一定是合理的，在这个适配的过程中很容易出现类型转换。

例外：在调用某个函数的时候，如果我们不想处理函数结果，首先要考虑这个是否是你的最好的选择。如果确实不想处理函数的返回值，那么可以使用(void)转换来解决。

## 规则9.3.1 如果确定要使用类型转换，请使用有C++提供的类型转换，而不是C风格的类型转换

说明：

C++提供的类型转换操作比C风格更有针对性，更易读，也更加安全，C++提供的转换有：

- 类型转换：
- `dynamic_cast`：主要用于继承体系下行转换，`dynamic_cast` 具有类型检查的功能，请做好基类和派生类的设计，避免使用dynamic\_cast来进行转换。
- `static_cast`：和C风格转换相似可做值的强制转换，或上行转换(把派生类的指针或引用转换成基类的指针或引用)。该转换经常用于消除多重继承带来的类型歧义，是相对安全的。如果是纯粹的算数转换，那么请使用后面



的大括号转换方式。

- `reinterpret_cast`：用于转换不相关的类型。`reinterpret_cast` 强制编译器将某个类型对象的内存重新解释成另一种类型，这是一种不安全的转换，建议尽可能少用 `reinterpret_cast`。
- `const_cast`：用于移除对象的 `const` 属性，使对象变得可修改，这样会破坏数据的不变性，建议尽可能少用。
- 算数转换：（C++11开始支持）对于那种算数转换，并且类型信息没有丢失的，比如float到double，int32到int64的转换，推荐使用大括号的初始方式。

```
1. double d{ someFloat };
2. int64_t i{ someInt32 };
```

### 建议9.3.1 避免使用dynamic\_cast

- `dynamic_cast` 依赖于C++的RTTI，让程序员在运行时识别C++类对象的类型。
- `dynamic_cast` 的出现一般说明我们的基类和派生类设计出现了问题，派生类破坏了基类的契约，不得不通过 `dynamic_cast` 转换到子类进行特殊处理，这个时候更希望来改善类的设计，而不是通过 `dynamic_cast` 来解决问题。

### 建议9.3.2 避免使用reinterpret\_cast

说明：`reinterpret_cast` 用于转换不相关类型。尝试用 `reinterpret_cast` 将一种类型强制转换另一种类型，这破坏了类型的安全性与可靠性，是一种不安全的转换。不同类型之间尽量避免转换。

### 建议9.3.3 避免使用const\_cast

说明：`const_cast` 用于移除对象的 `const` 和 `volatile` 性质。

使用const\_cast转换后的指针或者引用来修改const对象，行为是未定义的。

```
1. // 不好的例子
2. const int i = 1024;
3. int* p = const_cast<int*>(&i);
4. *p = 2048;      // 未定义行为
```

```
1. // 不好的例子
2. class Foo {
3. public:
4.     Foo() : i(3) {}
5.
6.     void Fun(int v) {
7.         i = v;
8.     }
9.
10. private:
11.     int i;
12. };
13.
14. int main(void) {
15.     const Foo f;
16.     Foo* p = const_cast<Foo*>(&f);
```

```
17.     p->Fun(8); // 未定义行为
18. }
```

## 资源分配和释放

### 规则9.4.1 单个对象释放使用delete，数组对象释放使用delete []

说明：单个对象删除使用delete， 数组对象删除使用delete []，原因：

- 调用new所包含的动作：从系统中申请一块内存，并调用此类型的构造函数。
- 调用new[n]所包含的动作：申请可容纳n个对象的内存，并且对每一个对象调用其构造函数。
- 调用delete所包含的动作：先调用相应的析构函数，再将内存归还系统。
- 调用delete[]所包含的动作：对每一个对象调用析构函数，再释放所有内存如果new和delete的格式不匹配，结果是未知的。对于非class类型， new和delete不会调用构造与析构函数。

错误写法：

```
1. const int KMaxArraySize = 100;
2. int* numberArray = new int[KMaxArraySize];
3. ...
4. delete numberArray;
5. numberArray = NULL;
```

正确写法：

```
1. const int KMaxArraySize = 100;
2. int* numberArray = new int[KMaxArraySize];
3. ...
4. delete[] numberArray;
5. numberArray = NULL;
```

### 建议9.4.1 使用 RAII 特性来帮助追踪动态分配

说明：RAII是“资源获取就是初始化”的缩写(Resource Acquisition Is Initialization)，是一种利用对象生命周期来控制程序资源(如内存、文件句柄、网络连接、互斥量等等)的简单技术。

RAII 的一般做法是这样的：在对象构造时获取资源，接着控制对资源的访问使之在对象的生命周期内始终保持有效，最后在对象析构的时候释放资源。这种做法有两大好处：

- 我们不需要显式地释放资源。
  - 对象所需的资源在其生命期内始终保持有效。这样，就不必检查资源有效性的问题，可以简化逻辑、提高效率。
- 示例：使用RAII不需要显式地释放互斥资源。

```
1. class LockGuard {
2. public:
3.     LockGuard(const LockType& lockType): lock(lockType) {
4.         lock.Acquire();
5.     }
```

```

6.
7.   ~LockGuard() {
8.       lock.Release();
9.   }
10.
11. private:
12.     LockType lock;
13. };
14.
15.
16. bool Update() {
17.     LockGuard lockGuard(mutex);
18.     if (...) {
19.         return false;
20.     } else {
21.         // 操作数据
22.     }
23.
24.     return true;
25. }

```

## 标准库

STL标准模板库在不同模块使用程度不同，这里列出一些基本规则和建议。

### 规则9.5.1 不要保存std::string的c\_str()返回的指针

说明：在C++标准中并未规定string::c\_str()指针持久有效，因此特定STL实现完全可以在调用string::c\_str()时返回一个临时存储区并很快释放。所以为了保证程序的可移植性，不要保存string::c\_str()的结果，而是在每次需要时直接调用。

示例：

```

1. void Fun1() {
2.     std::string name = "demo";
3.     const char* text = name.c_str(); // 表达式结束以后，name的生命周期还在，指针有效
4.
5.     // 如果中间调用了string的非const成员函数，导致string被修改，比如operator[], begin()等
6.     // 可能会导致text的内容不可用，或者不是原来的字符串
7.     name = "test";
8.     name[1] = '2';
9.
10.    // 后续使用text指针，其字符串内容不再是"demo"
11. }
12.
13. void Fun2() {
14.     std::string name = "demo";
15.     std::string test = "test";
16.     const char* text = (name + test).c_str(); // 表达式结束以后，+号产生的临时对象被销毁，指针无效
17.
18.     // 后续使用text指针，其已不再指向合法内存空间

```

```
19. }
```

例外：在少数对性能要求非常高的代码中，为了适配已有的只接受`const char*`类型入参的函数，可以临时保存`string::c_str()`返回的指针。但是必须严格保证`string`对象的生命周期长于所保存指针的生命周期，并且保证在所保存指针的生命周期内，`string`对象不会被修改。

## 建议9.5.1 使用`std::string`代替`char*`

说明：使用`string`代替 `char*` 有很多优势，比如：

- 不用考虑结尾的`'\0'`；
- 可以直接使用`+`，`=`，`==`等运算符以及其它字符串操作函数；
- 不需要考虑内存分配操作，避免了显式的`new/delete`，以及由此导致的错误；需要注意的是某些`stl`实现中`string`是基于写时复制策略的，这会带来2个问题，一是某些版本的写时复制策略没有实现线程安全，在多线程环境下会引起程序崩溃；二是当与动态链接库相互传递基于写时复制策略的`string`时，由于引用计数在动态链接库被卸载时无法减少可能导致悬挂指针。因此，慎重选择一个可靠的`stl`实现对于保证程序稳定是很重要的。

例外：当调用系统或者其它第三方库的API时，针对已经定义好的接口，只能使用 `char*` 。但是在调用接口之前都可以使用`string`，在调用接口时使用`string::c_str()`获得字符指针。当在栈上分配字符数组当作缓冲区使用时，可以直接定义字符数组，不要使用`string`，也没有必要使用类似 `vector<char>` 等容器。

## 规则9.5.2 禁止使用`auto_ptr`

说明：在`stl`库中的`std::auto_ptr`具有一个隐式的所有权转移行为，如下代码：

```
1. auto_ptr<T> p1(new T);
2. auto_ptr<T> p2 = p1;
```

当执行完第2行语句后，`p1`已经不再指向第1行中分配的对象，而是变为`NULL`。正因为如此，`auto_ptr`不能被置于各种标准容器中。转移所有权的行为通常不是期望的结果。对于必须转移所有权的场景，也不应该使用隐式转移的方式。这往往需要程序员对使用`auto_ptr`的代码保持额外的谨慎，否则出现对空指针的访问。使用`auto_ptr`常见的有两种场景，一是作为智能指针传递到产生`auto_ptr`的函数外部，二是使用`auto_ptr`作为RAII管理类，在超出`auto_ptr`的生命周期时自动释放资源。对于第1种场景，可以使用`std::shared_ptr`来代替。对于第2种场景，可以使用C++11标准中的`std::unique_ptr`来代替。其中`std::unique_ptr`是`std::auto_ptr`的代替品，支持显式的所有权转移。

例外：在C++11标准得到普遍使用之前，在一定需要对所有权进行转移的场景下，可以使用`std::auto_ptr`，但是建议对`std::auto_ptr`进行封装，并禁用封装类的拷贝构造函数和赋值运算符，以使该封装类无法用于标准容器。

## 建议9.5.2 使用新的标准头文件

说明：使用C++的标准头文件时，请使用 `<cstdlib>` 这样的，而不是 `<stdlib.h>` 这种的。

## const的用法

在声明的变量或参数前加上关键字 `const` 用于指明变量值不可被篡改（如 `const int foo` ）。为类中的函数加上

`const` 限定符表明该函数不会修改类成员变量的状态（如 `class Foo { int Bar(char c) const; };` ）。`const` 变量，数据成员，函数和参数为编译时类型检测增加了一层保障，便于尽早发现错误。因此，我们强烈建议在任何可能的情况下使用 `const`。有时候，使用C++11的`constexpr`来定义真正的常量可能更好。

## 规则9.6.1 对于指针和引用类型的形参，如果是不需要修改的，请使用`const`

不变的值更易于理解/跟踪和分析，把`const`作为默认选项，在编译时会对其进行检查，使代码更牢固/更安全。

```
1. class Foo;
2.
3. void PrintFoo(const Foo& foo);
```

## 规则9.6.2 对于不会修改成员变量的成员函数请使用`const`修饰

尽可能将成员函数声明为 `const`。访问函数应该总是 `const`。只要不修改数据成员的成员函数，都声明为`const`。

```
1. class Foo {
2. public:
3.
4.     // ...
5.
6.     int PrintValue() const { // const修饰成员函数，不会修改成员变量
7.         std::cout << value << std::endl;
8.     }
9.
10.    int GetValue() const { // const修饰成员函数，不会修改成员变量
11.        return value;
12.    }
13.
14. private:
15.     int value;
16. };
```

## 建议9.6.1 初始化后不会再修改的成员变量定义为`const`

```
1. class Foo {
2. public:
3.     Foo(int length) : dataLength(length) {}
4. private:
5.     const int dataLength;
6. };
```

## 异常

### 建议9.7.1 C++11中，如果函数不会抛出异常，声明为`noexcept`

理由

- 如果函数不会抛出异常，声明为 `noexcept` 可以让编译器最大程度的优化函数，如减少执行路径，提高错误退出的效率。
- `vector` 等STL容器，为了保证接口的健壮性，如果保存元素的 `move运算符` 没有声明为 `noexcept`，则在容器扩张搬移元素时不会使用 `move机制`，而使用 `copy机制`，带来性能损失的风险。如果一个函数不能抛出异常，或者一个程序并没有截获某个函数所抛出的异常并进行处理，那么这个函数可以用新的 `noexcept` 关键字对其进行修饰，表示这个函数不会抛出异常或者抛出的异常不会被截获并处理。例如：

```
1. extern "C" double sqrt(double) noexcept; // 永远不会抛出异常
2.
3. // 即使可能抛出异常，也可以使用 noexcept
4. // 这里不准备处理内存耗尽的异常，简单地将函数声明为noexcept
5. std::vector<int> MyComputation(const std::vector<int>& v) noexcept {
6.     std::vector<int> res = v; // 可能会抛出异常
7.     // do something
8.     return res;
9. }
```

## 示例

```
1. RetType Function(Type params) noexcept; // 最大的优化
2. RetType Function(Type params); // 更少的优化
3.
4. // std::vector 的 move 操作需要声明 noexcept
5. class Foo1 {
6. public:
7.     Foo1(Foo1&& other); // no noexcept
8. };
9.
10. std::vector<Foo1> a1;
11. a1.push_back(Foo1());
12. a1.push_back(Foo1()); // 触发容器扩张，搬移已有元素时调用copy constructor
13.
14. class Foo2 {
15. public:
16.     Foo2(Foo2&& other) noexcept;
17. };
18.
19. std::vector<Foo2> a2;
20. a2.push_back(Foo2());
21. a2.push_back(Foo2()); // 触发容器扩张，搬移已有元素时调用move constructor
```

注意默认构造函数、析构函数、`swap` 函数，`move操作符` 都不应该抛出异常。

## 模板

模板能够实现非常灵活简洁的类型安全的接口，实现类型不同但是行为相同的代码复用。

模板编程的缺点：

- 模板编程所使用的技巧对于使用c++不是很熟练的人是比较晦涩难懂的。在复杂的地方使用模板的代码让人更不

容易读懂，并且debug 和维护起来都很麻烦。

- 模板编程经常会导致编译出错的信息非常不友好：在代码出错的时候，即使这个接口非常的简单，模板内部复杂的实现细节也会在出错信息显示。导致这个编译出错信息看起来非常难以理解。
- 模板如果使用不当，会导致运行时代码过度膨胀。
- 模板代码难以修改和重构。模板的代码会在很多上下文里面展开来，所以很难确认重构对所有的这些展开的代码有用。所以，建议模板编程最好只用在少量的基础组件，基础数据结构上面。并且使用模板编程的时候尽可能把复杂度最小化，尽量不要让模板对外暴露。最好只在实现里面使用模板，然后给用户暴露的接口里面并不使用模板，这样能提高你的接口的可读性。并且你应该在这些使用模板的代码上写尽可能详细的注释。

## 宏

在C++语言中，我们强烈建议尽可能少使用复杂的宏

- 对于常量定义，请按照前面章节所述，使用const或者枚举；
- 对于宏函数，尽可能简单，并且遵循下面的原则，并且优先使用内联函数，模板函数等进行替换。

```
1. // 不推荐使用宏函数
2. #define SQUARE(a, b) ((a) * (b))
3.
4. // 请使用模板函数，内联函数等来替换。
5. template<typename T> T Square(T a, T b) { return a * b; }
```

如果需要使用宏，请参考C语言规范的相关章节。例外：一些通用且成熟的应用，如：对 new, delete 的封装处理，可以保留对宏的使用。

# 现代C++特性

随着 ISO 在2011年发布 C++11 语言标准，以及2017年3月发布 C++17 ，现代C++(C++11/14/17等)增加了大量提高编程效率、代码质量的新语言特性和标准库。本章节描述了一些可以帮助团队更有效率的使用现代C++，规避语言陷阱的指导意见。

## 代码简洁性和安全性提升

### 建议10.1.1 合理使用auto

理由

- `auto` 可以避免编写冗长、重复的类型名，也可以保证定义变量时初始化。
- `auto` 类型推导规则复杂，需要仔细理解。
- 如果能够使代码更清晰，继续使用明确的类型，且只在局部变量使用 `auto` 。示例

```
1. // 避免冗长的类型名
2. std::map<string, int>::iterator iter = m.find(val);
3. auto iter = m.find(val);
4.
5. // 避免重复类型名
6. class Foo {...};
7. Foo* p = new Foo;
8. auto p = new Foo;
9.
10. // 保证初始化
11. int x;    // 编译正确，没有初始化
12. auto x;   // 编译失败，必须初始化
```

`auto` 的类型推导可能导致困惑：

```
1. auto a = 3;           // int
2. const auto ca = a;    // const int
3. const auto& ra = a;   // const int&
4. auto aa = ca;         // int, 忽略 const 和 reference
5. auto ila1 = { 10 };   // std::initializer_list<int>
6. auto ila2{ 10 };      // std::initializer_list<int>
7.
8. auto&& ura1 = x;       // int&
9. auto&& ura2 = ca;      // const int&
10. auto&& ura3 = 10;     // int&&
11.
12. const int b[10];
13. auto arr1 = b;        // const int*
14. auto& arr2 = b;       // const int(&)[10]
```

如果没有注意 `auto` 类型推导时忽略引用，可能引入难以发现的性能问题：



```
1. std::vector<std::string> v;
2. auto s1 = v[0]; // auto 推导为 std::string, 拷贝 v[0]
```

如果使用 `auto` 定义接口，如头文件中的常量，可能因为开发人员修改了值，而导致类型发生变化。

## 规则10.1.1 在重写虚函数时请使用override关键字

理由 `override` 关键字保证函数是虚函数，且重写了基类的虚函数。如果子类函数与基类函数原型不一致，则产生编译警告。

如果修改了基类虚函数原型，但忘记修改子类重写的虚函数，在编译期就可以发现。也可以避免有多个子类时，重写函数的修改遗漏。

示例

```
1. class Base {
2.     public:
3.         virtual void Foo();
4.         void Bar();
5. };
6.
7. class Derived : public Base {
8.     public:
9.         void Foo() const override; // 编译失败: derived::Foo 和 base::Foo 原型不一致, 不是重写
10.        void Foo() override;       // 正确: derived::Foo 重写 base::Foo
11.        void Bar() override;       // 编译失败: base::Bar 不是虚函数
12.};
```

总结

- 基类首次定义虚函数，使用 `virtual` 关键字
- 子类重写基类虚函数，使用 `override` 关键字
- 非虚函数，`virtual` 和 `override` 都不使用

## 规则10.1.2 使用delete关键字删除函数

理由相比于将类成员函数声明为 `private` 但不实现，`delete` 关键字更明确，且适用范围更广。

示例

```
1. class Foo {
2.     private:
3.         // 只看头文件不知道拷贝构造是否被删除
4.         Foo(const Foo&);
5. };
6.
7. class Foo {
8.     public:
9.         // 明确删除拷贝赋值函数
10.        Foo& operator=(const Foo&) = delete;
```

```
11. };
```

`delete` 关键字还支持删除非成员函数

```
1. template<typename T>
2. void Process(T value);
3.
4. template<>
5. void Process<void>(void) = delete;
```

## 规则10.1.3 使用nullptr，而不是NULL或0

理由长期以来，C++没有一个代表空指针的关键字，这是一件很尴尬的事：

```
1. #define NULL ((void *)0)
2.
3. char* str = NULL; // 错误: void* 不能自动转换为 char*
4.
5. void(C::*pmf)() = &C::Func;
6. if (pmf == NULL) {} // 错误: void* 不能自动转换为指向成员函数的指针
```

如果把 `NULL` 被定义为 `0` 或 `0L`。可以解决上面的问题。

或者在需要空指针的地方直接使用 `0`。但这引入另一个问题，代码不清晰，特别是使用 `auto` 自动推导：

```
1. auto result = Find(id);
2. if (result == 0) { // Find() 返回的是 指针 还是 整数?
3.     // do something
4. }
```

`0` 字面上是 `int` 类型( `0L` 是 `long` )，所以 `NULL` 和 `0` 都不是指针类型。当重载指针和整数类型的函数时，传递 `NULL` 或 `0` 都调用到整数类型重载的函数：

```
1. void F(int);
2. void F(int*);
3.
4. F(0); // 调用 F(int), 而非 F(int*)
5. F(NULL); // 调用 F(int), 而非 F(int*)
```

另外，`sizeof(NULL) == sizeof(void*)` 并不一定总是成立的，这也是一个潜在的风险。

总结：直接使用 `0` 或 `0L`，代码不清晰，且无法做到类型安全；使用 `NULL` 无法做到类型安全。这些都是潜在的风险。

`nullptr` 的优势不仅仅是在字面上代表了空指针，使代码清晰，而且它不再是一个整数类型。

`nullptr` 是 `std::nullptr_t` 类型，而 `std::nullptr_t` 可以隐式的转换为所有的原始指针类型，这使得 `nullptr` 可以表现成指向任意类型的空指针。

```

1. void F(int);
2. void F(int*);
3. F(nullptr);    // 调用 F(int*)
4.
5. auto result = Find(id);
6. if (result == nullptr) { // Find() 返回的是 指针
7.     // do something
8. }

```

## 建议10.1.2 使用using而非typedef

在 C++11 之前，可以通过 `typedef` 定义类型的别名。没人愿意多次重复 `std::map<uint32_t, std::vector<int>>` 这样的代码。

```

1. typedef std::map<uint32_t, std::vector<int>> SomeType;

```

类型的别名实际是对类型的封装。而通过封装，可以让代码更清晰，同时在很大程度上避免类型变化带来的散弹式修改。在 C++11 之后，提供 `using`，实现 `声明别名(alias declarations)`：

```

1. using SomeType = std::map<uint32_t, std::vector<int>>;

```

对比两者的格式：

```

1. typedef Type Alias;    // Type 在前，还是 Alias 在前
2. using Alias = Type;    // 符合 '赋值' 的用法，容易理解，不易出错

```

如果觉得这点还不足以切换到 `using`，我们接着看看 `模板别名(alias template)`：

```

1. // 定义模板的别名，一行代码
2. template<class T>
3. using MyAllocatorVector = std::vector<T, MyAllocator<T>>;
4.
5. MyAllocatorVector<int> data;    // 使用 using 定义的别名
6.
7. template<class T>
8. class MyClass {
9.     private:
10.         MyAllocatorVector<int> data_;    // 模板类中使用 using 定义的别名
11. };

```

而 `typedef` 不支持带模板参数的别名，只能"曲线救国"：

```

1. // 通过模板包装 typedef，需要实现一个模板类
2. template<class T>
3. struct MyAllocatorVector {
4.     typedef std::vector<T, MyAllocator<T>> type;
5. };
6.

```

```

7. MyAllocatorVector<int>::type data; // 使用 typedef 定义的别名, 多写 ::type
8.
9. template<class T>
10. class MyClass {
11. private:
12.     typename MyAllocatorVector<int>::type data_; // 模板类中使用, 除了 ::type, 还需要加上 typename
13. };

```

## 规则10.1.4 禁止使用std::move操作const对象

从字面上看，`std::move` 的意思是要移动一个对象。而const对象是不允许修改的，自然也无法移动。因此用 `std::move` 操作const对象会给代码阅读者带来困惑。在实际功能上，`std::move` 会把对象转换成右值引用类型；对于const对象，会将其转换成const的右值引用。由于极少有类型会定义以const右值引用为参数的移动构造函数和移动赋值操作符，因此代码实际功能往往退化成了对象拷贝而不是对象移动，带来了性能上的损失。

错误示例：

```

1. std::string gString;
2. std::vector<std::string> gStringList;
3.
4. void func() {
5.     const std::string myString = "String content";
6.     gString = std::move(myString); // bad:并没有移动myString, 而是进行了复制
7.     const std::string anotherString = "Another string content";
8.     gStringList.push_back(std::move(anotherString)); // bad:并没有移动anotherString, 而是进行了复制
9. }

```

## 智能指针

### 建议10.2.1 优先使用智能指针而不是原始指针管理资源

理由避免资源泄露。

示例

```

1. void Use(int i) {
2.     auto p = new int {7}; // 不好: 通过 new 初始化局部指针
3.     auto q = std::make_unique<int>(9); // 好: 保证释放内存
4.     if (i > 0) {
5.         return; // 可能 return, 导致内存泄露
6.     }
7.     delete p; // 太晚了
8. }

```

例外在性能敏感、兼容性等场景可以使用原始指针。

### 规则10.2.1 优先使用unique\_ptr而不是shared\_ptr

理由

- `shared_ptr` 引用计数的原子操作存在可测量的开销，大量使用 `shared_ptr` 影响性能。
- 共享所有权在某些情况（如循环依赖）可能导致对象永远得不到释放。
- 相比于谨慎设计所有权，共享所有权是一种诱人的替代方案，但它可能使系统变得混乱。

## 规则10.2.2 使用std::make\_unique而不是new创建unique\_ptr

理由

- `make_unique` 提供了更简洁的创建方式
- 保证了复杂表达式的异常安全示例

```
1. // 不好：两次出现 MyClass，重复导致不一致风险
2. std::unique_ptr<MyClass> ptr(new MyClass(0, 1));
3. // 好：只出现一次 MyClass，不存在不一致的可能
4. auto ptr = std::make_unique<MyClass>(0, 1);
```

重复出现类型可能导致非常严重的问题，且很难发现：

```
1. // 编译正确，但new和delete不配套
2. std::unique_ptr<uint8_t> ptr(new uint8_t[10]);
3. std::unique_ptr<uint8_t[]> ptr(new uint8_t);
4. // 非异常安全：编译器可能按如下顺序计算参数：
5. // 1. 分配 Foo 的内存，
6. // 2. 构造 Foo，
7. // 3. 调用 Bar，
8. // 4. 构造 unique_ptr<Foo>。
9. // 如果 Bar 抛出异常，Foo 不会被销毁，产生内存泄露。
10. F(unique_ptr<Foo>(new Foo()), Bar());
11.
12. // 异常安全：调用函数不会被打断。
13. F(make_unique<Foo>(), Bar());
```

例外 `std::make_unique` 不支持自定义 `deleter`。在需要自定义 `deleter` 的场景，建议在自己的命名空间实现定制版本的 `make_unique`。使用 `new` 创建自定义 `deleter` 的 `unique_ptr` 是最后的选择。

## 规则10.2.3 使用std::make\_shared而不是new创建shared\_ptr

理由使用 `std::make_shared` 除了类似 `std::make_unique` 一致性等原因外，还有性能的因素。`std::shared_ptr` 管理两个实体：

- 控制块（存储引用计数，`deleter` 等）
- 管理对象 `std::make_shared` 创建 `std::shared_ptr`，会一次性在堆上分配足够容纳控制块和管理对象的内存。而使用 `std::shared_ptr<MyClass>(new MyClass)` 创建 `std::shared_ptr`，除了 `new MyClass` 会触发一次堆分配外，`std::shared_ptr` 的构造函数还会触发第二次堆分配，产生额外的开销。

例外类似 `std::make_unique`，`std::make_shared` 不支持定制 `deleter`

## Lambda

## 建议10.3.1 当函数不能工作时选择使用lambda(捕获局部变量，或编写局部函数)

理由函数无法捕获局部变量或在局部范围内声明；如果需要这些东西，尽可能选择 `lambda`，而不是手写的 `functor`。另一方面，`lambda` 和 `functor` 不会重载；如果需要重载，则使用函数。如果 `lambda` 和函数都可以的场景，则优先使用函数；尽可能使用最简单的工具。

### 示例

```
1. // 编写一个只接受 int 或 string 的函数
2. // -- 重载是自然的选择
3. void F(int);
4. void F(const string&);
5.
6. // 需要捕获局部状态，或出现在语句或表达式范围
7. // -- lambda 是自然的选择
8. vector<Work> v = LotsOfWork();
9. for (int taskNum = 0; taskNum < max; ++taskNum) {
10.     pool.Run([=, &v] {...});
11. }
12. pool.Join();
```

## 规则10.3.1 非局部范围使用lambdas，避免使用按引用捕获

理由非局部范围使用 `lambdas` 包括返回值，存储在堆上，或者传递给其它线程。局部的指针和引用不应该在它们的范围外存在。`lambdas` 按引用捕获就是把局部对象的引用存储起来。如果这会导致超过局部变量生命周期的引用存在，则不应该按引用捕获。

### 示例

```
1. // 不好
2. void Foo() {
3.     int local = 42;
4.     // 按引用捕获 local.
5.     // 当函数返回后，local 不再存在，
6.     // 因此 Process() 的行为未定义!
7.     threadPool.QueueWork([&]{ Process(local); });
8. }
9.
10. // 好
11. void Foo() {
12.     int local = 42;
13.     // 按值捕获 local。
14.     // 因为拷贝，Process() 调用过程中，local 总是有效的
15.     threadPool.QueueWork([=]{ Process(local); });
16. }
```

## 建议10.3.2 如果捕获this，则显式捕获所有变量

理由在成员函数中的 `[=]` 看起来是按值捕获。但因为是隐式的按值获取了 `this` 指针，并能够操作所有成员变量，数据成员实际是按引用捕获的，一般情况下建议避免。如果的确需要这样做，明确写出对 `this` 的捕获。

## 示例

```
1. class MyClass {
2. public:
3.     void Foo() {
4.         int i = 0;
5.
6.         auto Lambda = [=]() { Use(i, data_); }; // 不好：看起来像是拷贝/按值捕获，成员变量实际上是按引用捕获
7.
8.         data_ = 42;
9.         Lambda(); // 调用 use(42);
10.        data_ = 43;
11.        Lambda(); // 调用 use(43);
12.
13.        auto Lambda2 = [i, this]() { Use(i, data_); }; // 好，显式指定按值捕获，最明确，最少的混淆
14.    }
15.
16. private:
17.     int data_ = 0;
18. };
```

## 建议10.3.3 避免使用默认捕获模式

理由lambda表达式提供了两种默认捕获模式：按引用（&）和按值（=）。默认按引用捕获会隐式的捕获所有局部变量的引用，容易导致访问悬空引用。相比之下，显式的写出需要捕获的变量可以更容易的检查对象生命周期，减小犯错可能。默认按值捕获会隐式的捕获this指针，且难以看出lambda函数所依赖的变量是哪些。如果存在静态变量，还会让阅读者误以为lambda拷贝了一份静态变量。因此，通常应当明确写出lambda需要捕获的变量，而不是使用默认捕获模式。

## 错误示例

```
1. auto func() {
2.     int addend = 5;
3.     static int baseValue = 3;
4.
5.     return [=]() { // 实际上只复制了addend
6.         ++baseValue; // 修改会影响静态变量的值
7.         return baseValue + addend;
8.     };
9. }
```

## 正确示例

```
1. auto func() {
2.     int addend = 5;
3.     static int baseValue = 3;
4. }
```

```

5.     return [addend, baseValue = baseValue]() mutable { // 使用C++14的捕获初始化拷贝一份变量
6.         ++baseValue;    // 修改自己的拷贝，不会影响静态变量的值
7.         return baseValue + addend;
8.     };
9. }

```

参考：《Effective Modern C++》：Item 31: Avoid default capture modes.

## 接口

### 建议10.4.1 不涉及所有权的场景，使用T\*或T&作为参数，而不是智能指针

理由

- 只在需要明确所有权机制时，才通过智能指针转移或共享所有权。
- 通过智能指针传递，限制了函数调用者必须使用智能指针(如调用者希望传递 `this`)。
- 传递共享所有权的智能指针存在运行时的开销。示例

```

1. // 接受任何 int*
2. void F(int*);
3.
4. // 只能接受希望转移所有权的 int
5. void G(unique_ptr<int>);
6.
7. // 只能接受希望共享所有权的 int
8. void G(shared_ptr<int>);
9.
10. // 不改变所有权，但需要特定所有权的调用者
11. void H(const unique_ptr<int>&);
12.
13. // 接受任何 int
14. void H(int&);
15.
16. // 不好
17. void F(shared_ptr<Widget>& w) {
18.     // ...
19.     Use(*w); // 只使用 w -- 完全不涉及生命周期管理
20.     // ...
21. };

```

### 建议10.4.2 在接口层面明确指针不会为nullptr

理由

- 避免解引用空指针的错误。
- 避免重复检查空指针，提高代码效率。建议使用 `gsl::not_null`，或参考实现自己的版本（如使用 `NotNullObject` 模式）。对于集合，在不违反已有的接口约定的情况下，建议返回空集合而避免返回空指针，当返回字符串时，建议返回空串 `""`。

示例



```
1. int Length(const char* p);    // 不清楚 length == nullptr 是否是有效的
2. Length(nullptr);             // 可以吗？
3.
4. int Length(not_null<const char*> p); // 更好：可以认为 p 不会是空指针
5. int Length(const char* p);      // 必须假设 p 可能是空指针
```

通过在代码中表明意图(指针不能为空)，工具可以提供更好的诊断，如通过静态分析找到一些错误。也可以实现代码优化，如移除判空的测试和分支。

注意 `not_null` 由 `guideline support library(gsl)` 提供。

# 如何贡献

---

## 有哪些形式的贡献

方舟编译器项目诚挚欢迎每个人加入社区并为项目作出贡献。在方舟编译器项目中做贡献，不局限于代码贡献一种形式，您可以通过如下多种形式参与到其中，而我们对您做出的任何贡献表示真诚的感谢：

### 报告错误

如果您发现了OpenArkCompiler的问题，您有必要访问[Issue](#)跟踪系统并上报。登录后创建Issue，将您发现的问题和出现条件进行详细的描述，这将有利于对问题进行重现和快速定位。

### 贡献代码

代码贡献仍是OpenArkCompiler项目最主要的贡献形式。我们欢迎社区志愿者贡献自己的代码，以此获得代码的维护和增强。为保障项目的健康发展和保持高质量的代码库，我们需要遵循如下贡献事宜：

请随时提问。通过在Issue跟踪系统中发表评论或通过[邮件列表](#)提出问题。

重要提示：对OpenArkCompiler项目贡献代码需要遵循流程，并非直接拉取请求和修改代码。首先请与我们取得联系，就需要优化的内容进行讨论，就修改方法达成共识。如果没有与OpenArkCompiler提交者达成共识，贡献可能需要大量返工或不予审核。

具体操作指南请参阅[代码贡献指南](#)。

### 贡献用例

精美的用例能够帮助更多的开发者和用户快速参与到项目中。因此，OpenArkCompiler社区需要大量的用例贡献。针对用例，OpenArkCompiler社区开辟了专门的子项目用于存放代码，请参考[用例代码库](#)，代码提交方式参考上节内容。

### 贡献文档

OpenArkCompiler是较为底层的复杂软件系统。高质量的文档对于项目至关重要。方舟编译器社区旨在提供清晰、准确和完整的文档，以指引社区用户和开发者对项目有更加准确的了解。我们欢迎任何以改进OpenArkCompiler文档的贡献。具体操作指导请参阅[文档贡献指南](#)。

### 帮助代码检视

要保证项目代码的高质量，定期检视代码是必须的。同时，对他人代码进行客观公正的评论及给出优化建议也是一种贡献，并且这是从贡献者晋升为Committer的一个重要考量部分。

### 回答开发者的问题和疑惑

- 回复用户邮件列表中的问题
- 在社区论坛中的问题

- 检查[Issue](#)中的最新问题是否存在实际上是用户问题的故障单

## 传播关于方舟编译器的话题

- 组织或参加OpenArkCompiler Meetup
- 参与OpenArkCompiler微博、微信公众号、知乎等官方讨论渠道
- 在community@openarkcompiler.cn邮件列表上分享您的会议。

## 签署贡献者协议

- 请先下载[贡献者协议](#)；
- 签署后，通过邮件发送至邮箱：[openarkcompiler@huawei.com](mailto:openarkcompiler@huawei.com)；
- 收到OpenArkCompiler确认邮件。

## 注册会员

### 操作步骤

- 进入[华为云](#)首页，单击页面右上角的“注册”。
- 设置手机号并单击“获取短信验证码”。
  - 该手机号未注册过华为云账号，或已注册过华为云账号，但账号数量未达到系统允许的最大数量输入短信验证码、密码并勾选“我已阅读并同意《华为云用户协议》和《隐私政策声明》”。
  - 该手机号已注册过华为云账号，且账号数量已达到系统允许的最大数量将不支持再使用该手机号注册新的华为云账号。
- 单击“同意协议并注册”。账号提示注册成功。

# 文档贡献指南

## 在线编辑文档

开源代码托管平台支持在线编辑预览markdown格式的文档，选择需要修改的文件，点击右上角编辑按钮即可在线编辑。页面分左右两屏显示编辑窗口和预览窗口，编辑完成后再提交信息框内输入修改备注信息后点击提交即可完成文档在线编辑更新操作。

## 在线提交合并请求

提交完成后点击合并请求页签，点击右上角的新建合并请求按钮。在出现的分支选择框中，源分支为您修改文档所在的分支，目标分支为目标仓库的待合入分支。

## 注意事项

### 1、文档目录只支持Atx格式标题提取

Markdown支持两种标题的语法，类 Setext 和类 atx 形式。

类 Setext 形式是用底线的形式，利用 = （最高阶标题）和 - （第二阶标题），例如：

```
1. This is an H1
```

```
1. This is an H2
2. -----</code></pre>
```

任何数量的 = 和 - 都可以有效果。

类 Atx 形式则是在行首插入 1 到 6 个 # ，对应到标题 1 到 6 阶，例如：

```
1. # 这是 H1
2.
3. ## 这是 H2
4.
5. ##### 这是 H6
```

文档项目目录目前只能提取Atx的标题，不能提取Setext这样的标题。

### 2、同一文档内图片不可重名

由于文档项目默认使用统一的文件夹存放图片，因此在上传图片的过程中请特别注意：图片不能重名！否则，新上传的图片将会替换重名的原有图片。

我们正在着手解决这一问题，未更新之前，还请自行图片命名。



# 代码贡献指南

OpenArkCompiler 通过志愿者的代码贡献得到维护，改进和扩展。我们欢迎所有对 OpenArkCompiler 的贡献，但由于项目的规模和保持高质量代码库的原则，我们遵循本文档中解释的贡献过程。

请随时提问。在项目中新建 issue 或者对正在处理的 issue 发表评论，以此同 OpenArkCompiler 提交者达成共识。

重要提示：在开始处理代码贡献之前，请仔细阅读本文档。请遵循以下说明的流程和指南。对 OpenArkCompiler 的贡献不是从提交合并请求（Merge Request）开始的。我们希望贡献者首先与我们联系，共同讨论整体方法后达成一致。如果没有与 OpenArkCompiler 提交者达成共识，贡献可能需要大量返工或不予审核。

## 代码贡献流程

- 创建 issue 并达成共识
- 实施您的更改
- 创建合并请求并等待检视
- 合入合并请求以下是在实施时要记住的一些要点：
- 登录/创建[华为开源平台账号](#)
- [设置开发环境](#)
- 请遵循[编程规范](#)
- 有一个合格的合并请求的贡献更容易被社区接受

## 创建issue

进入到[OpenArkCompiler](#)项目的 issue 列表新建对应的 issue，描述你发现的一个 bug 或者一个好的功能特性。项目 committer 会定期查看 issue 列表中新增的 issue，并进行回复。与 committer 意见达成一致后就可以开始 fork 仓库编写代码了。

**issue 沟通建议：**

- 如果你刚好看到一个未关闭的 **issue**，恰是你打算解决的， 添加评论，告诉他人你将负责这个。这样的话，可以避免他人重复劳动。
- 如果说某个 **issue** 处于未关闭状态很久了， 这可能是已经有人正在解决中，又或者是早已经解决过了，所以也请添加评论，在打算开启工作之前，最好是确认一下。
- 如果你创建了一 **issue**，但是没多久自己解决了， 也要添加评论，让其他人知道，然后关闭该 issue。记录本身就是为社区的贡献。

## 实施您的更改

与项目 committer 达成一致后，您就可以按照如下步骤来实施代码变更

- fork 代码仓库并克隆到本地

- 进入openarkcompiler目录，为本地的仓库配置远端仓库，这样您可以在更新代码时保持和主仓同步，以此减少很多解决冲突的时间。

```
1. git remote add upstream https://code.opensource.huaweicloud.com/HarmonyOS/OpenArkCompiler.git
```

- 创建一个本地分支用于提交您的变更

```
1. git checkout -b mybranch
```

- 同步合入主仓主干分支代码

```
1. git remote update upstream
2. git merge upstream/master
```

- 测试您的改动！若测试用例存在的话，跑一遍，以覆盖你的更改，若没有的话，则创建相应的用例。无论测试是否存在，一定要确保您的改动不会破坏掉现有的项目。
- 尽您最大的努力同项目现有[编程规范](#)保持一致，这便于您的贡献能够更快的被合入，以及未来他人更好的理解和维护。

## 创建合并请求并等待检视

新建合并请求之前的注意事项：

- 确保不包含不相关或不必要的重新格式化更改。
- 确保您的提交历史符合要求。
- 确保您的更改已合入主仓主干分支中的最新提交。
- 确保合并请求关联相应的 issue，并且每个 issue 问题都分配给一个合并请求（如果一个 issue 需要多个合并请求处理，首先要做的是将它拆分成多个 issue）

## 合入合并请求

审核完成后，代码将由 OpenArkCompiler 的 Committer 合并，对应的 issue 将在合并完成同时关闭。

## 邮件列表指南

---

订阅开发者邮件列表，请访问：<https://mail.openarkcompiler.cn/cgi-bin/mailman/listinfo/developer>



## FAQ

### Q1：方舟编译器开源有官方网站吗？

A1：

孵化期间，官方网站为：<https://www.openarkcompiler.cn>；正式毕业后的官网待定。

### Q2：方舟编译器是一次性全部代码吗？

A2：

方舟编译器开源范围示意图



首次开源范围是编译器 IR ( Intermediate Representation )、RC ( Reference Counting ) 和多语言设计思想等，用于与业界、学术界沟通交流。后续将陆续开源编译器前端、后端，支持其它语言（如 JavaScript）的编译等，当前部分Java语言特性和JVM虚拟机特性的支持未包括在本次开源代码中，包括：annotation、lambda表达式、泛型等。目前仍有很多地方不完善，会在社区陆续迭代，遇到问题请在社区提交 issue，欢迎在社区继续讨论设计和代码共建。。

### Q3：通过哪些渠道可以获取方舟编译器的框架源码？

A3:

孵化期间，开发者可以通过如下代码托管网址获得相关代码和文档，供参考学习、了解方舟编译器的架构和代码。

华为云托管地址：<https://code.opensource.huaweicloud.com/HarmonyOS/OpenArkCompiler>。

开源中国码云托管地址：<https://gitee.com/harmonyos/OpenArkCompiler>

Q4：方舟编译器开源路标是怎样的？

A4:

方舟编译器从 2019 年 8 月开始开源，欢迎广大的应用开发者、手机厂商、软件从业人员一起共建开源社区和生态。开源计划大体上这样安排：2019 年 8 月重点开源框架部分；后续将陆续开源编译器前端、后端；支持 Java 程序编译、JavaScript 语言应用的编译等。

Q5：还有哪些方式可以获取方舟编译器的相关信息？

A5:



方舟编译器有官方微信公众号：开源方舟编译器  
相关信息。

大家可以关注，及时获得

Q6：方舟编译开源面向全球开发者吗？英文网站何时上线？

A6:

方舟编译器是向全球开发者进行开源，因此，官方网站和代码托管平台均面向全球开发者开放。目前英文版网站正在紧张的开发及内测，计划将于 2019 年 Q4 上线。

Q7：方舟编译器源码托管计划是怎么的？

A7:

孵化委员会期间，方舟编译器代码是托管在华为云上面。在未来走向开放治理以后，将由所挂靠机构选择的平台进行

托管。如果挂靠的机构采取了 X 平台，我们就采用这个机构的代码托管模式进行托管。

当前代码托管的地址：<https://code.opensource.huaweicloud.com/HarmonyOS/OpenArkCompiler>

所有相关信息请以方舟编译器开源官方网站发布为准。

## Q8：方舟编译器开源使用的是什么License？

A8：

方舟编译器采用木兰 License。木兰 License 由业界众多专家和公司集体共建，并非华为主导。木兰 License 具体的好处，详见专业点评：

- 开源中国：<https://www.oschina.net/news/108840/mulanpsl-the-1st-open-source-license-of-china>
- 开源社：<https://mp.weixin.qq.com/s/1xqxxiNKhcgw4RRspAaW2A>
- InfoQ：[https://www.infoq.cn/article/KrZAbAcJrxCU27vKj\\*JE](https://www.infoq.cn/article/KrZAbAcJrxCU27vKj*JE)

## Q9：如何确认下载包来源可靠？MD5校验码是多少？

A9：

OpenArkCompiler-0.2-ubuntu-16.04-x86\_64.tar.gz 包的 MD5码是：  
e7dc8b8c242f2d26429e2ad71dae3112

OpenArkCompiler-0.2.tar.gz 包的MD5码是：1bf859a67c8b9a2d4080bdaf25ef6181