
DragonOS

发布 *dev*

fslongjin

2023 年 03 月 13 日

1	DragonOS 简介	1
2	构建 DragonOS	7
3	DragonOS 镜像站	15
4	引导加载	17
5	核心 API 文档	21
6	锁	69
7	进程管理模块	89
8	DragonOS 调度	95
9	进程间通信	107
10	内存管理文档	111
11	文件系统	119
12	内核调试模块	121
13	内核测试	123
14	处理器架构	127
15	LibC 文档	129
16	系统调用 API	145
17	参与开发	147

18 与社区建立联系	161
19 发行日志	163
20 Indices and tables	217

DragonOS 龙操作系统（以下简称“DragonOS”）是一个面向服务器领域的，从 0 开发内核及用户态环境，并提供 Linux 兼容性的 64 位操作系统。它使用 Rust 与 C 语言进行编写，并正在逐步淘汰原有的 C 代码，以在将来提供更好的安全性与可靠性。

DragonOS 的目标是，构建一个完全独立自主的、开源的、高性能及高可靠性的服务器操作系统，为国家数字基础设施建设提供完全独立自主的底层核心动力。

作为一个社区驱动的开源操作系统，为了促进其发展，避免让其遭受一些不遵守开源协议的商业公司的侵权，我们决定使用 GPLv2 协议开放源代码，以严格的开源协议来保护 DragonOS。

您可能对 DragonOS 中已经实现了哪些功能感兴趣，您可以转到这里：[功能特性](#)

1.1 DragonOS 的功能

1.1.1 规范

- 启动引导：Multiboot2
- 接口：posix 2008

1.1.2 内核层

内存管理

- 页分配器
- slab 分配器
- VMA
- MMIO 地址空间自动分配

多核

- 多核引导
- ipi 框架

进程管理

- 进程创建
- 进程回收
- 内核线程
- fork
- exec
- 进程睡眠（支持高精度睡眠）
- kthread 机制

同步原语

- mutex 互斥量
- semaphore 信号量
- atomic 原子变量
- spinlock 自旋锁
- wait_queue 等待队列

调度

- CFS 调度器
- 单核调度

IPC

- 匿名 pipe 管道
- signal 信号

文件系统

- VFS
- fat32
- devfs
- rootfs

异常及中断处理

- APIC
- softirq 软中断
- 内核栈 traceback

内核数据结构

- 普通二叉树
- kfifo 缓冲区
- 循环链表
- IDR

内核实用库

- LZ4 压缩库 (1.9.3)
- 字符串操作库
- ELF 可执行文件支持
- printk
- 基础数学库
- 屏幕管理器
- textui 框架
- CRC 函数库

系统调用

请见系统调用文档

测试框架

- ktest

驱动程序

- ACPI 高级电源配置模块
- IDE 硬盘
- AHCI 硬盘
- PCI
- XHCI (usb3.0)
- ps/2 键盘
- ps/2 鼠标
- HPET 高精度定时器
- RTC 时钟
- local apic 定时器
- UART 串口
- VBE 显示

1.1.3 用户层

LibC

- 基础系统调用
- 基础标准库函数
- 部分数学函数

shell 命令程序

- 基于字符串匹配的解析
- 基本的几个命令

驱动程序

- ps/2 键盘用户态驱动

1.1.4 软件移植

- GCC 11.3.0（暂时只支持了 x86_64 的 Cross Compiler）<https://github.com/DragonOS-Community/gcc>
- binutils 2.38（暂时只支持了 x86_64 的 Cross Compiler）<https://github.com/DragonOS-Community/binutils>
- gmp 6.2.1 <https://github.com/DragonOS-Community/gmp-6.2.1>
- mpfr 4.1.1 <https://github.com/DragonOS-Community/mpfr>
- mpc 1.2.1 <https://github.com/DragonOS-Community/mpc>

2.1 1. 写在前面

无论您采用后文中的何种方式来编译 DragonOS，您必须先按照本小节中的步骤，初始化您的开发环境。

开始之前，您需要一台运行 Linux 或 MacOS 的计算机，并且处理器架构为 X86-64.

对于 Linux 发行版，建议使用 Ubuntu22、Debian、Arch Linux 这样的，仓库软件版本较新的发行版，这能为您减少很多麻烦。

2.1.1 使用一键初始化脚本进行安装

我们提供了一键初始化脚本，可以一键安装，只需要在控制台运行以下命令：

```
cd tools
bash bootstrap.sh # 这里请不要加上sudo, ↵
↵ 因为需要安装的开发依赖包是安装在用户环境而非全局环境
```

备注：一键配置脚本目前只支持以下系统：

- Ubuntu/Debian/Deepin/UOS 等基于 Debian 的衍生版本

欢迎您为其他的系统完善构建脚本！

如果一键初始化脚本能够正常运行，并输出最终的“祝贺”界面(如下所示)，那么恭喜你，可以直接跳到[这里](#)进行阅读！

```
|-----Congratulations!-----|
|                                |
|  你成功安装了DragonOS所需的依赖项！    |
|  您可以通过以下命令运行它：              |
|                                |
|  make run-docker -j 你的cpu核心数        |
|                                |
|-----|
```

2.1.2 依赖清单（手动安装）

如果自动安装脚本不能支持您的操作系统，那么您需要手动安装依赖程序。以下是依赖项的清单：

在以下依赖项中，除了 `docker-ce` 和 `Rust` 及其工具链以外，其他的都能通过系统自带的包管理器进行安装。关于 `docker` 以及 `rust` 的安装，请看后文。

- `docker-ce`
- `llvm-dev`
- `libclang-dev`
- `clang`
- `gcc-multilib`
- `qemu qemu-system qemu-kvm`
- `build-essential`
- `fdisk`
- `lsb-release`
- `git`
- `dosfstools`
- `Rust` 以及其工具链

请注意，若您的 `Linux` 系统是在虚拟机中运行的，还请您在您的 `VMware/Virtual Box` 虚拟机的处理器设置选项卡中，开启 `Intel VT-x` 或 `AMD-V` 选项，否则，`DragonOS` 将无法运行。

备注：在某些 `Linux` 发行版的软件仓库中构建的 `Qemu` 可能由于版本过低而不兼容 `DragonOS`，如果遇到这种问题，请卸载 `Qemu`，并采用编译安装的方式重新安装 `Qemu`

在该地址下载 `Qemu` 源代码：<https://download.qemu.org/>

解压后进入源代码目录，然后执行下列命令：

```
# 安装编译依赖项
sudo apt install -y autoconf automake autotools-dev curl libmpc-dev libmpfr-dev \
↳ libgmp-dev \
    gawk build-essential bison flex texinfo gperf libtool patchutils bc \
    zlib1g-dev libexpat-dev pkg-config libglib2.0-dev libpixman-1-dev \
↳ libSDL2-dev \
    git tmux python3 python3-pip ninja-build

./configure --enable-kvm
make -j 8
sudo make install
# 编译安装完成
```

请注意，编译安装的 QEMU，将通过 VNC 模式进行链接，因此，您还需要在您的计算机上安装 VNC viewer 以连接至 QEMU 虚拟机。

2.1.3 安装 Docker

您可以在 docker 官网下载安装 docker-ce.

详细信息请转到：<https://docs.docker.com/engine/install/>

2.1.4 安装 Rust

警告：【常见误区】：如果您打算采用 docker 进行编译，尽管 docker 镜像中已经安装了 Rust 编译环境，但是，为了能够在 VSCode 中使用 Rust-Analyzer 进行代码提示，以及 make clean 命令能正常运行，您的客户机上仍然需要安装 rust 环境。

您可以在控制台输入以下命令，安装 rust。

```
# 这两行用于换源，加速Rust的安装过程
export RUSTUP_DIST_SERVER=https://mirrors.ustc.edu.cn/rust-static
export RUSTUP_UPDATE_ROOT=https://mirrors.ustc.edu.cn/rust-static/rustup
# 安装Rust
curl https://sh.rustup.rs -sSf | sh -s -- --default-toolchain nightly
# 把Rustup加到环境变量
echo "export PATH=\"$HOME/.cargo/bin:$PATH\"" >> ~/.bashrc
source ~/.cargo/env
source "$HOME/.cargo/env"
```

(下页继续)

(续上页)

```
# 更换cargo的索引源
touch ~/.cargo/config
echo -e "[source.crates-io] \n \
registry = \"https://github.com/rust-lang/crates.io-index\" \n \
\n \
replace-with = 'dragonos-gitee' \n \
[source.dragonos-gitee] \n \
registry = \"https://gitee.com/DragonOS/crates.io-index.git\" \n \
" > ~/.cargo/config

# 安装DragonOS所需的工具链
cargo install cargo-binutils
rustup toolchain install nightly
rustup default nightly
rustup component add rust-src
rustup component add llvm-tools-preview
rustup target add x86_64-unknown-none
# Rust安装完成
```

2.1.5 解决 KVM 权限问题

在部分计算机上，可能由于权限问题而无法启动虚拟机，我们可以通过把当前用户加到 `kvm` 用户组的方式，解决该问题：

```
# 解决kvm权限问题
USR=$USER
sudo adduser $USR kvm
sudo chown $USR /dev/kvm
```

2.1.6 下载 DragonOS 的源代码

假设您的计算机上已经安装了 `git`，您可以通过以下命令，获得 DragonOS 的最新的源代码：

```
git clone https://github.com/fslongjin/DragonOS
cd DragonOS
```

至此，公共依赖项已经安装完成，您可以根据自己的需要，阅读后续章节

2.2 2. 从 Docker 构建（推荐）

为减轻配置环境的负担，DragonOS 发布了一个 Docker 编译环境，便于开发者运行 DragonOS。我们强烈建议您采用这种方式来运行 DragonOS。

本节假设以下操作均在 Linux 下进行。

2.2.1 获取 DragonOS 编译镜像

当您成功安装了 docker 之后，您可以通过以下命令，下载 DragonOS 的编译镜像：

```
docker pull dragonos/dragonos-dev:v1.1.0-beta3
```

2.2.2 安装 qemu 虚拟机

在本节中，我们建议您采用命令行安装 qemu：

```
sudo apt install -y qemu qemu-system qemu-kvm
```

2.2.3 创建磁盘镜像

首先，您需要使用 tools 文件夹下的 create_hdd_image.sh，创建一块虚拟磁盘镜像。您需要在 tools 文件夹下运行此命令。

```
bash create_hdd_image.sh
```

2.2.4 运行 DragonOS

如果不出意外的话，这将是运行 DragonOS 的最后一步。您只需要在 DragonOS 的根目录下方，执行以下命令，即可运行 DragonOS。

```
make run-docker
```

稍等片刻，DragonOS 将会被运行。

在 qemu 虚拟机被启动后，我们需要在控制台输入字母 c，然后回车。这样，虚拟机就会开始执行。

备注：(1) 首次编译时，由于需要下载 Rust 相关的索引（几百 MB 大小），因此需要一定的时间，请耐心等待！
(2) 输入命令可能需要加上 sudo

关于编译命令的用法，请见：[编译命令讲解](#)

2.3 3. 在本机中直接编译

若您追求快速的编译速度，以及完整的开发调试支持，且愿意花费半个小时来配置开发环境的话，该小节的内容能帮助您。

2.3.1 软件依赖

您需要通过以下命令，获取您本机安装的 Grub 的版本：

```
grub-install --version
```

如果显示的版本号为 **2.06** 及以上，且您已经按照第一小节中的内容，安装相关的依赖，那么，恭喜您，您可以直接在本机编译 **DragonOS**！

否则，您需要编译安装 Grub-2.06。它的源代码可以通过<https://ftp.gnu.org/gnu/grub/grub-2.06.tar.gz>获得。

- grub 2.06 (不必使用 sudo 权限进行 install)

2.3.2 创建磁盘镜像

首先，您需要使用 `sudo` 权限运行 `tools/create_hdd_image.sh`，为 DragonOS 创建一块磁盘镜像文件。该脚本会自动完成创建磁盘镜像的工作，并将其移动到 `bin/`目录下。

2.3.3 编译、运行 DragonOS

1. 安装编译及运行环境
2. 进入 DragonOS 文件夹
3. 输入 `make run` 即可编译并写入磁盘镜像，并运行

在 `qemu` 虚拟机被启动后，我们需要在控制台输入字母 `c`，然后回车。这样，虚拟机就会开始执行。

备注：首次编译时，由于需要下载 Rust 相关的索引（几百 MB 大小），因此需要一定的时间，请耐心等待！

关于编译命令的用法，请见：[编译命令讲解](#)

2.4 4. 编译命令讲解

- 本地编译, 不运行: `make all -j 您的 CPU 核心数`
- 本地编译, 并写入磁盘镜像, 不运行: `make build`
- 本地编译, 写入磁盘镜像, 并在 QEMU 中运行: `make run`
- Docker 编译, 并写入磁盘镜像: `make docker`
- Docker 编译, 写入磁盘镜像, 并在 QEMU 中运行: `make run-docker`
- 不编译, 直接从已有的磁盘镜像启动: `make qemu`

CHAPTER 3

DragonOS 镜像站

您可以从以下镜像站下载到 DragonOS 的源代码和其他文件：

- [DragonOS 镜像站](#)
- [DragonOS 国内镜像站 \(RinGoTek\)](#)

DragonOS 采用 GRUB2 作为其引导加载程序，支持 Multiboot2 协议引导。目前仅支持 GRUB2.06 版本。

4.1 引导加载程序

4.1.1 原理

目前，DragonOS 支持 Legacy BIOS 以及 UEFI 两种方式，进行启动引导。

在 `head.S` 的头部包含了 Multiboot2 引导头，里面标志了一些 Multiboot2 相关的特定信息，以及一些配置命令。

在 DragonOS 的启动初期，会存储由 GRUB2 传来的 magic number 以及 `multiboot2_boot_info_addr`。当系统进入 `Start_Kernel` 函数之后，将会把这两个信息保存到 `multiboot2` 驱动程序之中。信息的具体含义请参照 Multiboot2 Specification 进行理解，该部分难度不大，相信读者经过思考能理解其中的原理。

4.1.2 参考资料

- [Multiboot2 Specification](#)
- [GNU GRUB Manual 2.06](#)
- [UEFI/Legacy 启动 - yujianwu - DragonOS 社区](#)

4.2 Multiboot2 支持模块

Multiboot2 支持模块提供对 Multiboot2 协议的支持。位于 `kernel/driver/multiboot2` 文件夹中。

根据 Multiboot2 协议，操作系统能够从 BootLoader 处获得一些信息，比如基本的硬件信息以及 ACPI 表的起始地址等。

4.2.1 数据结构

`kernel/driver/multiboot2/multiboot2.h` 中按照 Multiboot2 协议的规定，定义了大部分的数据结构，具体细节可查看该文件: [DragonOS/multiboot2.h at master · fslongjin/DragonOS · GitHub](#)

4.2.2 迭代器

由于 Multiboot2 的信息存储在自 `multiboot2_boot_info_addr` 开始的一段连续的内存空间之中，且不同类型的 header 的长度不同，因此设计了一迭代器 `multiboot2_iter`。

函数原型

```
void multiboot2_iter(bool (*_fun) (const struct iter_data_t *, void *, unsigned int *),
                    void *data, unsigned int *count)
```

`_fun`

指定的 handler。当某个 header 的 tag 与该 handler 所处理的 tag 相同时，handler 将处理该 header，并返回 true。

其第一个参数为 tag 类型，第二个参数为返回的数据的指针，第三个值为计数（某些没有用到该值的地方，该值可以为空）

`data`

传递给 `_fun` 的第二个参数，`_fun` 函数将填充该指针所指向的内存区域，从而返回数据。

count

当返回的 **data** 为一个列表时，通过该值来指示列表中有多少项。

4.2.3 迭代工作函数

在模块中，按照我们需要获取不同类型的 **tag** 的需要，定义了一些迭代器工作函数。

这里是 DragonOS 的核心 api 文档。

5.1 DragonOS 内核核心 API

5.1.1 循环链表管理函数

循环链表是内核的重要的数据结构之一。包含在 `kernel/common/list.h` 中。

```
void list_init(struct List *list)
```

描述

初始化一个 List 结构体，使其 prev 和 next 指针指向自身

参数

list

要被初始化的 List 结构体

```
void list_add(struct List *entry, struct List *node)
```

描述

将 node 插入到 entry 的后方

参数

entry

已存在于循环链表中的一个结点

node

待插入的结点

```
void list_append(struct List *entry, struct List *node)
```

描述

将 node 插入到 entry 的前方

参数

entry

已存在于循环链表中的一个结点

node

待插入的结点

```
void list_del(struct List *entry)
```

描述

从链表中删除结点 entry

参数

entry

待删除的结点

```
list_del_init(struct List *entry)
```

描述

从链表中删除结点 entry，并将这个 entry 使用 list_init() 进行重新初始化。

参数

entry

待删除的结点

```
bool list_empty(struct List *entry)
```

描述

判断链表是否为空

参数

entry

链表中的一个结点

```
struct List *list_prev(struct List *entry)
```

描述

获取 entry 的前一个结点

参数

entry

链表中的一个结点

```
struct List *list_next(struct List *entry)
```

描述

获取 entry 的后一个结点

参数

entry

链表中的一个结点

```
void list_replace(struct List *old, struct List *new)
```

描述

将链表中的 old 结点替换成 new 结点

参数

old

要被换下来的结点

new

要被换入链表的新的结点

```
list_entry(ptr, type, member)
```

描述

该宏能通过 ptr 指向的 List 获取到 List 所处的结构体的地址

参数

ptr

指向 List 结构体的指针

type

要被换入链表的新的结点

member

List 结构体在上述的“包裹 list 结构体的结构体”中的变量名

```
list_first_entry(ptr, type, member)
```

描述

获取链表中的第一个元素。请注意，该宏要求链表非空，否则会出错。

参数

与`list_entry()`相同

```
list_first_entry_or_null(ptr, type, member)
```

描述

获取链表中的第一个元素。若链表为空，则返回 NULL。

参数

与`list_entry()`相同

```
list_last_entry(ptr, type, member)
```

描述

获取链表中的最后一个元素。请注意，该宏要求链表非空，否则会出错。

参数

与`list_entry()`相同

```
list_last_entry_or_full(ptr, type, member)
```

描述

获取链表中的最后一个元素。若链表为空，则返回 NULL。

参数

与`list_entry()`相同

```
list_next_entry(pos, member)
```

描述

获取链表中的下一个元素

参数

pos

指向当前的外层结构体的指针

member

链表结构体在外层结构体内的变量名

```
list_prev_entry(pos, member)
```

描述

获取链表中的上一个元素

参数

与`list_next_entry()`相同

```
list_for_each(ptr, head)
```

描述

遍历整个链表（从前往后）

参数

ptr

指向 List 结构体的指针

head

指向链表头结点的指针 (struct List*)

```
list_for_each_prev(ptr, head)
```

描述

遍历整个链表（从后往前）

参数

与`list_for_each()`相同

`list_for_each_safe(ptr, n, head)`

描述

从前往后遍历整个链表（支持删除当前链表结点）

该宏通过暂存中间变量，防止在迭代链表的过程中，由于删除了当前 `ptr` 所指向的链表结点而造成错误。

参数

ptr

指向 List 结构体的指针

n

用于存储临时值的 List 类型的指针

head

指向链表头结点的指针 (struct List*)

`list_for_each_prev_safe(ptr, n, head)`

描述

从后往前遍历整个链表。（支持删除当前链表结点）

该宏通过暂存中间变量，防止在迭代链表的过程中，由于删除了当前 `ptr` 所指向的链表结点而造成错误。

参数

与 `list_for_each_safe()` 相同

`list_for_each_entry(pos, head, member)`

描述

从头开始迭代给定类型的链表

参数**pos**

指向特定类型的结构体的指针

head

指向链表头结点的指针 (struct List*)

member

struct List 在 pos 的结构体中的成员变量名

```
list_for_each_entry_reverse(pos, head, member)
```

描述

逆序迭代给定类型的链表

参数

与`list_for_each_entry()`相同

```
list_for_each_entry_safe(pos, n, head, member)
```

描述

从头开始迭代给定类型的链表（支持删除当前链表结点）

参数**pos**

指向特定类型的结构体的指针

n

用于存储临时值的，和 pos 相同类型的指针

head

指向链表头结点的指针 (struct List*)

member

struct List 在 pos 的结构体中的成员变量名

```
list_prepare_entry(pos, head, member)
```

描述

为`list_for_each_entry_continue()`准备一个' pos' 结构体

参数

pos

指向特定类型的结构体的，用作迭代起点的指针

head

指向要开始迭代的 struct List 结构体的指针

member

struct List 在 pos 的结构体中的成员变量名

```
list_for_each_entry_continue(pos, head, member)
```

描述

从指定的位置的【下一个元素开始】，继续迭代给定的链表

参数

pos

指向特定类型的结构体的指针。该指针用作迭代的指针。

head

指向要开始迭代的 struct List 结构体的指针

member

struct List 在 pos 的结构体中的成员变量名

```
list_for_each_entry_continue_reverse(pos, head, member)
```

描述

从指定的位置的【上一个元素开始】，【逆序】迭代给定的链表

参数

与`list_for_each_entry_continue()`的相同

```
list_for_each_entry_from(pos, head, member)
```

描述

从指定的位置开始, 继续迭代给定的链表

参数

与`list_for_each_entry_continue()`的相同

```
list_for_each_entry_safe_continue(pos, n, head, member)
```

描述

从指定的位置的【下一个元素开始】，继续迭代给定的链表。（支持删除当前链表结点）

参数

pos

指向特定类型的结构体的指针。该指针用作迭代的指针。

n

用于存储临时值的，和 pos 相同类型的指针

head

指向要开始迭代的 struct List 结构体的指针

member

struct List 在 pos 的结构体中的成员变量名

```
list_for_each_entry_safe_continue_reverse(pos, n, head, member)
```

描述

从指定的位置的【上一个元素开始】，【逆序】迭代给定的链表。（支持删除当前链表结点）

参数

与`list_for_each_entry_safe_continue()`的相同

```
list_for_each_entry_safe_from(pos, n, head, member)
```

描述

从指定的位置开始,继续迭代给定的链表。（支持删除当前链表结点）

参数

与`list_for_each_entry_safe_continue()`的相同

5.1.2 基础 C 函数库

内核编程与应用层编程不同，你将无法使用 LibC 中的函数来进行编程。为此，内核实现了一些常用的 C 语言函数，并尽量使其与标准 C 库中的函数行为相近。值得注意的是，这些函数的行为可能与标准 C 库函数不同，请在使用时仔细阅读以下文档，这将会为你带来帮助。

字符串操作

```
int strlen(const char *s)
```

描述

测量并返回字符串长度。

参数

src

源字符串

```
long strnlen(const char *src, unsigned long maxlen)
```

描述

测量并返回字符串长度。当字符串长度大于 maxlen 时，返回 maxlen

参数

src

源字符串

maxlen

最大长度

```
long strnlen_user(const char *src, unsigned long maxlen)
```

描述

测量并返回字符串长度。当字符串长度大于 maxlen 时，返回 maxlen。

该函数会进行地址空间校验，要求 src 字符串必须来自用户空间。当源字符串来自内核空间时，将返回 0。

参数

src

源字符串，地址位于用户空间

maxlen

最大长度

```
char *strncpy(char *dst, const char *src, long count)
```

描述

拷贝长度为 count 个字节的字符串，返回 dst 字符串

参数

src

源字符串

dst

目标字符串

count

要拷贝的源字符串的长度

```
char *strcpy(char *dst, const char *src)
```

描述

拷贝源字符串，返回 dst 字符串

参数

src

源字符串

dst

目标字符串

```
long strncpy_from_user(char *dst, const char *src, unsigned long size)
```

描述

从用户空间拷贝长度为 count 个字节的字符串到内核空间，返回拷贝的字符串的大小

该函数会对字符串的地址空间进行校验，防止出现地址空间越界的问题。

参数

src

源字符串

dst

目标字符串

size

要拷贝的源字符串的长度

```
int strcmp(char *FirstPart, char *SecondPart)
```

描述

比较两个字符串的大小。

返回值

情况	返回值
FirstPart == SecondPart	0
FirstPart > SecondPart	1
FirstPart < SecondPart	-1

参数

FirstPart

第一个字符串

SecondPart

第二个字符串

```
printk(const char* fmt, ...)
```

描述

该宏能够在控制台上以黑底白字格式化输出字符串.

参数

fmt

源格式字符串

...

可变参数

```
printk_color(unsigned int FRcolor, unsigned int BKcolor, const char* fmt, ...)
```

描述

在控制台上以指定前景色和背景色格式化输出字符串.

参数

FRcolor

前景色

BKcolor

背景色

fmt

源格式字符串

...

可变参数

```
int vsprintf(char *buf, const char *fmt, va_list args)
```

描述

按照 `fmt` 格式化字符串，并将结果输出到 `buf` 中，返回写入 `buf` 的字符数量。

参数

buf

输出缓冲区

fmt

源格式字符串

args

可变参数列表

```
int sprintk(char *buf, const char *fmt, ...)
```

描述

按照 `fmt` 格式化字符串，并将结果输出到 `buf` 中，返回写入 `buf` 的字符数量。

参数

buf

输出缓冲区

fmt

源格式字符串

...

可变参数

内存操作

```
void *memcpy(void *dst, const void *src, uint64_t size)
```

描述

将内存从 `src` 处拷贝到 `dst` 处。

参数

dst

指向目标地址的指针

src

指向源地址的指针

size

待拷贝的数据大小

```
void *memmove(void *dst, const void *src, uint64_t size)
```

描述

与 `memcpy()` 类似，但是在源数据区域与目标数据区域之间存在重合时，该函数能防止数据被错误的覆盖。

参数

dst

指向目标地址的指针

src

指向源地址的指针

size

待拷贝的数据大小

5.1.3 CRC 函数

函数列表

```
uint8_t crc7(uint8_t crc, const uint8_t *buffer, size_t len)
```

```
uint8_t crc8(uint8_t crc, const uint8_t *buffer, size_t len)
```

```
uint16_t crc16(uint16_t crc, const uint8_t *buffer, size_t len)
```

```
uint32_t crc32(uint32_t crc, const uint8_t *buffer, size_t len)
```

```
uint64_t crc64(uint64_t crc, const uint8_t *buffer, size_t len)
```

描述

用于计算循环冗余校验码

参数说明

crc

传入的 CRC 初始值

buffer

待处理的数据缓冲区

len

缓冲区大小（字节）

5.2 原子变量

5.2.1 简介

DragonOS 实现了原子变量，类型为 `atomic_t`。原子变量是基于具体体系结构的原子操作指令实现的。具体实现在 `kernel/common/atomic.h` 中。

5.2.2 API

请注意，以下 API 均为原子操作。

```
inline void atomic_add(atomic_t *ato, long val)
```

描述

原子变量增加指定值

参数

ato

原子变量对象

val

变量要增加的值

```
inline void atomic_sub(atomic_t *ato, long val)
```

描述

原子变量减去指定值

参数

ato

原子变量对象

val

变量要被减去的值

```
void atomic_inc(atomic_t *ato)
```

描述

原子变量自增 1

参数

ato

原子变量对象

```
void atomic_dec(atomic_t *ato)
```

描述

原子变量自减 1

参数

ato

原子变量对象

```
inline void atomic_set_mask(atomic_t *ato, long mask)
```

描述

将原子变量的值与 mask 变量进行 or 运算

参数

ato

原子变量对象

mask

与原子变量进行 or 运算的变量

```
inline void atomic_clear_mask(atomic_t *ato, long mask)
```

描述

将原子变量的值与 mask 变量进行 and 运算

参数

ato

原子变量对象

mask

与原子变量进行 and 运算的变量

5.3 内核数据结构

内核中实现了常用的几种数据结构，这里是他们的 api 文档。

5.3.1 kfifo 先进先出缓冲区

kfifo 先进先出缓冲区定义于 `common/kfifo.h` 中。您可以使用它，创建指定大小的 fifo 缓冲区（最大大小为 4GB）

kfifo_alloc

```
int kfifo_alloc(struct kfifo_t *fifo, uint32_t size, uint64_t reserved)
```

描述

通过动态方式初始化 kfifo 缓冲队列。fifo 缓冲区的 buffer 将由该函数进行申请。

参数

fifo

kfifo 队列结构体的指针

size

缓冲区大小（单位：bytes）

reserved

当前字段保留，请将其置为 0

返回值

当返回值为 0 时，表示正常初始化成功，否则返回对应的 `errno`

kfifo_init

```
void kfifo_init(struct kfifo_t *fifo, void *buffer, uint32_t size)
```

描述

使用指定的缓冲区来初始化 kfifo 缓冲队列

参数

fifo

kfifo 队列结构体的指针

buffer

缓冲区基地址指针

size

缓冲区大小（单位：bytes）

kfifo_free_alloc

```
void kfifo_free_alloc(struct kfifo_t* fifo)
```

描述

释放通过 kfifo_alloc 创建的 fifo 缓冲区。请勿通过该函数释放其他方式创建的 kfifo 缓冲区。

参数

fifo

kfifo 队列结构体的指针

kfifo_in

```
uint32_t kfifo_in(struct kfifo_t *fifo, const void *from, uint32_t size)
```

描述

向 kfifo 缓冲区推入指定大小的数据。当队列中空间不足时，则不推入数据。

参数

fifo

kfifo 队列结构体的指针

from

源数据基地址指针

size

数据大小（单位：bytes）

返回值

返回成功被推入的数据的大小。

kfifo_out

```
uint32_t kfifo_out(struct kfifo_t *fifo, void *to, uint32_t size)
```

描述

从 kfifo 缓冲区取出数据，并从队列中删除数据。当队列中数据量不足时，则不取出。

参数

fifo

kfifo 队列结构体的指针

to

目标缓冲区基地址指针

size

数据大小（单位：bytes）

返回值

返回成功被取出的数据的大小。

kfifo_out_peek

```
uint32_t kfifo_out_peek(struct kfifo_t *fifo, void *to, uint32_t size)
```


描述

从 kfifo 缓冲区取出数据，但是不从队列中删除数据。当队列中数据量不足时，则不取出。

参数

fifo

kfifo 队列结构体的指针

to

目标缓冲区基地址指针

size

数据大小（单位：bytes）

返回值

返回成功被取出的数据的大小。

kfifo_reset

```
kfifo_reset(fifo)
```

描述

忽略 kfifo 队列中的所有内容，并把输入和输出偏移量都归零

参数

fifo

kfifo 队列结构体的指针

kfifo_reset_out

```
kfifo_reset_out(fifo)
```

描述

忽略 kfifo 队列中的所有内容，并将输入偏移量赋值给输出偏移量

参数

fifo

kfifo 队列结构体的指针

kfifo_total_size

```
kfifo_total_size(fifo)
```

描述

获取 kfifo 缓冲区的最大大小

参数

fifo

kfifo 队列结构体的指针

返回值

缓冲区最大大小

kfifo_size

```
kfifo_size(fifo)
```

描述

获取 kfifo 缓冲区当前已使用的大小

参数

fifo

kfifo 队列结构体的指针

返回值

缓冲区当前已使用的大小

kfifo_empty

```
kfifo_empty(fifo)
```

描述

判断 kfifo 缓冲区当前是否为空

参数

fifo

kfifo 队列结构体的指针

返回值

情况	返回值
空	1
非空	0

kfifo_full

```
kfifo_full(fifo)
```

描述

判断 kfifo 缓冲区当前是否为满

参数

fifo

kfifo 队列结构体的指针

返回值

情况	返回值
满	1
不满	0

5.3.2 ID Allocation

ida 的主要作用是分配 + 管理 id. 它能分配一个最小的, 未被分配出去的 id. 当您需要管理某个数据结构时, 可能需要使用 id 来区分不同的目标. 这个时候, ida 将会是很好的选择. 因为 ida 的十分高效, 运行常数相对数组更小, 而且提供了基本管理 id 需要用到的功能, 值得您试一试.

IDA 定义于 idr.h 文件中. 您通过 DECLARE_IDA(my_ida) 来创建一个 ida 对象, 或者 struct ida my_ida; ida_init(&my_ida); 来初始化一个 ida.

ida_init

```
void ida_init(struct ida *ida_p)
```

描述

通初始化 IDA, 你需要保证调用函数之前, ida 的 free_list 为空, 否则会导致内存泄漏.

参数

ida_p

指向 ida 的指针

返回值

无返回值

ida_preload

```
int ida_preload(struct ida *ida_p, gfp_t gfp_mask)
```

描述

为 ida 预分配空间. 您可以不自行调用, 因为当 ida 需要空间的时候, 内部会自行使用 `kmalloc` 函数获取空间. 当然, 设计这个函数的目的是为了让您有更多的选择. 当您提前调用这个函数, 可以避免之后在开辟空间上的时间开销.

参数

ida_p

指向 ida 的指针

gfp_mask

保留参数, 目前尚未使用.

返回值

如果分配成功, 将返回 0; 否则返回负数错误码, 有可能是内存空间不够.

ida_alloc

```
int ida_alloc(struct ida *ida_p, int *p_id)
```

描述

获取一个空闲 ID. 您需要注意, 返回值是成功/错误码.

参数

ida_p

指向 ida 的指针

p_id

您需要传入一个 int 变量的指针, 如果成功分配 ID, ID 将会存储在该指针所指向的地址.

返回值

如果分配成功, 将返回 0; 否则返回负数错误码, 有可能是内存空间不够.

ida_count

```
bool ida_count(struct ida *ida_p, int id)
```

描述

查询一个 ID 是否被分配.

参数

ida_p

指向 ida 的指针

id

您查询该 ID 是否被分配.

返回值

如果分配, 将返回 true; 否则返回 false.

ida_remove

```
void ida_remove(struct ida *ida_p, int id)
```

描述

删除一个已经分配的 ID. 如果该 ID 不存在, 该函数不会产生异常错误, 因为在检测到该 ID 不存在的时候, 函数将会自动退出.

参数

ida_p

指向 ida 的指针

id

您要删除的 id.

返回值

无返回值.

ida_destroy

```
void ida_destroy(struct ida *ida_p)
```

描述

释放一个 IDA 所有的空间, 同时删除 ida 的所有已经分配的 id.(所以您不用担心删除 id 之后, ida 还会占用大量空间.)

参数

ida_p

指向 ida 的指针

返回值

无返回值

ida_empty

```
void ida_empty(struct ida *ida_p)
```

描述

查询一个 ida 是否为空

参数

ida_p

指向 ida 的指针

返回值

ida 为空则返回 true，否则返回 false。

5.3.3 IDR

idr 是一个基于 radix-tree 的 ID-pointer 的数据结构. 该数据结构提供了建 id 与数据指针绑定的功能, 它的主要功能有以下 4 个:

1. 获取一个 ID, 并且将该 ID 与一个指针绑定
2. 删除一个已分配的 ID
3. 根据 ID 查找对应的指针
4. 根据 ID 使用新的 ptr 替换旧的 ptr

您可以使用 `DECLARE_idr(my_idr)` 来创建一个 idr。或者您也可以使用 `struct idr my_idr; idr_init(my_idr);` 这两句话创建一个 idr。至于什么是 radix-tree, 您可以把他简单理解为一个向上生长的多叉树, 在实现中, 我们选取了 64 叉树。

idr_init

```
void idr_init(struct idr *idp)
```

描述

通初始化 IDR, 你需要保证调用函数之前, idr 的 free_list 为空, 否则会导致内存泄漏.

参数

idp

指向 idr 的指针

返回值

无返回值

idr_preload

```
int idr_preload(struct idr *idp, gfp_t gfp_mask)
```

描述

为 idr 预分配空间. 您可以不自行调用, 因为当 idr 需要空间的时候, 内部会自行使用 kmalloc 函数获取空间. 当然, 设计这个函数的目的是为了让您有更多的选择. 当您提前调用这个函数, 可以避免之后在开辟空间上的时间开销.

参数

idp

指向 idr 的指针

gfp_mask

保留参数, 目前尚未使用.

返回值

如果分配成功, 将返回 0; 否则返回负数错误码, 有可能是内存空间不够.

idr_alloc

```
int idr_alloc(struct idr *idp, void *ptr, int *id)
```

描述

获取一个空闲 ID. 您需要注意, 返回值是成功/错误码. 调用这个函数, 需要您保证 `ptr` 是非空的, 即: `ptr != NULL`, 否则将会影响 `idr_find/idr_find_next/idr_find_next_getid/...` 等函数的使用. (具体请看这三个函数的说明, 当然, 只会影响到您的使用体验, 并不会影响到 `idr` 内部函数的决策和逻辑)

参数

idp

指向 `ida` 的指针

ptr

指向数据的指针

id

您需要传入一个 `int` 变量的指针, 如果成功分配 ID, ID 将会存储在该指针所指向的地址.

返回值

如果分配成功, 将返回 0; 否则返回负数错误码, 有可能是内存空间不够.

idr_remove

```
void* idr_remove(struct idr *idp, int id)
```

描述

删除一个 id, 但是不释放对应的 ptr 指向的空间, 同时返回这个被删除 id 所对应的 ptr。 如果该 ID 不存在, 该函数不会产生异常错误, 因为在检测到该 ID 不存在的时候, 函数将会自动退出, 并返回 NULL。

参数

idp

指向 idr 的指针

id

您要删除的 id.

返回值

如果删除成功, 就返回被删除 id 所对应的 ptr; 否则返回 NULL。注意: 如果这个 id 本来就和 NULL 绑定, 那么也会返回 NULL

idr_remove_all

```
void idr_remove_all(struct idr *idp)
```

描述

删除 idr 的所有已经分配的 id.(所以您不用担心删除 id 之后, idr 还会占用大量空间。)

但是您需要注意的是, 调用这个函数是不会释放数据指针指向的空间的。所以您调用该函数之前, 确保 IDR 内部的数据指针被保存。否则当 IDR 删除所有 ID 之后, 将会造成内存泄漏。

参数

idp

指向 idr 的指针

返回值

无返回值

idr_destroy

```
void idr_destroy(struct idr *idp)
```

描述

释放一个 IDR 所有的空间, 同时删除 idr 的所有已经分配的 id.(所以您不用担心删除 id 之后, ida 还会占用大量空间.) - 和 `idr_remove_all` 的区别是, 释放掉所有的空间 (包括 `free_list` 的预分配空间)。

参数

idp

指向 idr 的指针

返回值

无返回值

idr_find

```
void *idr_find(struct idr *idp, int id)
```

描述

查询一个 ID 所绑定的数据指针

参数

idp

指向 idr 的指针

id

您查询该 ID 的数据指针

返回值

如果分配, 将返回该 ID 对应的数据指针; 否则返回 NULL.(注意, 返回 NULL 不一定代表这 ID 不存在, 有可能该 ID 就是与空指针绑定。) 当然, 我们也提供了 `idr_count` 函数来判断 id 是否被分配, 具体请查看 `idr_count` 介绍。

idr_find_next

```
void *idr_find_next(struct idr *idp, int start_id)
```

描述

传进一个 `start_id`, 返回满足 “id 大于 `start_id` 的最小 id” 所对应的数据指针。

参数

idp

指向 idr 的指针

start_id

您提供的 ID 限制

返回值

如果分配, 将返回该 ID 对应的数据指针; 否则返回 NULL.(注意, 返回 NULL 不一定代表这 ID 不存在, 有可能该 ID 就是与空指针绑定。) 当然, 我们也提供了 `idr_count` 函数来判断 id 是否被分配, 具体请查看 `idr_count` 介绍。

idr_find_next_getid

```
void *idr_find_next_getid(struct idr *idp, int start_id, int *nextid)
```

描述

传进一个 `start_id`, 返回满足 “id 大于 `start_id` 的最小 id” 所对应的数据指针。同时, 你获取到这个满足条件的最小 id, 即参数中的 `*nextid`。

参数

idp

指向 idr 的指针

start_id

您提供的 ID 限制

返回值

如果分配, 将返回该 ID 对应的数据指针; 否则返回 NULL.(注意, 返回 NULL 不一定代表这 ID 不存在, 有可能该 ID 就是与空指针绑定。) 当然, 我们也提供了 `idr_count` 函数来判断 id 是否被分配, 具体请查看 `idr_count` 介绍。

idr_replace

```
int idr_replace(struct idr *idp, void *ptr, int id)
```

描述

传进一个 ptr, 使用该 ptr 替换掉 id 所对应的 Old_ptr。

参数

idp

指向 idr 的指针

ptr

您要替换原来的 old_ptr 的新指针

id

您要替换的指针所对应的 id

返回值

0 代表成功，否则就是错误码 - 代表错误。

idr_replace_get_old

```
int idr_replace_get_old(struct idr *idp, void *ptr, int id, void **oldptr)
```

描述

传进一个 ptr，使用该 ptr 替换掉 id 所对应的 Old_ptr，同时你可以获取到 old_ptr。

参数

idp

指向 idr 的指针

ptr

您要替换原来的 old_ptr 的新指针

id

您要替换的指针所对应的 id

old_ptr

您需要传进该 (void**) 指针，old_ptr 将会存放在该指针所指向的地址。

返回值

0 代表成功，否则就是错误码 - 代表错误。

idr_empty

```
void idr_empty(struct idr *idp)
```

描述

查询一个 idr 是否为空

参数

idp

指向 idr 的指针

返回值

idr 为空则返回 true，否则返回 false。

idr_count

```
bool idr_count(struct idr *idp, int id)
```

描述

查询一个 ID 是否被分配.

参数

ida_p

指向 idr 的指针

id

您查询该 ID 是否被分配.

返回值

如果分配, 将返回 true; 否则返回 false.

5.4 内存管理

这里快速讲解了如何在 DragonOS 中分配、使用内存。以便您能快速的了解这个模块。

详细的内存管理模块的文档请参见：[内存管理文档](#)

5.4.1 内存分配指南

DragonOS 提供了一些用于内存分配的 api。您可以使用 *kmalloc* 来分配小的内存块，也可以使用 *alloc_pages* 分配连续的 2MB 大小的内存页面。

选择合适的内存分配器

在内核中，最直接、最简单的分配内存的方式就是，使用 *kmalloc()* 函数进行分配。并且，出于安全起见，除非内存存在分配后一定会被覆盖，且您能确保内存中的脏数据不会对程序造成影响，在其余情况下，我们建议使用 *kzalloc()* 进行内存分配，它将会在 *kmalloc()* 的基础上，把申请到的内存进行清零。

您可以通过 *kmalloc()* 函数分配得到 32bytes 到 1MBytes 之间的内存对象。并且，这些内存对象具有以下性质：

- 内存起始地址及大小按照 2 次幂对齐。（比如，申请的是 80bytes 的内存空间，那么获得的内存对象大小为 128bytes 且内存地址按照 128bytes 对齐）

对于需要大量连续内存的分配，可以使用 *alloc_pages()* 向页面分配器申请连续的内存页。

当内存空间不再被使用时，那么必须释放他们。若您使用的是 *kmalloc()* 分配的内存，那么您需要使用 *kfree()* 释放它。若是使用 *alloc_pages()* 分配的内存，则需要使用 *free_pages()* 来释放它们。

5.4.2 内存管理 API

SLAB 内存池

SLAB 内存池提供小内存对象的分配功能。

```
void *kmalloc(unsigned long size, gfp_t gfp)
```

获取小块的内存。

描述

`kmalloc` 用于获取那些小于 2M 内存页大小的内存对象。可分配的内存对象大小为 32bytes~1MBytes. 且分配的内存块大小、起始地址按照 2 的 `n` 次幂进行对齐。(比如, 申请的是 80bytes 的内存空间, 那么获得的内存对象大小为 128bytes 且内存地址按照 128bytes 对齐)

参数

size

内存对象的大小

gfp

标志位

```
void *kzalloc(unsigned long size, gfp_t gfp)
```

描述

获取小块的内存, 并将其清零。其余功能与 `kmalloc` 相同。

参数

size

内存对象的大小

gfp

标志位

```
unsigned long kfree(void *address)
```

释放从 slab 分配的内存。

描述

该函数用于释放通过 `kmalloc` 申请的内存。如果 `address` 为 `NULL`, 则函数被调用后, 无事发生。

请不要通过这个函数释放那些不是从 `kmalloc()` 或 `kzalloc()` 申请的内存, 否则将会导致系统崩溃。

参数

address

指向内存对象的起始地址的指针

物理页管理

DragonOS 支持对物理页的直接操作

```
struct Page *alloc_pages(unsigned int zone_select, int num, ul flags)
```

描述

从物理页管理单元中申请一段连续的物理页

参数

zone_select

要申请的物理页所位于的内存区域

可选值：

- ZONE_DMA DMA 映射专用区域
- ZONE_NORMAL 正常的物理内存区域，已在页表高地址处映射
- ZONE_UNMAPPED_IN_PGT 尚未在页表中映射的区域

num

要申请的连续物理页的数目，该值应当小于 64

flags

分配的页面要被设置成的属性

可选值：

- PAGE_PGT_MAPPED 页面在页表中已被映射
- PAGE_KERNEL_INIT 内核初始化所占用的页
- PAGE_DEVICE 设备 MMIO 映射的内存
- PAGE_KERNEL 内核层页
- PAGE_SHARED 共享页

返回值

成功

成功申请则返回指向起始页面的 Page 结构体的指针

失败

当 ZONE 错误或内存不足时, 返回 NULL

```
void free_pages(struct Page *page, int number)
```

描述

从物理页管理单元中释放一段连续的物理页。

参数

page

要释放的第一个物理页的 Page 结构体

number

要释放的连续内存页的数量。该值应小于 64

页表管理

```
int mm_map_phys_addr(ul virt_addr_start, ul phys_addr_start, ul length, ul  
flags, bool use4k)
```

描述

将一段物理地址映射到当前页表的指定虚拟地址处

参数

virt_addr_start

虚拟地址的起始地址

phys_addr_start

物理地址的起始地址

length

要映射的地址空间的长度

flags

页表项的属性

use4k

使用 4 级页表，将地址区域映射为若干 4K 页

```
int mm_map_proc_page_table(ul proc_page_table_addr, bool is_phys, ul
virt_addr_start, ul phys_addr_start, ul length, ul flags, bool user, bool
flush, bool use4k)
```

描述

将一段物理地址映射到指定页表的指定虚拟地址处

参数

proc_page_table_addr

指定的顶层页表的起始地址

is_phys

该顶层页表地址是否为物理地址

virt_addr_start

虚拟地址的起始地址

phys_addr_start

物理地址的起始地址

length

要映射的地址空间的长度

flags

页表项的属性

user

页面是否为用户态可访问

flush

完成映射后，是否刷新 TLB

use4k

使用 4 级页表，将地址区域映射为若干 4K 页

返回值

- 映射成功：0
- 映射失败：-EFAULT

```
void mm_unmap_proc_table(ul proc_page_table_addr, bool is_phys, ul  
virt_addr_start, ul length)
```

描述

取消给定页表中的指定地址空间的页表项映射。

参数

proc_page_table_addr

指定的顶层页表的基地址

is_phys

该顶层页表地址是否为物理地址

virt_addr_start

虚拟地址的起始地址

length

要取消映射的地址空间的长度

```
mm_unmap_addr(virt_addr, length)
```

描述

该宏定义用于取消当前进程的页表中的指定地址空间的页表项映射。

参数

virt_addr

虚拟地址的起始地址

length

要取消映射的地址空间的长度

内存信息获取

```
struct mm_stat_t mm_stat()
```

描述

获取计算机目前的内存空间使用情况

参数

无

返回值

返回值是一个 mm_mstat_t 结构体，该结构体定义于 mm/mm.h 中。其中包含了以下信息（单位均为字节）：

参数名	解释
total	计算机的总内存数量大小
used	已使用的内存大小
free	空闲物理页所占的内存大小
shared	共享的内存大小
cache_used	位于 slab 缓冲区中的已使用的内存大小
cache_free	位于 slab 缓冲区中的空闲的内存大小
available	系统总空闲内存大小（包括 kmalloc 缓冲区）

这里是 DragonOS 的锁变量的说明文档。

6.1 锁的类型及其规则

6.1.1 简介

DragonOS 内核实现了一些锁，大致可以分为两类：

- 休眠锁
- 自旋锁

6.1.2 锁的类型

休眠锁

休眠锁只能在可抢占的上下文之中被获取。

在 DragonOS 之中，实现了以下的休眠锁：

- semaphore
- mutex_t

自旋锁

- `spinlock_t`
- *RawSpinLock* (Rust 版本的 `spinlock_t`, 但与 `spinlock_t` 不兼容)
- *SpinLock* ——在 *RawSpinLock* 的基础上, 封装了一层守卫 (Guard), 将锁及其要保护到的数据绑定在一个结构体内, 并能在编译期避免未加锁就访问数据的问题。

进程在获取自旋锁后, 将改变 `pcb` 中的锁变量持有计数, 从而隐式地禁止了抢占。为了获得更多灵活的操作, `spinlock` 还提供了以下的方法:

后缀	说明
<code>_irq()</code>	在加锁时关闭中断/在放锁时开启中断
<code>_irqsave()/_irqrestore()</code>	在加锁时保存中断状态, 并关中断/在放锁时恢复中断状态

当您同时需要使用自旋锁以及引用计数时, 一个好的方法是: 使用 `lockref`. 这是一种额外的加速技术, 能额外提供“无锁修改引用计数”的功能。详情请见: *lockref*

6.1.3 详细介绍

自旋锁的详细介绍

关于自旋锁的详细介绍, 请见文档: *自旋锁*

semaphore 信号量

semaphore 信号量是基于计数实现的。

当可用资源不足时, 尝试对 semaphore 执行 down 操作的进程将会被休眠, 直到资源可用。

mutex 互斥量

请见*Mutex* 文档

备注: 作者: 龙进 longjin@RinGoTek.cn

6.2 自旋锁

6.2.1 1. 简介

自旋锁是用于多线程同步的一种锁，线程反复检查锁变量是否可用。由于线程在这一过程中保持运行的状态，因此是一种忙等待。一旦获取了自旋锁，线程会一直保持该锁，直至显式释放自旋锁。

DragonOS 在 `kernel/src/lib/spinlock.rs` 文件中，实现了自旋锁。根据功能特性的略微差异，分别提供了 `RawSpinLock` 和 `SpinLock` 两种自旋锁。

6.2.2 2. RawSpinLock - 原始自旋锁

`RawSpinLock` 是原始的自旋锁，其数据部分包含一个 `AtomicBool`，实现了自旋锁的基本功能。其加锁、放锁需要手动确定对应的时机，也就是说，和我们在其他语言中使用的自旋锁一样，需要先调用 `lock()` 方法，然后当离开临界区时，手动调用 `unlock()` 方法。我们并没有向编译器显式地指定该自旋锁到底保护的是哪些数据。

`RawSpinLock` 为程序员提供了非常自由的加锁、放锁控制。但是，正是由于它过于自由，因此在使用它的时候，我们很容易出错。很容易出现“未加锁就访问临界区的数据”、“忘记放锁”、“双重释放”等问题。当使用 `RawSpinLock` 时，编译器并不能对这些情况进行检查，这些问题只能在运行时被发现。

警告： `RawSpinLock` 与 C 版本的 `spinlock_t` 不具有二进制兼容性。如果由于暂时的兼容性的需求，要操作 C 版本的 `spinlock_t`，请使用 `spinlock.rs` 中提供的 C 版本的 `spinlock_t` 的操作函数。

但是，对于新开发的功能，请不要使用 C 版本的 `spinlock_t`，因为随着代码重构的进行，我们将会移除它。

6.2.3 3. SpinLock - 具备守卫的自旋锁

`SpinLock` 在 `RawSpinLock` 的基础上，进行了封装，能够在编译期检查出“未加锁就访问临界区的数据”、“忘记放锁”、“双重释放”等问题；并且，支持数据的内部可变性。

其结构体原型如下：

```
#[derive(Debug)]
pub struct SpinLock<T> {
    lock: RawSpinlock,
    /// 自旋锁保护的数据
    data: UnsafeCell<T>,
}
```

3.1. 使用方法

您可以这样初始化一个 SpinLock:

```
let x = SpinLock::new(Vec::new());
```

在初始化这个 SpinLock 时, 必须把要保护的数据传入 SpinLock, 由 SpinLock 进行管理。

当需要读取、修改 SpinLock 保护的数据时, 请先使用 SpinLock 的 lock() 方法。该方法会返回一个 SpinLockGuard。您可以使用被保护的数据的成员函数来进行一些操作。或者是直接读取、写入被保护的数据。(相当于您获得了被保护的数据的可变引用)

完整示例如下方代码所示:

```
let x : SpinLock<Vec<i32>> = SpinLock::new(Vec::new());
{
    let mut g : SpinLockGuard<Vec<i32>> = x.lock();
    g.push(1);
    g.push(2);
    g.push(2);
    assert!(g.as_slice() == [1, 2, 2] || g.as_slice() == [2, 2, 1]);
    // 在此处, SpinLock是加锁的状态
    kdebug!("x={:?}", x);
}
//_
→ 由于上方的变量`g`, 也就是SpinLock守卫的生命周期结束, 自动释放了SpinLock。因此, 在此处, SpinLock是放
```

```
kdebug!("x={:?}", x);
```

对于结构体内部的变量, 我们可以使用 SpinLock 进行细粒度的加锁, 也就是使用 SpinLock 包裹需要细致加锁的成员变量, 比如这样:

```
pub struct a {
    pub data: SpinLock<data_struct>,
}
```

当然, 我们也可以对整个结构体进行加锁:

```
struct MyStruct {
    pub data: data_struct,
}
/// 被全局加锁的结构体
pub struct LockedMyStruct (SpinLock<MyStruct>);
```

3.2. 原理

SpinLock 之所以能够实现编译期检查，是因为它引入了一个 SpinLockGuard 作为守卫。我们在编写代码的时候，保证只有调用 SpinLock 的 lock() 方法加锁后，才能生成一个 SpinLockGuard。并且，当我们想要访问受保护的数据的时候，都必须获得一个守卫。然后，我们为 SpinLockGuard 实现了 Drop trait，当守卫的生命周期结束时，将会自动释放锁。除此以外，没有别的方法能够释放锁。因此我们能够得知，一个上下文中，只要 SpinLockGuard 的生命周期没有结束，那么它就拥有临界区数据的访问权，数据访问就是安全的。

3.3. 存在的问题

3.3.1. 双重加锁

请注意，SpinLock 支持的编译期检查并不是万能的。它目前无法在编译期检查出“双重加锁”问题。试看这样一个场景：函数 A 中，获得了锁。然后函数 B 中继续尝试加锁，那么就造成了“双重加锁”问题。这样在编译期是无法检测出来的。

针对这个问题，我们建议采用这样的编程方法：

- 如果函数 B 需要访问临界区内的数据，那么，函数 B 应当接收一个类型为 &SpinLockGuard 的参数，这个守卫由函数 A 获得。这样一来，函数 B 就能访问临界区内的数据。

6.3 lockref

lockref 是将自旋锁与引用计数变量融合在连续、对齐的 8 字节内的一种技术。

目前，DragonOS 中，通过 C、Rust 各实现了一个版本的 lockref。请注意，二者不兼容。对于新的功能模块，请使用 Rust 版本的 lockref。随着代码重构工作的进行，我们将会删除 C 版本的 lockref。

6.3.1 1. lockref 结构

1.1. Rust 版本

```
/// 仅在x86_64架构下使用cmpxchg
#[cfg(target_arch = "x86_64")]
/// 由于需要cmpxchg，所以整个lockref按照8字节对齐
#[repr(align(8))]
#[derive(Debug)]
pub struct LockRef {
    pub lock: RawSpinlock,
    pub count: i32,
}
```

(下页继续)

(续上页)

```

/// 除了x86_64以外的架构，不使用cmpxchg进行优化
#[cfg(not(target_arch = "x86_64"))]
pub struct LockRef {
    lock: RawSpinlock,
    count: i32,
}

```

1.2. C 版本

```

struct lockref
{
    union
    {
#ifdef __LOCKREF_ENABLE_CMPXCHG__
        aligned_u64 lock_count; //通过该变量的声明，使得整个lockref的地址按照8字节对齐
#endif

        struct
        {
            spinlock_t lock;
            int count;
        };
    };
};

```

6.3.2 2. 特性描述

由于在高负载的情况下，系统会频繁的执行“锁定-改变引用变量-解锁”的操作，这期间很可能会出现 spinlock 和引用计数跨缓存行的情况，这将会大大降低性能。lockref 通过强制对齐，尽可能的降低缓存行的占用数量，使得性能得到提升。

并且，在 x64 体系结构下，还通过 cmpxchg() 指令，实现了无锁快速路径。不需要对自旋锁加锁即可更改引用计数的值，进一步提升性能。当快速路径不存在（对于未支持的体系结构）或者尝试超时后，将会退化成“锁定-改变引用变量-解锁”的操作。此时由于 lockref 强制对齐，只涉及到 1 个缓存行，因此性能比原先的 spinlock+ref_count 的模式要高。

6.3.3 3. 关于 cmpxchg_loop

在改变引用计数时，cmpxchg 先确保没有别的线程持有锁，然后改变引用计数，同时通过 lock cmpxchg 指令验证在更改发生时，没有其他线程持有锁，并且当前的目标 lockref 的值与 old 变量中存储的一致，从而将新值存储到目标 lockref。这种无锁操作能极大的提升性能。如果不符合上述条件，在多次尝试后，将退化成传统的加锁方式来更改引用计数。

6.3.4 4. Rust 版本的 API

4.1. 引用计数自增

- pub fn inc(&mut self)
- pub fn inc_not_zero(&mut self) -> Result<i32, i32>
- pub fn inc_not_dead(&mut self) -> Result<i32, i32>

4.1.1. inc

说明

原子的将引用计数加 1。

返回值

无

4.1.2. inc_not_zero

说明

原子地将引用计数加 1。如果原来的 count≤0，则操作失败。

返回值

返回值	说明
Ok(self.count)	成功，返回新的引用计数
Err(-1)	失败，返回-1

4.1.3. inc_not_dead

说明

引用计数自增 1。(除非该 lockref 已经被标记为死亡)

返回值

返回值	说明
Ok(self.count)	成功, 返回新的引用计数
Err(-1)	失败, 返回-1

4.2. 引用计数自减

- `pub fn dec(&mut self) -> Result<i32, i32>`
- `pub fn dec_return(&mut self) -> Result<i32, i32>`
- `pub fn dec_not_zero(&mut self) -> Result<i32, i32>`
- `pub fn dec_or_lock_not_zero(&mut self) -> Result<i32, i32>`

4.2.1. dec

说明

原子地将引用计数-1。如果已处于 `count≤0` 的状态, 则返回 `Err(-1)`

本函数与 `lockref_dec_return()` 的区别在于, 当在 `cmpxchg()` 中检测到 `count≤0` 或已加锁, 本函数会再次尝试通过加锁来执行操作, 而 `lockref_dec_return()` 会直接返回错误

返回值

返回值	说明
Ok(self.count)	成功, 返回新的引用计数
Err(-1)	失败, 返回-1

4.2.2. dec_return

原子地将引用计数减 1。如果处于已加锁或 `count ≤ 0` 的状态，则返回 -1

本函数与 `lockref_dec()` 的区别在于，当在 `cmpxchg()` 中检测到 `count ≤ 0` 或已加锁，本函数会直接返回错误，而 `lockref_dec()` 会再次尝试通过加锁来执行操作。

备注：若当前处理器架构不支持 `cmpxchg`，则退化为 `self.dec()`

返回值

返回值	说明
Ok(self.count)	成功，返回新的引用计数
Err(-1)	失败，返回 -1

4.2.3. dec_not_zero

说明

原子地将引用计数减 1。若当前的引用计数 `≤ 1`，则操作失败。

该函数与 `lockref_dec_or_lock_not_zero()` 的区别在于，当 `cmpxchg()` 时发现 `old.count ≤ 1` 时，该函数会直接返回 `Err(-1)`，而 `lockref_dec_or_lock_not_zero()` 在这种情况下，会尝试加锁来进行操作。

返回值

返回值	说明
Ok(self.count)	成功，返回新的引用计数
Err(-1)	失败，返回 -1

4.2.4. dec_or_lock_not_zero

说明

原子地将引用计数减 1。若当前的引用计数 ≤ 1 ，则操作失败。

该函数与 `lockref_dec_not_zero()` 的区别在于，当 `cmpxchg()` 时发现 `old.count ≤ 1` 时，该函数会尝试加锁来进行操作，而 `lockref_dec_not_zero()` 在这种情况下，会直接返回 `Err(-1)`。

返回值

返回值	说明
<code>Ok(self.count)</code>	成功，返回新的引用计数
<code>Err(-1)</code>	失败，返回-1

4.3. 其他

- `pub fn mark_dead(&mut self)`

4.3.1. mark_dead

说明

将引用计数原子地标记为死亡状态。

6.3.5 参考资料

[Introducing lockrefs - LWN.net](#), Jonathan Corbet

备注：作者：龙进 longjin@RinGoTek.cn

6.4 mutex 互斥量

mutex 是一种轻量级的同步原语，只有被加锁、空闲两种状态。

当 mutex 被占用时，尝试对 mutex 进行加锁操作的进程将会被休眠，直到资源可用。

6.4.1 1. 特性

- 同一时间只有 1 个任务可以持有 mutex
- 不允许递归地加锁、解锁
- 只允许通过 mutex 的 api 来操作 mutex
- 在硬中断、软中断中不能使用 mutex

6.4.2 2. 定义

mutex 定义在 lib/mutex.rs 中，定义如下所示：

```
/// @brief Mutex互斥量结构体
///_
→ 请注意！由于Mutex属于休眠锁，因此，如果您的代码可能在中断上下文内执行，请勿采用Mutex!
#[derive(Debug)]
pub struct Mutex<T> {
    /// 该Mutex保护的数据
    data: UnsafeCell<T>,
    /// Mutex内部的信息
    inner: SpinLock<MutexInner>,
}

#[derive(Debug)]
struct MutexInner {
    /// 当前Mutex是否已经被上锁(上锁时，为true)
    is_locked: bool,
    /// 等待获得这个锁的进程的链表
    wait_list: LinkedList<static mut process_control_block>,
}
```

6.4.3 3. 使用

与 SpinLock 类似，Rust 版本的 Mutex 具有一个守卫。使用的时候，需要将要被保护的数据的所有权移交 Mutex。并且，守卫只能在加锁成功后产生，因此，每个时刻，每个 Mutex 最多存在 1 个守卫。

当需要读取、修改 Mutex 保护的数据时，请先使用 Mutex 的 lock() 方法。该方法会返回一个 MutexGuard。您可以使用被保护的数据的成员函数来进行一些操作。或者是直接读取、写入被保护的数据。（相当于您获得了被保护的数据的可变引用）

完整示例如下方代码所示：

```
let x :Mutex<Vec<i32>>= Mutex::new(Vec::new());
{
    let mut g :MutexGuard<Vec<i32>>= x.lock();
    g.push(1);
    g.push(2);
    g.push(2);
    assert!(g.as_slice() == [1, 2, 2] || g.as_slice() == [2, 2, 1]);
    // 在此处，Mutex是加锁的状态
    kdebug!("x={:?}", x);
}
//_
```

→ 由于上方的变量`g`，也就是Mutex守卫的生命周期结束，自动释放了Mutex。因此，在此处，Mutex是放锁的状态

```
kdebug!("x={:?}", x);
```

对于结构体内部的变量，我们可以使用 Mutex 进行细粒度的加锁，也就是使用 Mutex 包裹需要细致加锁的成员变量，比如这样：

```
pub struct a {
    pub data: Mutex<data_struct>,
}
```

当然，我们也可以对整个结构体进行加锁：

```
struct MyStruct {
    pub data: data_struct,
}
/// 被全局加锁的结构体
pub struct LockedMyStruct (Mutex<MyStruct>);
```

6.4.4 4. API

4.1. new - 初始化 Mutex

原型

```
pub const fn new(value: T) -> Self
```

说明

`new()` 方法用于初始化一个 `Mutex`。该方法需要一个被保护的数据作为参数。并且，该方法会返回一个 `Mutex`。

4.2. lock - 加锁

原型

```
pub fn lock(&self) -> MutexGuard<T>
```

说明

对 `Mutex` 加锁，返回 `Mutex` 的守卫，您可以使用这个守卫来操作被保护的数据。

如果 `Mutex` 已经被加锁，那么，该方法会阻塞当前进程，直到 `Mutex` 被释放。

4.3. try_lock - 尝试加锁

原型

```
pub fn try_lock(&self) -> Result<MutexGuard<T>, i32>
```

说明

尝试对 `Mutex` 加锁。如果加锁失败，不会将当前进程加入等待队列。如果加锁成功，返回 `Mutex` 的守卫；如果当前 `Mutex` 已经被加锁，返回 `Err`(错误码)。

6.4.5 5. C 版本的 Mutex（在将来会被废弃）

mutex 定义在 `common/mutex.h` 中。其数据类型如下所示：

```
typedef struct
{
    atomic_t count; // 锁计数。1->已解锁。 0->已上锁,且有可能存在等待者
    spinlock_t wait_lock; // mutex操作锁,用于对mutex的list的操作进行加锁
    struct List wait_list; // Mutex的等待队列
} mutex_t;
```

5.1. API

mutex_init

```
void mutex_init(mutex_t *lock)
```

初始化一个 mutex 对象。

mutex_lock

```
void mutex_lock(mutex_t *lock)
```

对一个 mutex 对象加锁。若 mutex 当前被其他进程持有，则当前进程进入休眠状态。

mutex_unlock

```
void mutex_unlock(mutex_t *lock)
```

对一个 mutex 对象解锁。若 mutex 的等待队列中有其他的进程，则唤醒下一个进程。

mutex_trylock

```
void mutex_trylock(mutex_t *lock)
```

尝试对一个 mutex 对象加锁。若 mutex 当前被其他进程持有，则返回 0. 否则，加锁成功，返回 1.

mutex_is_locked

```
void mutex_is_locked(mutex_t *lock)
```

判断 mutex 是否已被加锁。若给定的 mutex 已处于上锁状态，则返回 1，否则返回 0。

6.5 RwLock 读写锁

备注：本文作者: sujintao

Email: sujintao@dragonos.org

6.5.1 1. 简介

读写锁是一种在并发环境下保护多进程间共享数据的机制。相比于普通的 spinlock，读写锁将对共享数据的访问分为读和写两种类型：只读取共享数据的访问使用读锁控制，修改共享数据的访问使用写锁控制。读写锁设计允许同时存在多个“读者”（只读取共享数据的访问）和一个“写者”（修改共享数据的访问），对于一些大部分情况都是读访问的共享数据来说，使用读写锁控制访问可以一定程度上提升性能。

6.5.2 2. DragonOS 中读写锁的实现

2.1 读写锁的机理

读写锁的目的是维护多线程系统中的共享变量的一致性。数据会被包裹在一个 RwLock 的数据结构中，一切的访问必须通过 RwLock 的数据结构进行访问和修改。每个要访问共享数据的会获得一个守卫 (guard)，只读进程获得 READER(读者守卫)，需要修改共享变量的进程获得 WRITER(写者守卫)，作为 RwLock 的“影子”，线程都根据 guard 来进行访问和修改操作。

在实践中，读写锁除了 READER, WRITER，还增加了 UPGRADER；这是一种介于 READER 和 WRITER 之间的守卫，这个守卫的作用就是防止 WRITER 的饿死 (Starvation)。当进程获得 UPGRADER 时，进程把它当成 READER 来使用；但是 UPGRADER 可以进行升级处理，升级后的 UPGRADER 相当于是一个 WRITER 守卫，可以对共享数据执行写操作。

所有守卫都满足 rust 原生的 RAII 机理，当守卫所在的作用域结束时，守卫将自动释放。

2.2 读写锁守卫之间的关系

同一时间点, 可以存在多个 READER, 即可以同时有多个进程对共享数据进行访问; 同一时间只能存在一个 WRITER, 而且当有一个进程获得 WRITER 时, 不能存在 READER 和 UPGRADER; 进程获得 UPGRADER 的前提条件是, 不能有 UPGRADER 或 WRITER 存在, 但是当有一个进程获得 UPGRADER 时, 进程无法成功申请 READER.

2.3 设计的细节

2.3.1 RwLock 数据结构

```
pub struct RwLock<T> {  
    lock: AtomicU32, //原子变量  
    data: UnsafeCell<T>,  
}
```

2.3.2 READER 守卫的数据结构

```
pub struct RwLockReadGuard<'a, T: 'a> {  
    data: *const T,  
    lock: &'a AtomicU32,  
}
```

2.3.3 UPGRADER 守卫的数据结构

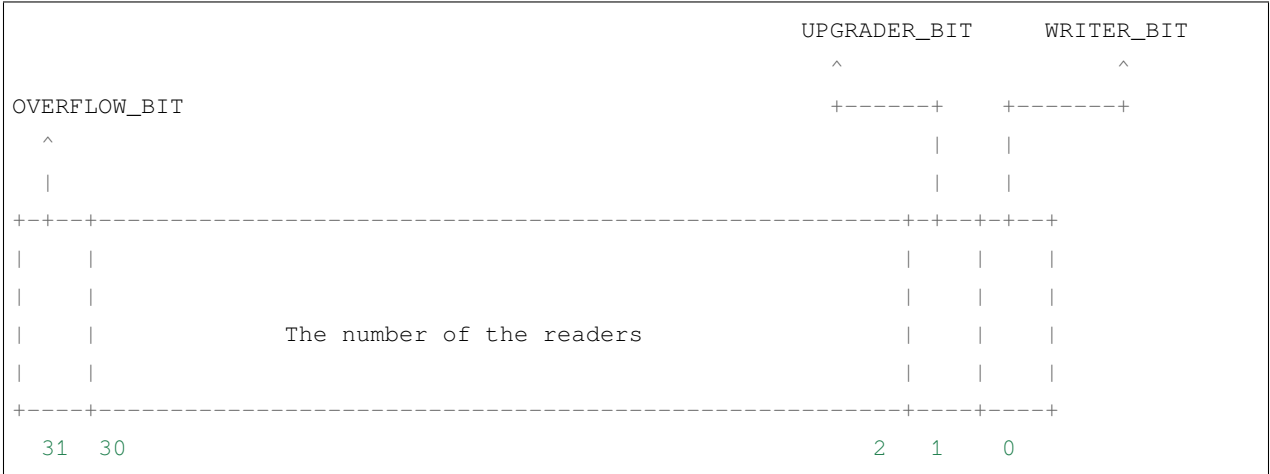
```
pub struct RwLockUpgradableGuard<'a, T: 'a> {  
    data: *const T,  
    inner: &'a RwLock<T>,  
}
```

2.3.4 WRITER 守卫的数据结构

```
pub struct RwLockWriteGuard<'a, T: 'a> {  
    data: *mut T,  
    inner: &'a RwLock<T>,  
}
```


2.3.5 RwLock 的 lock 的结构介绍

lock 是一个 32 位原子变量 AtomicU32, 它的比特位分配如下:



(从右到左) 第 0 位表征 WRITER 是否有效, 若 WRITER_BIT=1, 则存在一个进程获得了 WRITER 守卫; 若 UPGRADER_BIT=1, 则存在一个进程获得了 UPGRADER 守卫, 第 2 位到第 30 位用来二进制表示获得 READER 守卫的进程数; 第 31 位是溢出判断位, 若 OVERFLOW_BIT=1, 则不再接受新的读者守卫的获得申请.

6.5.3 3. 读写锁的主要 API

3.1 RwLock 的主要 API

```
///功能： 输入需要保护的数据类型data, 返回一个新的RwLock类型。
pub const fn new(data: T) -> Self

///功能： 获得READER守卫
pub fn read(&self) -> RwLockReadGuard<T>

///功能： 尝试获得READER守卫
pub fn try_read(&self) -> Option<RwLockReadGuard<T>>

///功能： 获得WRITER守卫
pub fn write(&self) -> RwLockWriteGuard<T>

///功能： 尝试获得WRITER守卫
pub fn try_write(&self) -> Option<RwLockWriteGuard<T>>
```

```
///功能：获得UPGRADER守卫
pub fn upgradeable_read(&self) -> RwLockUpgradableGuard<T>
```

```
///功能：尝试获得UPGRADER守卫
pub fn try_upgradeable_read(&self) -> Option<RwLockUpgradableGuard<T>>
```

3.2 WRITER 守卫 RwLockWriteGuard 的主要 API

```
///功能：将WRITER降级为READER
pub fn downgrade(self) -> RwLockReadGuard<'rwlock, T>
```

```
///功能：将WRITER降级为UPGRADER
pub fn downgrade_to_upgradeable(self) -> RwLockUpgradableGuard<'rwlock, T>
```

3.3 UPGRADER 守卫 RwLockUpgradableGuard 的主要 API

```
///功能：将UPGRADER升级为WRITER
pub fn upgrade(mut self) -> RwLockWriteGuard<'rwlock, T>
```

```
///功能：将UPGRADER降级为READER
pub fn downgrade(self) -> RwLockReadGuard<'rwlock, T>
```

6.5.4 4. 用法实例

```
static LOCK: RwLock<u32> = RwLock::new(100 as u32);

fn t_read1() {
    let guard = LOCK.read();
    let value = *guard;
    let readers_current = LOCK.reader_count();
    let writers_current = LOCK.writer_count();
    println!(
        "Reader1: the value is {value}
        There are totally {writers_current} writers, {readers_current} readers"
    );
}

fn t_read2() {
    let guard = LOCK.read();
```

(下页继续)

(续上页)

```

    let value = *guard;
    let readers_current = LOCK.reader_count();
    let writers_current = LOCK.writer_count();
    println!(
        "Reader2: the value is {value}
        There are totally {writers_current} writers, {readers_current} readers"
    );
}

fn t_write() {
    let mut guard = LOCK.write();
    *guard += 100;
    let writers_current = LOCK.writer_count();
    let readers_current = LOCK.reader_count();
    println!(
        "Writers: the value is {guard}
        There are totally {writers_current} writers, {readers_current} readers",
        guard = *guard
    );
    let read_guard=guard.downgrade();
    let value=*read_guard;
    println!("After downgraded to read_guard: {value}");
}

fn t_upgrade() {
    let guard = LOCK.upgradeable_read();
    let value = *guard;
    let readers_current = LOCK.reader_count();
    let writers_current = LOCK.writer_count();
    println!(
        "Upgrader1 before upgrade: the value is {value}
        There are totally {writers_current} writers, {readers_current} readers"
    );
    let mut upgraded_guard = guard.upgrade();
    *upgraded_guard += 100;
    let writers_current = LOCK.writer_count();
    let readers_current = LOCK.reader_count();
    println!(
        "Upgrader1 after upgrade: the value is {temp}
        There are totally {writers_current} writers, {readers_current} readers",
        temp = *upgraded_guard
    );
    let downgraded_guard=upgraded_guard.downgrade_to_upgradeable();

```

(下页继续)

(续上页)

```
    let value=*downgraded_guard;
    println!("value after downgraded: {value}");
    let read_guard=downgraded_guard.downgrade();
    let value_*=read_guard;
    println!("value after downgraded to read_guard: {value_}");
}

fn main() {
    let r2=thread::spawn(t_read2);
    let r1 = thread::spawn(t_read1);
    let t1 = thread::spawn(t_write);
    let g1 = thread::spawn(t_upgrade);
    r1.join().expect("r1");
    t1.join().expect("t1");
    g1.join().expect("g1");
    r2.join().expect("r2");
}
```

7.1 kthread 内核线程

内核线程模块定义在 `common/kthread.h` 中，提供对内核线程的支持功能。内核线程作为内核的“分身”，能够提升系统的并行化程度以及故障容错能力。

7.1.1 原理

每个内核线程都运行在内核态，执行其特定的任务。

内核线程的创建是通过调用 `kthread_create()` 或者 `kthread_run()` 宏，向 `kthreadd` 守护线程发送创建任务来实现的。也就是说，内核线程的创建，最终是由 `kthreadd` 来完成。

当内核线程被创建后，虽然会加入调度队列，但是当其被第一次调度，执行引导程序 `kthread()` 后，将进入休眠状态。直到其他模块使用 `process_wakeup()`，它才会真正开始运行。

当内核其他模块想要停止一个内核线程的时候，可以调用 `kthread_stop()` 函数。该函数将会置位内核线程的 `worker_private` 中的 `KTHREAD_SHOULD_STOP` 标志位，并等待内核线程的退出，然后获得返回值并清理内核线程的 `pcb`。

内核线程应当经常检查 `KTHREAD_SHOULD_STOP` 标志位，以确定其是否要退出。当检测到该标志位被置位时，内核线程应当完成数据清理工作，并调用 `kthread_exit()` 或直接返回一个返回码，以退出内核线程。

7.1.2 创建内核线程

kthread_create()

原型

```
kthread_create(thread_fn, data, name_fmt, arg...)
```

简介

在当前 NUMA 结点上创建一个内核线程（DragonOS 目前暂不支持 NUMA，因此 node 可忽略。）

请注意，该宏会创建一个内核线程，并将其设置为停止状态。

参数

thread_fn

该内核线程要执行的函数

data

传递给 *thread_fn* 的参数数据

name_fmt

printf-style format string for the thread name

arg

name_fmt 的参数

返回值

创建好的内核线程的 pcb

kthread_run()

原型

```
kthread_run(thread_fn, data, name_fmt, ...)
```

简介

创建内核线程并加入调度队列。

该宏定义是 `kthread_create()` 的简单封装，提供创建了内核线程后，立即运行的功能。

`kthread_run_rt()`

原型

```
kthread_run_rt(thread_fn, data, name_fmt, ...)
```

简介

创建内核实时线程并加入调度队列。

类似 `kthread_run()`，该宏定义也是 `kthread_create()` 的简单封装，提供创建了内核实时线程后，在设置实时进程的参数后，立即运行的功能。

7.1.3 停止内核线程

`kthread_stop()`

原型

```
int kthread_stop(struct process_control_block * pcb)
```

简介

当外部模块希望停止一个内核线程时，调用该函数，向 `kthread` 发送停止消息，请求其结束。并等待其退出，返回内核线程的退出返回值。

参数

`pcb`

内核线程的 `pcb`

返回值

内核线程的退出返回码。

kthread_should_stop()

原型

```
bool kthread_should_stop(void)
```

简介

内核线程可以调用该函数得知是否有其他进程请求结束当前内核线程。

返回值

一个 bool 变量

值	解释
true	有其他进程请求结束该内核线程
false	该内核线程没有收到停止消息

kthread_exit()

原型

```
void kthread_exit(long result)
```

简介

让当前内核线程退出，并返回 result 参数给 kthread_stop() 函数。

参数

result

内核线程的退出返回码

7.2 PCB 进程控制块

PCB 的全称为 process control block, 它是每个进程/线程的核心控制结构。定义于 `kernel/src/process/proc-types.h` 中。

7.2.1 PCB 详解

Todo:

7.2.2 与 PCB 的管理相关的 API

根据 pid 寻找 pcb

`process_find_pcb_by_pid`

该 API 提供了根据 pid 寻找 pcb 的功能, 定义在 `kernel/src/process/process.h` 中。

当找到目标的 pcb 时, 返回对应的 pcb, 否则返回 NULL。

参数

pid 进程 id

返回值

struct process_control_block 目标 pcb

这里是 DragonOS 中，与进程调度相关的说明文档。

8.1 与“等待”相关的 api (C 语言)

警告：随着内核的发展，我们将会逐步将 C 语言的等待机制替换为 Rust 语言的等待机制。在这个过程中，我们将会同时保留 C 语言和 Rust 语言的等待机制，以便于我们在开发过程中进行对比。待时机成熟，我们将会逐步将 C 语言的等待机制移除。

如果几个进程需要等待某个事件发生，才能被运行，那么就需要一种“等待”的机制，以实现进程同步。

8.1.1 一. wait_queue 等待队列

wait_queue 是一种进程同步机制，中文名为“等待队列”。它可以将当前进程挂起，并在时机成熟时，由另一个进程唤醒他们。

当您需要等待一个事件完成时，使用 wait_queue 机制能减少进程同步的开销。相比于滥用自旋锁以及信号量，或者是循环使用 usleep(1000) 这样的函数来完成同步，wait_queue 是一个高效的解决方案。

警告：wait_queue.h 中的等待队列的实现并没有把队列头独立出来，同时没有考虑为等待队列加锁。所以在后来的开发中加入了 wait_queue_head_t 的队列头实现，实质上就是链表 + 自旋锁。它与

wait_queue.h 中的队列 wait_queue_node_t 是兼容的, 当你使用 struct wait_queue_head 作为队列头时, 你同样可以使用等待队列添加节点的函数。

简单用法

等待队列的使用方法主要分为以下几部分:

- 创建并初始化一个等待队列
- 使用 wait_queue_sleep_on_ 系列的函数, 将当前进程挂起。晚挂起的进程将排在队列的尾部。
- 通过 wait_queue_wakeup() 函数, 依次唤醒在等待队列上的进程, 将其加入调度队列

要使用 wait_queue, 您需要 #include<common/wait_queue.h>, 并创建一个 wait_queue_node_t 类型的变量, 作为等待队列的头。这个结构体只包含两个成员变量:

```
typedef struct
{
    struct List wait_list;
    struct process_control_block *pcb;
} wait_queue_node_t;
```

对于等待队列, 这里有一个好的命名方法:

```
wait_queue_node_t wq_keyboard_interrupt_received;
```

这样的命名方式能增加代码的可读性, 更容易让人明白这里到底在等待什么。

初始化等待队列

函数 wait_queue_init(wait_queue_node_t *wait_queue, struct process_control_block *pcb) 提供了初始化 wait_queue 结点的功能。

当您初始化队列头部时, 您仅需要将 wait_queue 首部的结点指针传入, 第二个参数请设置为 NULL

将结点插入等待队列

您可以使用以下函数, 将当前进程挂起, 并插入到指定的等待队列。这些函数大体功能相同, 只是在一些细节上有所不同。

函数名	解释
wait_queue_sleep_on()	将当前进程挂起，并设置挂起状态为 PROC_UNINTERRUPTIBLE
wait_queue_sleep_on_unlock()	将当前进程挂起，并设置挂起状态为 PROC_UNINTERRUPTIBLE。待当前进程被插入等待队列后，解锁给定的自旋锁
wait_queue_sleep_on_interruptible()	将当前进程挂起，并设置挂起状态为 PROC_INTERRUPTIBLE

从等待队列唤醒一个进程

您可以使用 `void wait_queue_wakeup(wait_queue_node_t * wait_queue_head, int64_t state);` 函数，从指定的等待队列中，唤醒第一个挂起时的状态与指定的 `state` 相同的进程。

当没有符合条件的进程时，将不会唤醒任何进程，仿佛无事发生。

8.1.2 二. wait_queue_head 等待队列头

数据结构定义如下：

```
typedef struct
{
    struct List wait_list;
    spinlock_t lock; // 队列需要有一个自旋锁, 虽然目前内部并没有使用, 但是以后可能会用.
} wait_queue_head_t;
```

等待队列头的使用逻辑与等待队列实际是一样的，因为他同样也是等待队列的节点（仅仅多了一把锁）。且 `wait_queue_head` 的函数基本上与 `wait_queue` 一致，只不过多了 `***_with_node_***` 的字符串。

同时，`wait_queue.h` 文件中提供了很多的宏，可以方便您的工作。

提供的宏

宏	解释
DE- CLARE_WAIT_ON_STACK(name, pcb)	在栈上声明一个 wait_queue 节点，同时把 pcb 所代表的进程与该节点绑定
DE- CLARE_WAIT_ON_STACK_SELF(name)	传在栈上声明一个 wait_queue 节点，同时当前进程 (即自身进程) 与该节点绑定
DE- CLARE_WAIT_ALLOC(name, pcb)	使用 kcalloc 声明一个 wait_queue 节点，同时把 pcb 所代表的进程与该节点绑定，请记得使用 kfree 释放空间
DE- CLARE_WAIT_ALLOC_SELF(name)	使用 kcalloc 声明一个 wait_queue 节点，同时当前进程 (即自身进程) 与该节点绑定，请记得使用 kfree 释放空间

创建等待队列头

您可以直接调用宏

```
DECLARE_WAIT_QUEUE_HEAD(m_wait_queue_head); // 在栈上声明一个队列头变量
```

也可以手动声明

```
struct wait_queue_head_t m_wait_queue_head = {0};  
wait_queue_head_init(&m_wait_queue_head);
```

将结点插入等待队列

函数名	解释
wait_queue_sleep_with_node(wait_queue_head_t *head, wait_queue_node_t *wait_node)	传入一个等待队列节点，并设置该节点的挂起状态为 PROC_UNINTERRUPTIBLE
wait_queue_sleep_with_node_unlock(wait_queue_head_t *q, wait_queue_node_t *wait, void *lock)	传入一个等待队列节点，将该节点的 pcb 指向的进程挂起，并设置挂起状态为 PROC_UNINTERRUPTIBLE。待当前进程被插入等待队列后，解锁给定的自旋锁
wait_queue_sleep_with_node_interruptible(wait_queue_head_t *q, wait_queue_node_t *wait)	传入一个等待队列节点，将该节点的 pcb 指向的进程挂起，并设置挂起状态为 PROC_INTERRUPTIBLE

从等待队列唤醒一个进程

在 `wait_queue.h` 中的 `wait_queue_wakeup` 函数直接 `kfree` 掉了 `wait_node` 节点。对于在栈上的 `wait_node`, 您可以选择 `wait_queue_wakeup_on_stack(wait_queue_head_t *q, int64_t state)` 来唤醒队列里面的队列头节点。

8.1.3 三. completion 完成量

简单用法

完成量的使用方法主要分为以下几部分:

- 声明一个完成量 (可以在栈中/使用 `kmalloc`/使用数组)
- 使用 `wait_for_completion` 等待事件完成
- 使用 `complete` 唤醒等待的进程

等待操作

```
void wait_fun() {
    DECLARE_COMPLETION_ON_STACK(comp); // 声明一个 completion

    // .... do something here
    // 大部分情况是你使用 kthread_run() 创建了另一个线程
    // 你需要把 comp 变量传给这个线程, 然后当前线程就会等待他的完成

    if (!try_wait_for_completion(&comp)) // 进入等待
        wait_for_completion(&comp);
}
```

完成操作

```
void kthread_fun(struct completion *comp) {
    // ..... 做一些事 .....
    // 这里你确定你完成了目标事件

    complete(&comp);
    // 或者你使用 complete_all
    complete_all(&comp);
}
```

更多用法

kernel/sched/completion.c 文件夹中, 你可以看到 __test 开头的几个函数, 他们是 completion 模块的测试代码, 基本覆盖了 completion 的大部分函数. 你可以在这里[查询函数使用方法](#).

初始化完成量

函数 completion_init(struct completion *x) 提供了初始化 completion 的功能. 当你使用 DECLARE_COMPLETION_ON_STACK 来创建 (在栈上创建) 的时候, 会自动初始化.

关于完成量的 wait 系列函数

函数名	解释
wait_for_completion(struct completion *x)	将当前进程挂起, 并设置挂起状态为 PROC_UNINTERRUPTIBLE。
wait_for_completion_timeout(struct completion *x, long timeout)	将当前进程挂起, 并设置挂起状态为 PROC_UNINTERRUPTIBLE。当等待 timeout 时间 (jiffies 时间片) 之后, 自动唤醒进程。
wait_for_completion_interruptible(struct completion *x)	将当前进程挂起, 并设置挂起状态为 PROC_INTERRUPTIBLE。
wait_for_completion_interruptible_timeout(struct completion *x, long timeout)	将当前进程挂起, 并设置挂起状态为 PROC_INTERRUPTIBLE。当等待 timeout 时间 (jiffies 时间片) 之后, 自动唤醒进程。
wait_for_multicompletion(struct completion x[], int n)	将当前进程挂起, 并设置挂起状态为 PROC_UNINTERRUPTIBLE。(等待数组里面的 completion 的完成)

关于完成量的 complete 系列函数

函数名	解释
complete(struct completion *x)	表明一个事件被完成, 从等待队列中唤醒一个进程
complete_all(struct completion *x)	表明与该 completion 有关的事件被标记为永久完成, 并唤醒等待队列中的所有进程

其他用于查询信息的函数

函数名	解释
completion_done(struct completion *x)	查询 completion 的 done 变量是不是大于 0，如果大于 0，返回 true；否则返回 false。在等待前加上这个函数有可能加速？（暂未经过实验测试，有待证明）
try_wait_for_completion(struct completion *x)	查询 completion 的 done 变量是不是大于 0，如果大于 0，返回 true(同时令 done-=1)；否则返回 false。在等待前加上这个函数有可能加速？（该函数和 completion_done 代码逻辑基本一致，但是会主动令 completion 的 done 变量减 1）

8.2 与“等待”相关的 api (rust 语言)

如果几个进程需要等待某个事件发生，才能被运行，那么就需要一种“等待”的机制，以实现进程同步。

8.2.1 1. WaitQueue 等待队列

WaitQueue 是一种进程同步机制，中文名为“等待队列”。它可以将当前进程挂起，并在时机成熟时，由另一个进程唤醒他们。

当您需要等待一个事件完成时，使用 WaitQueue 机制能减少进程同步的开销。相比于滥用自旋锁以及信号量，或者是循环使用 usleep(1000) 这样的函数来完成同步，WaitQueue 是一个高效的解决方案。

1.1 WaitQueue 的使用

WaitQueue 的使用非常简单，只需要三步：

- 1. 初始化一个 WaitQueue 对象。
- 2. 调用这个 WaitQueue 的挂起相关的 API，将当前进程挂起。
- 3. 当事件发生时，由另一个进程，调用这个 WaitQueue 的唤醒相关的 API，唤醒一个进程。

下面是一个简单的例子：

1.1.1 初始化一个 WaitQueue 对象

WaitQueue 对象的初始化非常简单，只需要调用 WaitQueue::INIT 即可。

```
let mut wq = WaitQueue::INIT;
```

1.1.2 挂起进程

您可以这样挂起当前进程：

```
wq.sleep();
```

当前进程会被挂起，直到有另一个进程调用了 `wq.wakeup()`。

1.1.3 唤醒进程

您可以这样唤醒一个进程：

```
// 唤醒等待队列头部的进程（如果它的state & PROC_INTERRUPTIBLE 不为0）
wq.wakeup(PROC_INTERRUPTIBLE);

// 唤醒等待队列头部的进程（如果它的state & PROC_UNINTERRUPTIBLE 不为0）
wq.wakeup(PROC_UNINTERRUPTIBLE);

// 唤醒等待队列头部的进程（无论它的state是什么）
wq.wakeup((-1) as u64);
```

1.2 API

1.2.1 挂起进程

您可以使用以下函数，将当前进程挂起，并插入到指定的等待队列。这些函数大体功能相同，只是在一些细节上有所不同。

函数名	解释
<code>sleep()</code>	将当前进程挂起，并设置进程状态为 <code>PROC_INTERRUPTIBLE</code>
<code>sleep_uninterruptible()</code>	将当前进程挂起，并设置进程状态为 <code>PROC_UNINTERRUPTIBLE</code>
<code>sleep_unlock_spinlock()</code>	将当前进程挂起，并设置进程状态为 <code>PROC_INTERRUPTIBLE</code> 。待当前进程被插入等待队列后，解锁给定的自旋锁
<code>sleep_unlock_mutex()</code>	将当前进程挂起，并设置进程状态为 <code>PROC_INTERRUPTIBLE</code> 。待当前进程被插入等待队列后，解锁给定的 <code>Mutex</code>
<code>sleep_uninterruptible_unlock_spinlock()</code>	将当前进程挂起，并设置进程状态为 <code>PROC_UNINTERRUPTIBLE</code> 。待当前进程被插入等待队列后，解锁给定的自旋锁
<code>sleep_uninterruptible_unlock_mutex()</code>	将当前进程挂起，并设置进程状态为 <code>PROC_UNINTERRUPTIBLE</code> 。待当前进程被插入等待队列后，解锁给定的 <code>Mutex</code>

1.2.2 唤醒进程

您可以使用 `wakeup(state)` 函数，唤醒等待队列中的第一个进程。如果这个进程的 `state` 与给定的 `state` 进行 `and` 操作之后，结果不为 0, 则唤醒它。

返回值：如果有进程被唤醒，则返回 `true`，否则返回 `false`。

1.2.3 其它 API

函数名	解释
<code>len()</code>	返回等待队列中的进程数量

8.3 进程调度器相关的 api

定义了 DragonOS 的进程调度相关的 api，是系统进行进程调度的接口。同时也抽象出了 Scheduler 的 trait，以供具体的调度器实现

8.3.1 1. 调度器介绍

一般来说，一个系统会同时处理多个请求，但是其资源是优先的，调度就是用来协调每个请求对资源使用的方法。

1.1 主要函数

- 1. `cpu_executing()`: 获取指定的 `cpu` 上正在执行的进程的 `pcb`
- 2. `sched_enqueue()`: 将进程加入调度队列
- 3. `sched_init()`: 初始化进程调度器模块
- 4. `sched_update_jiffies()`: 当时钟中断到达时，更新时间片。请注意，该函数只能被时钟中断处理程序调用
- 5. `sys_sched()`: 让系统立即运行调度器的系统调用。请注意，该系统调用不能由 `ring3` 的程序发起

8.4 完全公平调度器相关的 api

CFS (Completely Fair Scheduler)，顾名思义，完全公平调度器。CFS 作为主线调度器之一，也是最典型的 $O(1)$ 调度器之一

8.4.1 1. CFSQueue 介绍

CFSQueue 是用来存放普通进程的调度队列，每个 CPU 维护一个 RTQueue，主要使用 Vec 作为主要存储结构来实现。

1.1 主要函数

1. enqueue(): 将 pcb 入队列
2. dequeue(): 将 pcb 从调度队列中弹出, 若队列为空, 则返回 IDLE 进程的 pcb
3. sort(): 将进程按照虚拟运行时间的升序进行排列

8.4.2 2. SchedulerCFS 介绍

CFS 调度器类，主要实现了 CFS 调度器类的初始化以及调度功能函数。

2.1 主要函数

1. sched(): 是对于 Scheduler trait 的 sched() 实现，是实时进程进行调度时的逻辑处理，该函数会返回接下来要执行的 pcb，若没有符合要求的 pcb，返回 None
2. enqueue(): 同样对于 Scheduler trait 的 sched() 实现，将一个 pcb 加入调度器的调度队列
3. update_cpu_exec_proc_jiffies(): 更新这个 cpu 上，这个进程的可执行时间。
4. timer_update_jiffies(): 时钟中断到来时，由 sched 的 core 模块中的函数，调用本函数，更新 CFS 进程的可执行时间

8.5 实时进程调度器相关的 api

RT (realtime scheduler)，实时调度器。实时调度是为了完成实时处理任务而分配 CPU 的调度方法。

DragonOS 的进程分为“实时进程”和“普通进程”两类；实时进程的优先级高于普通进程，如果当前的系统的执行队列中有“实时进程”，RT 调度器会优先选择实时进程；如果队列中会有多个实时进程，调度器会选择优先级最高的实时进程来执行；

8.5.1 1. RTQueue 介绍

RTQueue 是用来存放 state 为 running 的实时进程的调度队列，每个 CPU 维护一个 RTQueue，主要使用 Vec 作为主要存储结构来实现。

1.1 主要函数

1. enqueue(): 将 pcb 入队列
2. dequeue(): 将 pcb 出队列

8.5.2 2. SchedulerRT 介绍

RT 调度器类，主要实现了 RT 调度器类的初始化以及调度功能函数。

2.1 主要函数

1. pick_next_task_rt(): 获取当前 CPU 中的第一个需要执行的 RT pcb
2. sched(): 是对于 Scheduler trait 的 sched() 实现，是实时进程进行调度时的逻辑处理，该函数会返回接下来要执行的 pcb，若没有符合要求的 pcb，返回 None
3. enqueue(): 同样也是对于 Scheduler trait 的 sched() 实现，将一个 pcb 加入调度器的调度队列

2.2 内核调度策略

目前在 DragonOS 中，主要的调度策略有 SCHED_NORMAL 策略 | SCHED_FIFO 策略 | SCHED_RT 策略，具体的调度策略为：

1. SCHED_NORMAL 策略：SCHED_NORMAL 是“绝对公平调度策略”，该策略的进程使用 CFS 进行调度。
2. SCHED_FIFO 策略：SCHED_FIFO 是“实时进程调度策略”，这是一种先进先出的调度策略，该策略不涉及到 CPU 时间片机制，在没有更高优先级进程的前提下，只能等待其他进程主动释放 CPU 资源；在 SCHED_FIFO 策略中，被调度器调度运行的进程，其运行时长不受限制，可以运行任意长的时间。
3. SCHED_RR 策略：SCHED_RR 是“实时进程调度策略”，使用的是时间片轮转机制，对应进程的 time_slice 会在运行时减少，进程使用完 CPU 时间片后，会加入该 CPU 的与该进程优先级相同的执行队列中。同时，释放 CPU 资源，CPU 的使用权会被分配给下一个执行的进程

8.5.3 3. Q&A

几种常用的方法

1. 如何创建实时进程

```
struct process_control_block *pcb_name = kthread_run_rt(&fn_name, NULL, "test_↵  
↵create rt pcb");
```

其中 `kthread_run_rt`, 是创建内核实时线程的宏

2. pcb 中涉及到实时进程的字段含义

1. policy: 实时进程的策略, 目前有: `SCHED_FIFO` 与 `SCHED_RR`
2. priority: 实时进程的优先级, 范围为 0-99, 数字越大, 表示优先级越高
3. rt_time_slice: 实时进程的时间片, 默认为 100, 随着 CPU 运行而减少, 在 `rt_time_slice` 为 0 时, 将时间片赋初值并将该进程加入执行队列。

3. 如何实时进程存储队列

- 目前是使用 `Vec` 来保存, 因为具体实现的逻辑原因, 目前的入队列和出队列都是对队尾的操作, 因此会有如下现象: 系统中有多个优先级相同的实时进程等待运行时, 会出现饥饿现象, 也即上一个因为时间片耗尽的进程会在下一个执行, 造成同优先级等待的进程饥饿。

4. todo

1. 将存储实时进程的队列使用双向链表存储 (或者其他办法解决上述的饥饿问题)
2. 目前的实时调度是针对单 CPU 的, 需要实现多 CPU 的实时调度
3. 实现 RT 进程和普通进程之间的分配带宽的比例
4. 多个 CPU 之间实现负载均衡

这里是 DragonOS 进程间通信（IPC）的说明文档。

9.1 Signal 信号

备注：本文 Maintainer: 龙进

Email: longjin@RinGoTek.cn

信号是一种进程间通信机制，当我们发送一个信号给特定的进程，就能触发它的特定行为（例如退出程序，或者运行一些信号处理程序）。信号是发送到进程或同一进程内的特定线程的异步通知，用于通知它有事件发生。信号的常见用途是中断、挂起、终止或终止进程。发送信号时，操作系统会中断目标进程的正常执行流程以传递信号。可以在任何非原子指令期间中断执行。如果该进程之前注册了一个信号处理程序，则执行该例程。否则，将执行默认信号处理程序。

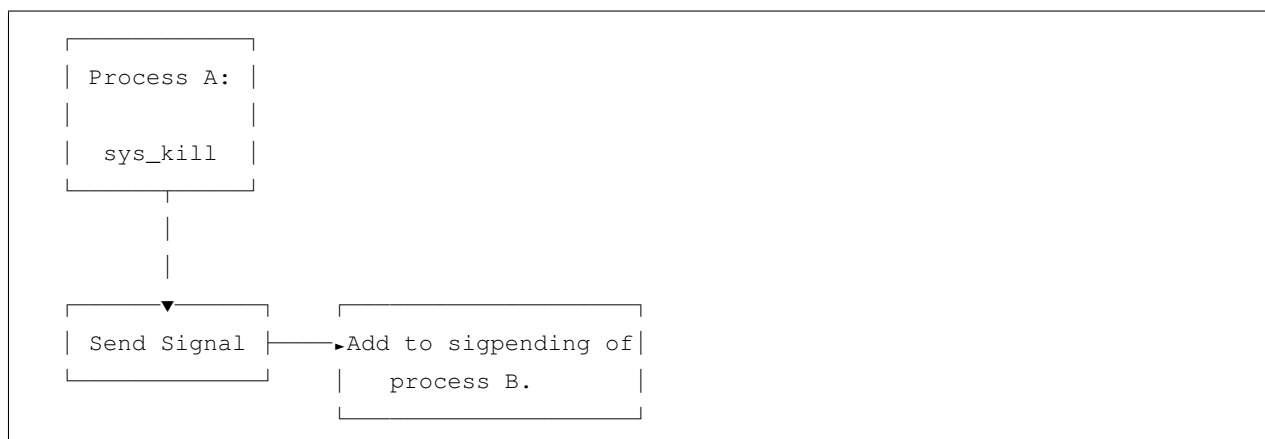
信号类似于中断，区别在于中断由 CPU 调解并由内核处理，而信号在内核中产生（也可通过系统调用让内核产生）并由各个进程或内核的默认处理函数处理。

9.1.1 1. 信号处理概要

1.1 信号发送

当进程 A 想发送信号给进程 B 的时候，使用 `kill(pid, signal)` 接口进行发送。然后陷入内核的 `sys_kill()` 函数中进行处理。然后内核将会把信号加入目标进程的 `pcb` 的 `sigpending` 中。

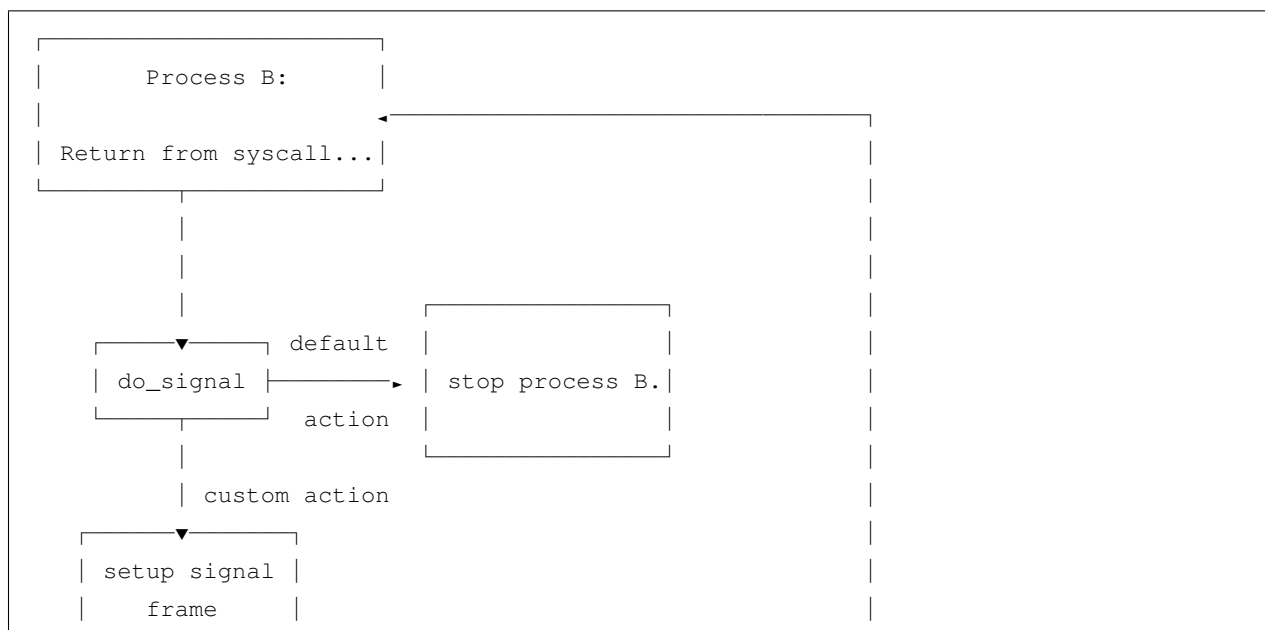
示意图如下：



1.2 信号处理

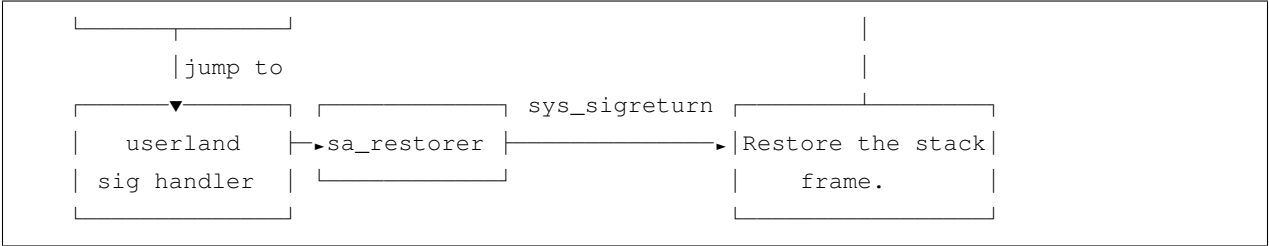
进程会在退出内核态的时候，跳转到 `do_signal()` 函数内，检查当前是否有需要被处理的信号，如果有的话，就会开启信号处理流程。

信号处理流程示意图：



(下页继续)

(续上页)



- 如果内核检查发现，进程没有指定信号处理函数，且信号处理动作不是“忽略”，那就会杀掉进程。
- 如果内核发现该信号没有被忽略，那么将会：
 - 备份当前内核栈
 - 设置信号处理的用户态栈帧
 - 回到用户态，执行信号处理函数
 - 信号处理函数结束之后，将会进入由 `libc` 提供的 `__sa_restorer` 中，发起 `sys_sigreturn()` 系统调用，回到内核态
 - 内核恢复处理信号之前的内核栈。
 - 信号处理流程结束，内核继续执行“返回用户态”的过程。
- 如果内核发现当前信号被忽略，那么就检查下一个信号。
- 发现没有任何需要处理的信号时，返回用户态。

9.1.2 2. 其他问题

暂无。

这里讲解了内存管理模块的一些设计及实现原理。

如果你正在寻找使用内存管理模块的方法，请转到：[内存管理 API 文档](#)

10.1 MMIO

MMIO 是“内存映射 IO”的缩写，它被广泛应用于与硬件设备的交互之中。

10.1.1 地址空间管理

DragonOS 中实现了 MMIO 地址空间的管理机制，本节将介绍它们。

为什么需要 MMIO 地址空间自动分配？

由于计算机上的很多设备都需要 MMIO 的地址空间，而每台计算机上所连接的各种设备的对 MMIO 地址空间的需求是不一样的。如果我们为每个类型的设备都手动指定一个 MMIO 地址，会使得虚拟地址空间被大大浪费，也会增加系统的复杂性。并且，我们在将来还需要为不同的虚拟内存区域做异常处理函数。因此，我们需要一套能够自动分配 MMIO 地址空间的机制。

这套机制提供了什么功能？

- 为驱动程序分配 4K 到 1GB 的 MMIO 虚拟地址空间
- 对于这些虚拟地址空间，添加到 VMA 中进行统一管理
- 可以批量释放这些地址空间

这套机制是如何实现的？

这套机制本质上是使用了伙伴系统来对 MMIO 虚拟地址空间进行维护。在 `mm/mm.h` 中指定了 MMIO 的虚拟地址空间范围，这个范围是 `0xfffffa1000000000` 开始的 1TB 的空间。也就是说，这个伙伴系统为 MMIO 维护了这 1TB 的虚拟地址空间。

地址空间分配过程

1. 初始化 MMIO-mapping 模块，在 `mmio` 的伙伴系统中创建 512 个 1GB 的 `__mmio_buddy_addr_region`
2. 驱动程序使用 `mmio_create` 请求分配地址空间。
3. `mmio_create` 对申请的地址空间大小按照 2 的 n 次幂进行对齐，然后从 `buddy` 中申请内存地址空间
4. 创建 VMA，并将 VMA 标记为 `VM_IO|VM_DONTCOPY`。MMIO 的 `vma` 只绑定在 `initial_mm` 下，且不会被拷贝。
5. 分配完成

一旦 MMIO 地址空间分配完成，它就像普通的 `vma` 一样，可以使用 `mmap` 系列函数进行操作。

MMIO 的映射过程

在得到了虚拟地址空间之后，当我们尝试往这块地址空间内映射内存时，我们可以调用 `mm_map` 函数，对这块区域进行映射。

该函数会对 MMIO 的 VMA 的映射做出特殊处理。即：创建 `Page` 结构体以及对应的 `anon_vma`。然后会将对应的物理地址，填写到页表之中。

MMIO 虚拟地址空间的释放

当设备被卸载时，驱动程序可以调用 `mmio_release` 函数对指定的 `mmio` 地址空间进行释放。

释放的过程中，`mmio_release` 将执行以下流程：

1. 取消 `mmio` 区域在页表中的映射。
2. 将释放 MMIO 区域的 VMA
3. 将地址空间归还给 `mmio` 的伙伴系统。

10.1.2 MMIO 的伙伴算法

伙伴的定义

同时满足以下三个条件的两个内存块被称为伙伴内存块：

- 1. 两个内存块的大小相同
- 2. 两个内存块的内存地址连续
- 3. 两个内存块由同一个大块内存分裂得到

伙伴算法

伙伴（buddy）算法的作用是维护以及组织大块连续内存块的分配和回收，以减少系统时运行产生的外部碎片。伙伴系统中的每个内存块的大小均为 2^n 。在 DragonOS 中，伙伴系统内存池共维护了 1TB 的连续存储空间，最大的内存块大小为 1G，即 $2^{30}B$ ，最小的内存块大小为 4K，即 $2^{12}B$ 。

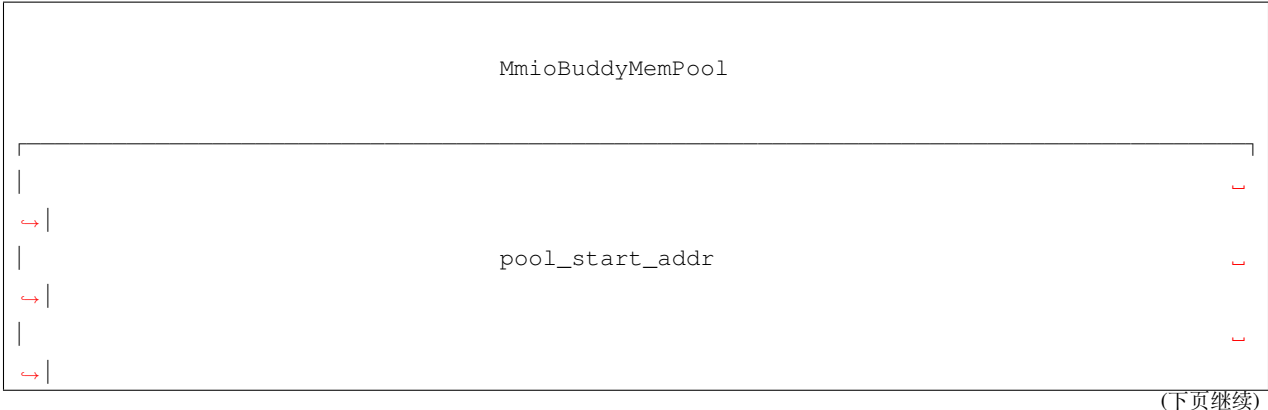
伙伴算法的核心思想是当应用申请内存时，每次都分配比申请的内存大小更大的最小内存块，同时分配出去的内存块大小为 $2^n B$ 。（e.g. 假设某应用申请了 $3B$ 内存，显然并没有整数 n ，使 $2^n = 3$ ，且 $3 \in [2^1, 2^2]$ ，所以系统会去取一块大小为 $2^2 B$ 的内存块，将它分配给申请的应用，本次申请内存操作顺利完成。）

那么当伙伴系统中没有如此“合适”的内存块时该怎么办呢？系统先会去寻找更大的内存块，如果找到了，则会将大内存块分裂成合适的内存块分配给应用。（e.g. 假设申请 $3B$ 内存，此时系统中比 $3B$ 大的最小内存块的大小为 $16B$ ，那么 $16B$ 会被分裂成两块 $8B$ 的内存块，一块放入内存池中，一块继续分裂成两块 $4B$ 的内存块。两块 $4B$ 的内存块，一块放入内存池中，一块分配给应用。至此，本次申请内存操作顺利完成。）

如果系统没有找到更大的内存块，系统将会尝试合并较小的内存块，直到符合申请空间的大小。（e.g. 假设申请 $3B$ 内存，系统检查内存池发现只有两个 $2B$ 的内存块，那么系统将会把这两个 $2B$ 的内存块合并成一块 $4B$ 的内存块，并分配给应用。至此，本次申请内存操作顺利完成。）

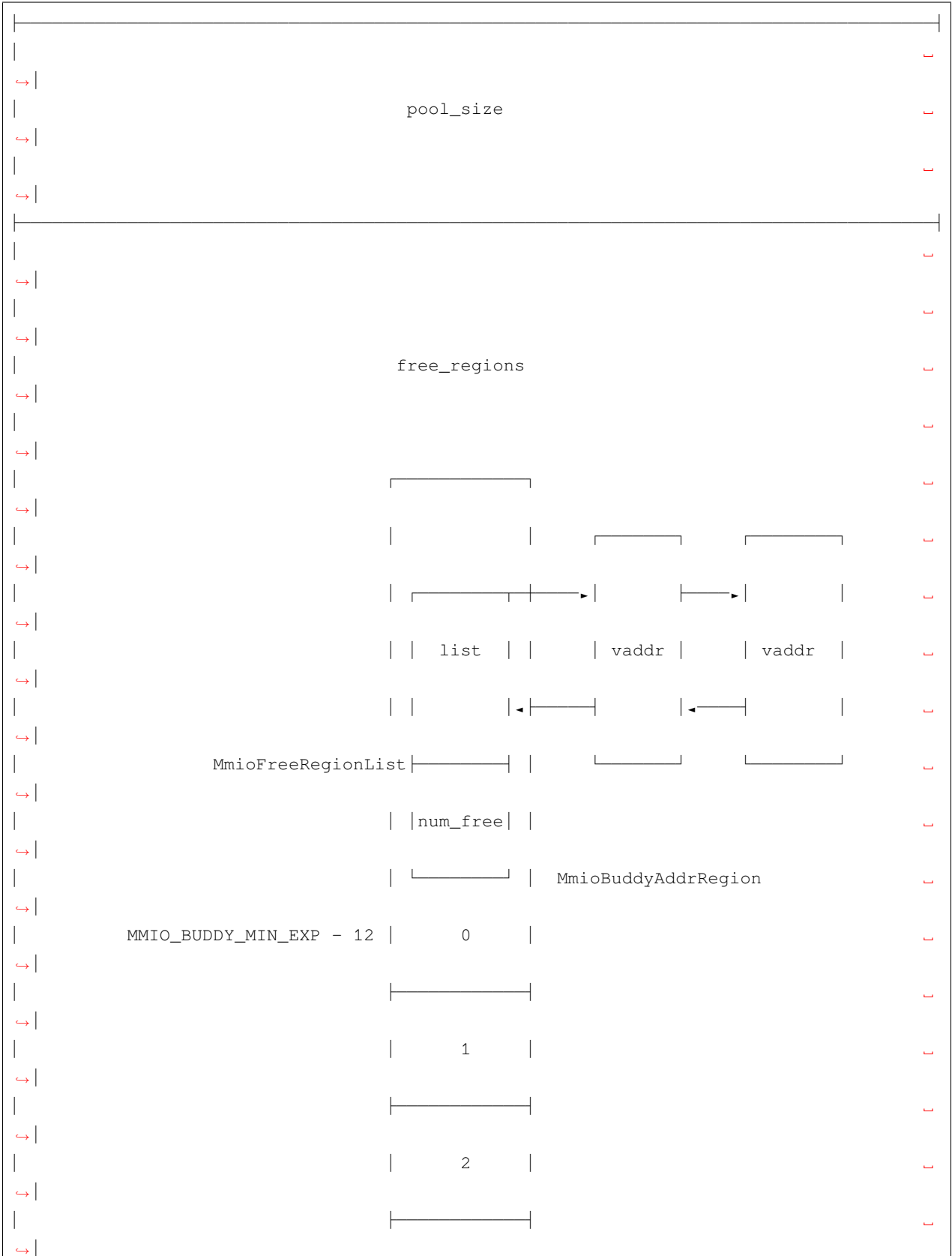
最后，当系统既没有找到大块内存，又无法成功合并小块内存时，就会通知应用内存不够，无法分配内存。

伙伴算法的数据结构



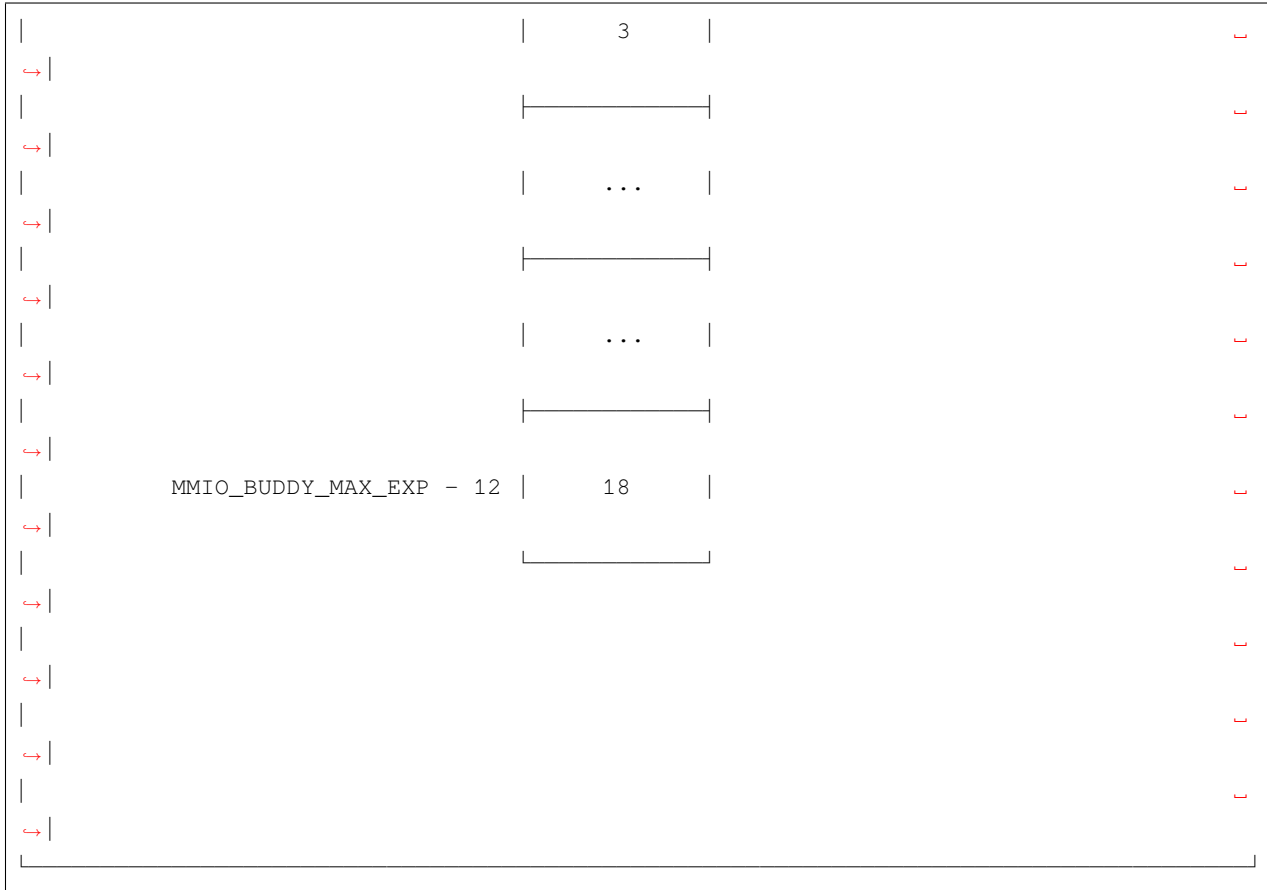
(下页继续)

(续上页)



(下页继续)

(续上页)



```

/// 最大的内存块为1G，其幂为30
const MMIO_BUDDY_MAX_EXP: u32 = PAGE_1G_SHIFT;
/// 最小的内存块为4K，其幂为12
const MMIO_BUDDY_MIN_EXP: u32 = PAGE_4K_SHIFT;
/// 内存池数组的大小为18
const MMIO_BUDDY_REGION_COUNT: u32 = MMIO_BUDDY_MAX_EXP - MMIO_BUDDY_MIN_EXP + 1;

/// buddy内存池
pub struct MmioBuddyMemPool {
    /// 内存池的起始地址
    pool_start_addr: u64,
    /// 内存池大小：初始化为1TB
    pool_size: u64,
    /// 空闲内存块链表数组
    /// MMIO_BUDDY_REGION_COUNT = MMIO_BUDDY_MAX_EXP - MMIO_BUDDY_MIN_EXP + 1
    free_regions: [SpinLock<MmioFreeRegionList>; MMIO_BUDDY_REGION_COUNT as usize],
}

/// 空闲内存块链表结构体

```

(下页继续)

(续上页)

```

pub struct MmioFreeRegionList {
    /// 存储了空闲内存块信息的结构体的链表
    list: LinkedList<Box<MmioBuddyAddrRegion>>,
    /// 当前链表空闲块的数量
    num_free: i64,
}

/// mmio伙伴系统内部的地址区域结构体
pub struct MmioBuddyAddrRegion {
    /// 内存块的起始地址
    vaddr: u64,
}

```

设计思路

DragonOS 中, 使用 `MmioBuddyMemPool` 结构体作为 `buddy` (为表述方便, 以下将伙伴算法简称为 `buddy`) 内存池的数据结构, 其记录了内存池的起始地址 (`pool_start_addr`) 以及内存池中内存块的总大小 (`pool_size`), 同时其维护了大小为 `MMIO_BUDDY_REGION_COUNT` 的双向链表数组 (`free_regions`)。 `free_regions` 中的各个链表维护了若干空闲内存块 (`MmioBuddyAddrRegion`)。

`free_regions` 的下标 (`index`) 与内存块的大小有关。由于每个内存块大小都为 2^n bytes, 那么可以令 $exp = n$ 。 `index` 与 `exp` 的换算公式如下: $index = exp - 12$ 。 e.g. 一个大小为 2^{12} bytes 的内存块, 其 $exp = 12$, 使用上述公式计算得 $index = 12 - 12 = 0$, 所以该内存块会被存入 `free_regions[0].list` 中。 通过上述换算公式, 每次取出或释放 2^n 大小的内存块, 只需要操作 `free_regions[n - 12]` 即可。 DragonOS 中, `buddy` 内存池最大的内存块大小为 $1G = 2^{30}bytes$, 最小的内存块大小为 $4K = 2^{12}bytes$, 所以 $index \in [0, 18]$ 。

作为内存分配机制, `buddy` 服务于所有进程, 为了解决在各个进程之间实现 `free_regions` 中的链表数据同步的问题, `free_regions` 中的链表类型采用加了自旋锁 (`SpinLock`) 的空闲内存块链表 (`MmioFreeRegionList`), `MmioFreeRegionList` 中封装有真正的存储了空闲内存块信息的结构体的链表 (`list`) 和对应链表长度 (`num_free`)。 有了自选锁后, 同一时刻只允许一个进程修改某个链表, 如取出链表元素 (申请内存) 或者向链表中插入元素 (释放内存)。

`MmioFreeRegionList` 中的元素类型为 `MmioBuddyAddrRegion` 结构体, `MmioBuddyAddrRegion` 记录了内存块的起始地址 (`vaddr`)。

伙伴算法内部 api

P.S 以下函数均为 MmioBuddyMemPool 的成员函数。系统中已经创建了一个 MmioBuddyMemPool 类型的全局引用 MMIO_POOL，如要使用以下函数，请以 MMIO_POOL.xxx() 形式使用，以此形式使用则不需要传入 self。

函数名	描述
__create_region(&self, vaddr)	将虚拟地址传入，创建新的内存块地址结构体
__give_back_block(&self, vaddr, exp)	将地址为 vaddr，幂为 exp 的内存块归还给 buddy
__buddy_split(&self,region,exp,list_guard)	将给定大小为 2^{exp} 的内存块一分为二，并插入内存块大小为 2^{exp-1} 的链表中
__query_addr_region(&self,exp,list_guard)	从 buddy 中申请一块大小为 2^{exp} 的内存块
mmio_buddy_query_addr_region(&self,exp)	对 query_addr_region 进行封装， 请使用这个函数，而不是 __query_addr_region
__buddy_add_region_obj(&self,region,list_guard)	往指定的地址空间链表中添加一个内存块
__buddy_block_vaddr(&self, vaddr, exp)	根据地址和内存块大小，计算伙伴块虚拟内存的地址
__pop_buddy_block(&self, vaddr,exp,list_guard)	寻找并弹出指定内存块的伙伴块
__buddy_pop_region(&self, list_guard)	从指定空闲链表中取出内存区域
__buddy_merge(&self,exp,list_guard,high_list_guard)	合并所有 2^{exp} 大小的内存块
__buddy_merge_blocks(&self,region_1,region_2,exp,high_list_guard)	合并两个已经从链表中取出的内存块

伙伴算法对外 api

函数名	描述
__mmio_buddy_init()	初始化 buddy 系统，在 mmio_init() 中调用，请勿随意调用
__exp2index(exp)	将 2^{exp} 的 exp 转换成内存池中的数组的下标（index）
mmio_create(size,vm_flags,res_vaddr,res_length)	创建一块根据 size 对齐后的大小的 mmio 区域，并将其 vma 绑定到 initial_mm
mmio_release(vaddr, length)	取消地址为 vaddr，大小为 length 的 mmio 的映射并将其归还到 buddy 中

DragonOS 的文件系统模块由 VFS（虚拟文件系统）及具体的文件系统组成。

todo: 由于文件系统模块重构，文档暂时不可用，预计在 2023 年 4 月 10 日前补齐。

11.1 VFS 虚拟文件系统

在 DragonOS 中，VFS 作为适配器，遮住了具体文件系统之间的差异，对外提供统一的文件操作接口抽象。

11.1.1 DragonOS 虚拟文件系统概述

简介

DragonOS 的虚拟文件系统是内核中的一层适配器，为用户程序（或者是系统程序）提供了通用的文件系统接口。同时对内核中的不同文件系统提供了统一的抽象。各种具体的文件系统可以挂载到 VFS 的框架之中。

与 VFS 相关的系统调用有 `open()`, `read()`, `write()`, `create()` 等。

TODO

VFS 的设计与实现讲解

11.1.2 VFS API 文档

这里是 DragonOS 的内核调试模块文档。

12.1 内核栈 traceback

12.1.1 简介

内核栈 traceback 的功能位于 `kernel/debug/traceback/` 文件夹中。为内核态提供 traceback 的功能，打印调用栈到屏幕上。

12.1.2 API

```
void traceback(struct pt_regs * regs)
```

作用

该接口定义于 `kernel/debug/traceback/traceback.h` 中，将会对给定内核栈进行 traceback，并打印跟踪结果到屏幕上。

参数

regs

要开始追踪的第一层内核栈栈帧（也就是栈的底端）

12.1.3 实现原理

当内核第一次链接之后，将会通过 **Makefile** 中的命令，运行 `kernel/debug/kallsyms` 程序，提取内核文件的符号表，然后生成 `kernel/debug/kallsyms.S`。该文件的 `rodata` 段中存储了 `text` 段的函数的符号表。接着，该文件将被编译为 `kallsyms.o`。最后，**Makefile** 中再次调用 `ld` 命令进行链接，将 `kallsyms.o` 链接至内核文件。

当调用 `traceback` 函数时，其将遍历该符号表，找到对应的符号并输出。

12.1.4 未来发展方向

- 增加写入到日志文件的功能

本章节将介绍如何测试内核，包括手动测试以及自动测试。

我们需要尽可能的对内核进行完善的测试，以便我们能够更好的保证内核的稳定性，且减少其他模块的 debug 难度。

设置完善的测试用例能帮助我们尽可能的检测到问题，防止我们在写新的模块的时候，被已有的模块的一些藏得很深的 bug “背刺一刀”。

由于您难以借助 GDB 等工具进行调试，因此在内核中进行手动测试比应用程序测试要困难一些。

对于一些模块，我们可以使用编写代码进行单元测试，并输出异常信息。遗憾的是，并非所有模块都可以进行单元测试。比如我们常见的内存管理、进程管理等模块都不能进行单元测试。

13.1 内核测试框架

DragonOS 提供了一个测试框架，旨在对内核的一些模块进行自动化测试。内核测试框架位于 `ktest/` 下。

我们可以使用这个测试框架，按照规范编写测试代码，然后在合适的地方使用 `ktest_start()` 创建一个全新的内核线程并发起测试。

13.1.1 使用方法

创建自动测试程序

假如您要对 kfifo 模块进行自动测试，您可以在 ktest/下，创建一个名为 test-kfifo.c 的测试文件，并编写 Makefile。

在 test-kfifo.c 中，包含 ktest_utils.h 和 ktest.h 这两个头文件。

您需要像下面这样，在 test-kfifo.c 中，创建一个测试用例函数表，并把测试用例函数填写到其中：

```
static ktest_case_table kt_kfifo_func_table[] = {
    ktest_kfifo_case0_1,
};
```

然后创建一个函数，作为 kfifo 测试的主函数。请注意，您需要将它的声明添加到 ktest.h 中。

```
uint64_t ktest_test_kfifo(uint64_t arg)
{
    kTEST("Testing kfifo...");
    for (int i = 0; i < sizeof(kt_kfifo_func_table) / sizeof(ktest_case_table); ++i)
    {
        kTEST("Testing case %d", i);
        kt_kfifo_func_table[i](i, 0);
    }
    kTEST("kfifo Test done.");
    return 0;
}
```

编写测试用例

您可以创建一个或多个测试用例，命名为：ktest_kfifo_case_xxxxx. 在这个例子中，我创建了一个测试用例，命名为：ktest_kfifo_case0_1. 如下所示：

```
static long ktest_kfifo_case0_1(uint64_t arg0, uint64_t arg1)
```

这里最多允许我们传递两个参数到测试函数里面。

那么，我们该如何编写测试用例呢？

我们主要是需要设置一些情节，以便能测试到目标组件的每个情况。为了检验模块的行为是否符合预期，我们需要使用 assert(condition) 宏函数，对目标 condition 进行校验。若 condition 为 1，则表明测试通过。否则，将会输出一行 assert failed 信息到屏幕上。

发起测试

我们可以在 `pid ≥ 1` 的内核线程中发起测试。由于 DragonOS 目前尚不完善，您可以在 `process/process.c` 中的 `initial_kernel_thread()` 函数内，发起内核自动测试。具体的代码如下：

```
ktest_start(ktest_test_kfifo, 0);
```

这样就发起了一个内核测试，它会创建一个新的内核线程进行自动测试，您不必担心第一个内核线程会被阻塞。

13.1.2 API 文档

ktest_start

```
pid_t ktest_start(uint64_t (*func)(uint64_t arg), uint64_t arg)
```

描述

开启一个新的内核线程以进行测试

参数

func

测试函数. 新的测试线程将会执行该函数，以进行测试。

arg

传递给测试函数的参数

返回值

测试线程的 pid

assert

```
#define assert(condition)
```

描述

判定 condition 是否为 1，若不为 1，则输出一行错误日志信息：

```
[ kTEST FAILED ] Ktest Assertion Failed, file:%s, Line:%d
```

kTEST

```
#define kTEST(...)
```

描述

格式化输出一行以 [kTEST] file:%s, Line:%d 开头的日志信息。

ktest_case_table

```
typedef long (*ktest_case_table)(uint64_t arg0, uint64_t arg1)
```

描述

ktest 用例函数的类型定义。

该部分文档提供了和处理器架构相关的一些编程实现细节的描述。

14.1 x86-64 相关文档

14.1.1 USB Legacy 支持

简介

usb legacy support 指的是，由 BIOS 提供的，对 USB 鼠标、USB 键盘的支持。在支持并启用 USB Legacy Support 的计算机上，USB 鼠标、键盘由 BIOS 提供模拟，在操作系统看来，就像接入了 PS/2 鼠标、键盘一样。

相关

- 在初始化 USB 控制器时，需要关闭它的 USB Legacy Support

15.1 简介

LibC 是连接用户程序和操作系统的纽带，LibC 为应用程序提供了一系列标准库函数。应用程序可以通过 DragonOS 的 LibC，快速地与操作系统进行交互。DragonOS 的 LibC 主要依照 POSIX 2008 规范实现，与 Linux 下的 glibc 具有相似之处。

15.2 API 文档

15.2.1 ctype.h

函数列表（这里只列出已实现的函数）：

```
``int isprint(int c)`` : 传入一个字符，判断是否可以被输出  
``int islower(int c)`` : 传入一个字符，判断是否是小写字母  
``int isupper(int c)`` : 传入一个字符，判断是否是大写字母  
``int isalpha(int c)`` : 传入一个字符，判断是否是字母  
``int isdigit(int c)`` : 传入一个字符，判断是否是数字
```

（下页继续）

(续上页)

```
``int toupper(int c)`` : 传入一个小写字母字符, 返回这个字母的大写形式

``int tolower(int c)`` : 传入一个大写字母字符, 返回这个字母的小写形式

``int isspace(int c)`` : 传入一个字符, 判断是否是空白字符
```

宏定义:

```
### 暂无用处

``#define _U 01``

``#define _L 02``

``#define _N 04``

``#define _S 010``

``#define _P 020``

``#define _C 040``

``#define _X 0100``

``#define _B 0200``
```

15.2.2 dirent.h

简介

与文件夹有关的头文件。

结构体列表:

```
``struct DIR`` :
```

变量列表:

``int fd`` : 文件夹id (不推荐修改)

``int buf_pos`` : 文件夹缓冲区指针的位置

``int buf_len`` : 文件夹缓冲区的大小 (默认为256)

```
``struct dirent`` :
```

变量列表:

``ino_t(see libc/sys/types.h) ino`` : 文件序列号 (不推荐修改)

``off_t d_off`` : dir偏移量 (不推荐修改)

``unsigned short d_reclen`` : 文件夹中的记录数

``unsigned char d_type`` : 目标的类型(有可能是文件, 文件夹, 磁盘)

``char d_name[]`` : 目标的名字

函数列表 (这里只列出已实现的函数):

```
``DIR opendir(const char *path)``
```

传入文件夹的路径, 返回文件夹结构体

```
``int closedir(DIR *dirp)``
```

传入文件夹结构体, 关闭文件夹, 释放内存

若失败, 返回-1

```
``dirent readdir(DIR *dir)``
```

传入文件夹结构体, 读入文件夹里的内容, 并打包为dirent结构体返回

宏定义:

文件夹类型:

```
``#define VFS_IF_FILE (1UL << 0)``  
  
``#define VFS_IF_DIR (1UL << 1)``  
  
``#define VFS_IF_DEVICE (1UL << 2)``  
  
缓冲区长度的默认值  
  
``#define DIR_BUF_SIZE 256``
```

15.2.3 errno.h

简介:

共享错误号码

属性:

```
``extern int errno`` : 通用错误代码
```

宏定义 (复制自代码, 了解即可):

```
#define E2BIG 1 /* 参数列表过长, 或者在输出buffer中缺少空间  
→ 或者参数比系统内建的最大值要大 Argument list too long.*/  
  
#define EACCES 2 /* 访问被拒绝 Permission denied*/  
  
#define EADDRINUSE 3 /* 地址正在被使用 Address in use.*/  
  
#define EADDRNOTAVAIL 4 /* 地址不可用 Address not available.*/  
  
#define EAFNOSUPPORT 5 /* 地址family不支持 Address family not supported.*/  
  
#define EAGAIN 6 /* 资源不可用, 请重试。 Resource unavailable, try again  
→ (may be the same value as [EWOULDBLOCK]).*/
```

(下页继续)

(续上页)

```

#define EALREADY 7          /* 连接已经在处理 Connection already in progress.*/

#define EBADF 8             /* 错误的文件描述符 Bad file descriptor.*/

#define EBADMSG 9           /* 错误的消息 Bad message.*/

#define EBUSY 10            /* 设备或资源忙 Device or resource busy.*/

#define ECANCELED 11        /* 操作被取消 Operation canceled.*/

#define ECHILD 12           /* 没有子进程 No child processes.*/

#define ECONNABORTED 13     /* 连接已断开 Connection aborted.*/

#define ECONNREFUSED 14     /* 连接被拒绝 Connection refused.*/

#define ECONNRESET 15       /* 连接被重置 Connection reset.*/

#define EDEADLK 16          /* 资源死锁将要发生 Resource deadlock would occur.*/

#define EDESTADDRREQ 17     /* 需要目标地址 Destination address required.*/

#define EDOM 18             /* 数学参数超出作用域 Mathematics argument out of domain.
↳of function.*/

#define EDQUOT 19           /* 保留使用 Reserved*/

#define EEXIST 20           /* 文件已存在 File exists.*/

#define EFAULT 21           /* 错误的地址 Bad address*/

#define EFBIG 22            /* 文件太大 File too large.*/

#define EHOSTUNREACH 23     /* 主机不可达 Host is unreachable.*/

#define EIDRM 24            /* 标志符被移除 Identifier removed.*/

#define EILSEQ 25           /* 不合法的字符序列 Illegal byte sequence.*/

#define EINPROGRESS 26      /* 操作正在处理 Operation in progress.*/

```

(下页继续)

(续上页)

```
#define EINTR 27          /* 被中断的函数 Interrupted function.*/

#define EINVAL 28         /* 不可用的参数 Invalid argument.*/

#define EIO 29            /* I/O错误 I/O error.*/


#define EISCONN 30        /* 套接字已连接 Socket is connected.*/

#define EISDIR 31         /* 是一个目录 Is a directory*/

#define ELOOP 32          /* 符号链接级别过多 Too many levels of symbolic links.*/

#define EMFILE 33         /* 文件描述符的值过大 File descriptor value too large.*/

#define EMLINK 34         /* 链接数过多 Too many links.*/

#define EMSGSIZE 35       /* 消息过大 Message too large.*/

#define EMULTIHOP 36       /* 保留使用 Reserved.*/

#define ENAMETOOLONG 37    /* 文件名过长 Filename too long.*/

#define ENETDOWN 38        /* 网络已关闭 Network is down.*/

#define ENETRESET 39       /* 网络连接已断开 Connection aborted by network.*/

#define ENETUNREACH 40     /* 网络不可达 Network unreachable.*/

#define ENFILE 41          /* 系统中打开的文件过多 Too many files open in system.*/

#define ENOBUFS 42         /* 缓冲区空间不足 No buffer space available.*/

#define ENODATA 43         /* 队列头没有可读取的消息 No message is available on the_
↳STREAM head read queue.*/

#define ENODEV 44          /* 没有指定的设备 No such device.*/

#define ENOENT 45          /* 没有指定的文件或目录 No such file or directory.*/

#define ENOEXEC 46         /* 可执行文件格式错误 Executable file format error.*/
```

(下页继续)

(续上页)

```

#define ENOLCK 47          /* 没有可用的锁 No locks available.*/

#define ENOLINK 48         /* 保留 Reserved.*/

#define ENOMEM 49          /* 没有足够的空间 Not enough space.*/

#define ENOMSG 50          /* 没有期待类型的消息 No message of the desired type.*/

#define ENOPROTOOPT 51     /* 协议不可用 Protocol not available.*/

#define ENOSPC 52          /* 设备上没有空间 No space left on device.*/

#define ENOSR 53           /* 没有STREAM资源 No STREAM resources.*/

#define ENOSTR 54          /* 不是STREAM Not a STREAM*/

#define ENOSYS 55          /* 功能不支持 Function not supported.*/

#define ENOTCONN 56        /* 套接字未连接 The socket is not connected.*/

#define ENOTDIR 57         /* 不是目录 Not a directory.*/

#define ENOTEMPTY 58       /* 目录非空 Directory not empty.*/

#define ENOTRECOVERABLE 59 /* 状态不可覆盖 State not recoverable.*/

#define ENOTSOCK 60        /* 不是一个套接字 Not a socket.*/

#define ENOTSUP 61          /* 不被支持 Not supported (may be the same value as
↪ [EOPNOTSUPP]).*/

#define ENOTTY 62          /* 不正确的I/O控制操作 Inappropriate I/O control operation.
↪ */

#define ENXIO 63           /* 没有这样的设备或地址 No such device or address.*/

#define EOPNOTSUPP 64       /* 套接字不支持该操作 Operation not supported on socket
↪ (may be the same value as [ENOTSUP]).*/

```

(下页继续)

```
#define EOVERFLOW 65      /* 数值过大，产生溢出 Value too large to be stored in data_
↳type.*/

#define EOWNERDEAD 66     /* 之前的拥有者挂了 Previous owner died.*/

#define EPERM 67          /* 操作不被允许 Operation not permitted.*/

#define EPIPE 68          /* 断开的管道 Broken pipe.*/

#define EPROTO 69         /* 协议错误 Protocol error.*/


#define EPROTONOSUPPORT 70 /* 协议不被支持 Protocol not supported.*/

#define EPROTOTYPE 71     /* 对于套接字而言，错误的协议 Protocol wrong type for_
↳socket.*/

#define ERANGE 72         /* 结果过大 Result too large.*/

#define EROFS 73          /* 只读的文件系统 Read-only file system.*/

#define ESPIPE 74         /* 错误的寻道 Invalid seek.*/

#define ESRCH 75          /* 没有这样的进程 No such process.*/

#define ESTALE 76         /* 保留 Reserved.*/

#define ETIME 77          /* 流式ioctl()超时 Stream ioctl() timeout*/

#define ETIMEDOUT 78      /* 连接超时 Connection timed out.*/

#define ETXTBSY 79        /* 文本文件忙 Text file busy.*/


#define EWOULDBLOCK 80     /* 操作将被禁止 Operation would block (may be the same_
↳value as [EAGAIN]).*/

#define EXDEV 81          /* 跨设备连接 Cross-device link.*/
```

15.2.4 fcntl.h

简介

文件操作

函数列表：

```
``int open(const char * path,int options, ...)``
```

传入文件路径，和文件类型（详细请看下面的宏定义），将文件打开并返回文件id。

宏定义（粘贴自代码，了解即可）：

```
#define O_RDONLY 00000000 // Open Read-only

#define O_WRONLY 00000001 // Open Write-only

#define O_RDWR 00000002 // Open read/write

#define O_ACCMODE 00000003 // Mask for file access modes

#define O_CREAT 00000100 // Create file if it does not exist

#define O_EXCL 00000200 // Fail if file already exists

#define O_NOCTTY 00000400 // Do not assign controlling terminal

#define O_TRUNC 00001000 // 文件存在且是普通文件，并以O_RDWR或O_
↪WRONLY打开，则它会被清空

#define O_APPEND 00002000 // 文件指针会被移动到文件末尾

#define O_NONBLOCK 00004000 // 非阻塞式IO模式

#define O_EXEC 00010000 // 以仅执行的方式打开（非目录文件）

#define O_SEARCH 00020000 // Open the directory for search only
```

(下页继续)

(续上页)

```
#define O_DIRECTORY 00040000 // 打开的必须是一个目录

#define O_NOFOLLOW 00100000 // Do not follow symbolic links
```

15.2.5 math.h

简介:

数学库

函数列表:

```
``double fabs(double x)`` : 返回 x 的绝对值

``float fabsf(float x)`` : 返回 x 的绝对值

``long double fabsl(long double x)``: 返回 x 的绝对值

``double round(double x)`` 四舍五入 x

``float roundf(float x)`` 四舍五入 x

``long double roundl(long double x)`` 四舍五入 x

``int64_t pow(int64_t x,int y)`` 返回 x 的 y 次方
```

15.2.6 stdio.h

简介:

向标准输入输出里操作

函数列表:

```
``int64_t put_string(char *str, uint64_t front_color, uint64_t bg_color)``
```

输出字符串 (带有前景色, 背景色)

```
``int printf(const char *fmt, ...)``
```

就是正常的 ``printf``

```
``int sprintf(char *buf, const char *fmt, ...)``
```

就是正常的 ``sprintf``

```
``int vsprintf(char *buf, const char *fmt, va_list args)``
```

格式化, 不建议调用, 请用 printf 或 sprintf 替代。

宏定义

```
### 字体颜色的宏定义
```

```
``#define COLOR_WHITE 0x00ffffff //白``
```

```
``#define COLOR_BLACK 0x00000000 //黑``
```

```
``#define COLOR_RED 0x00ff0000 //红``
```

```
``#define COLOR_ORANGE 0x00ff8000 //橙``
```

```
``#define COLOR_YELLOW 0x00ffff00 //黄``
```

```
``#define COLOR_GREEN 0x0000ff00 //绿``
```

```
``#define COLOR_BLUE 0x000000ff //蓝``
```

```
``#define COLOR_INDIGO 0x0000ffff //靛``
```

```
``#define COLOR_PURPLE 0x008000ff //紫``
```

```
### 无需使用
```

```
``#define SEEK_SET 0 /* Seek relative to start-of-file */``
```

(下页继续)

(续上页)

```
``#define SEEK_CUR 1 /* Seek relative to current position */``  
  
``#define SEEK_END 2 /* Seek relative to end-of-file */``  
  
``#define SEEK_MAX 3``
```

15.2.7 printf.h

不建议引用，需要 printf 函数请引用 stdio.h

15.2.8 stddef.h

简介：

定义了关于指针的常用类型

定义：

```
``typedef __PTRDIFF_TYPE__ ptrdiff_t`` : 两个指针相减的结果类型  
  
``NULL ((void *) 0)`` : 空指针
```

15.2.9 stdlib.h

简介：

一些常用函数

函数列表：

```
``void *malloc(ssize_t size)`` : 普通的 ``malloc``  
  
``void free(void *ptr)`` : 释放内存  
  
``int abs(int x)`` : x 的绝对值  
  
``long labs(long x)`` : x 的绝对值
```

(下页继续)

(续上页)

```
``long long labs(long long x)`` : x 的绝对值
```

```
``int atoi(const char *str)`` 字符串转数字
```

```
``void exit(int status)`` : 普通的 ``exit``
```

15.2.10 string.h

简介:

字符串操作

函数列表:

```
``size_t strlen(const char *s)`` : 返回字符串长度
```

```
``int strcmp(const char *a,const char *b)`` 比较字符串的字典序
```

```
``char* strncpy(char *dst,const char *src,size_t count)``
```

拷贝制定字节数的字符串

dst: 目标地址

src: 原字符串

count: 字节数

```
``char* strcpy(char *dst,const char *src)`` : 复制整个字符串
```

```
``char* strcat(char *dest,const char* src)`` : 拼接两个字符串
```

15.2.11 time.h

简介：

时间相关

时刻以纳秒为单位

结构体：

```
``struct timespec`` : 时间戳

    ### 变量列表：

        ``long int tv_sec`` : 秒

        ``long int tv_nsec`` : 纳秒
```

宏定义：

```
``#define CLOCKS_PER_SEC 1000000`` 每一秒有1000000个时刻（纳秒）
```

函数列表：

```
``int nanosleep(const struct timespec *rdtp, struct timespec *rmtp)``

    休眠指定时间

    rdtp : 指定休眠的时间

    rmtp : 返回剩余时间

``clock_t clock()`` : 获得当前系统时间
```

15.2.12 unistd.h

简介：

也是一些常用函数

函数列表：

`int close(int fd)` : 关闭文件

`ssize_t read(int fd, void *buf, size_t count)` : 从文件读取

传入文件id, 缓冲区, 以及字节数

返回成功读取的字节数

`ssize_t write(int fd, void const *buf, size_t count)` : 写入文件

传入文件id, 缓冲区, 字节数

返回成功写入的字节数

`off_t lseek(int fd, off_t offset, int whence)` : 调整文件访问位置

传入文件id, 偏移量, 调整模式

返回结束后的文件访问位置

`pid_t fork(void)` : fork 当前进程

`pid_t vfork(void)` : fork 当前进程, 与父进程共享 VM, flags, fd

`uint64_t brk(uint64_t end_brk)` : 将堆内存调整为end_brk

若end_brk 为 -1, 返回堆区域的起始地址

若end_brk 为 -2, 返回堆区域的结束地址

否则调整堆区的结束地址域, 并返回错误码

`void *sbrk(int64_t increment)` :

将堆内存空间加上offset (注意, 该系统调用只应在普通进程中调用, 而不能是内核线程)

(下页继续)

(续上页)

```
    increment : 偏移量

``int64_t chdir(char *dest_path)``

    切换工作目录 (传入目录路径)

``int execv(const char* path, char * const argv[])`` : 执行文件
    path : 路径
    argv : 执行参数列表

``extern int usleep(useconds_t usec)`` : 睡眠usec微秒
```

这里是所有 `libc` 头文件的集合, 在代码里可以这样引用: `#include<libc/src/xxx.h>`

15.3 设计文档

[内容待完善]

CHAPTER 16

系统调用 API

16.1 简介

DragonOS 社区欢迎您的加入！学习技术本身固然重要，但是以下这些文档将会帮助您了解 DragonOS 社区需要什么。

阅读这些文档将会帮助您参与到开发之中，并且能让您的代码能更快合并到主线。

17.1 开发流程介绍

作为一名想要参与开发的新人，您可能迫切想要了解如何参与开发，仔细阅读这篇文章将能帮助您了解整个开发流程，以及一些注意事项。

注：本文参考了 Linux 文档中的一些思想、内容，非常感谢 Linux 社区的工作者们的经验！

17.1.1 1. 概要

对于新人而言，参与开发的过程主要包括以下几步：

- **运行 DragonOS**：按照文档：[构建 DragonOS](#)中的教程，编译 DragonOS，并成功运行。在运行过程中，如果有任何的疑问，欢迎您在交流群或者 BBS 上进行反馈！
- **联系 Maintainer**：您可以通过邮箱longjin@DragonOS.org或者 QQ184829088 与龙进取得联系，或者是对应的模块的开发者进行联系（目前您可以通过发行日志上的邮箱与他们建立联系，在将来我们将编写一份“Maintainers List”以便于您能快速找到要联系的人）。为了节省您的时间，请简要说明：
 - 如何称呼您

- 您目前掌握的技术能力
 - 您希望为 DragonOS 贡献什么功能，或者进行某个方面的开发，亦或者是想要按照社区的需要来参与新功能开发及 bug 的修复。
 - 如果您是来自高校/科研单位/企业/组织的代表，希望与社区开展合作研究、开发。那么，除使用 QQ 交流之外，还请麻烦您通过您的教师邮箱/学生邮箱/企业邮箱向contact@DragonOS.org发送一封相关内容的邮件！这么做的目的是为了确认您是来自您的单位，而不是网络上其他人员冒充您的身份。
- **加入工作群**：在进一步了解，确认您愿意参与开发后，我们将邀请您加入工作群。
 - **开始开发**：正式开始代码开发工作。

备注：一些小功能的改进以及 Bug 修复并不一定需要提前与社区建立正式的联系，对于这些 patch，您可以直接开发，然后在 Github 上进行 Pull Request. 这也是可以的。

但是，如果您愿意的话，与 Maintainer 联系会是一个更好的选择。

17.1.2 2. 开发过程是如何运作的？

如今的 DragonOS 由于正处于开发的早期阶段，开发者数量不超过 50 人，因此，现在 DragonOS 的开发过程是通过比较松散的方式组织起来的。

2.1. 版本发布周期

自从 2022 年 11 月 6 日，DragonOS 发布第一个版本以来，版本发布就被定义为 15~21 天发布一个更新版本。由于开发人员数量仍然较少，因此，目前这个时间是 21 天。我们将版本号定义为 $x.y.z$ 的格式，每 21 天发布一个 z 版本。在积累了 2~3 个月后，当 DragonOS 引入了足够的新功能，则发布一个 y 版本。请注意，我们仍未定义 x 版本的发行周期。当前， x 版本号仍然是 0。

创建没有 BUG 的、具有尽可能少 BUG 的代码，是每个开发者的目标之一。我们希望在每个 y 版本发布时，能够修复已知的问题。然而，由于在操作系统中，影响问题的变量太多了，以至于我们只能尽全力去减少 BUG，我们无法保证 y 版本不存在 bug。

2.2. 每个补丁的生命周期

当您向 DragonOS 的仓库发起一次 PR，那么这次 PR 就是一个补丁。我们会对您的补丁进行 Review，以确保每个补丁都实现了一个希望在主线中进行的更改。并且，Maintainer 或者感兴趣的小伙伴会对您的补丁提出修改意见。当时机合适时，您的代码将被合入主线。

如果您的补丁的规模比较小，那么，它将会比较快的被合入主线。如果补丁的规模较大，或者存在一些争议，那么我们需要对其进行进一步的讨论及修改、审查，直到它符合要求。

每个 Patch 都会经历这么一个过程（这只是一个概述，详细的描述请看后文）：

- **设计**：在这个阶段，我们将明确，这个补丁将要实现什么功能，或者是解决什么问题，以及我们将要采用什么样的方式去完成它。通常来说，这项工作是开发者自己完成的。但是，**我们建议您，在设计了这个补丁之后，能够把您的设计方案公开，和大家讨论这项工作**。闭门造车容易出错，在与大家沟通的过程中，则能及早的发现设计上的错误，从而节省很多的时间。
- **代码编写**：经过了设计阶段，您已经能够明白自己要实现的到底是一个什么样的东西。您在这个阶段进行代码编写、调试。
- **代码审查**：当完成代码编写后，您可以通过 Github 发起一个 PR，然后您的补丁进入了代码审查阶段。在这一阶段，开发者们，或者是 Maintainer 会与您进行沟通交流，对您的代码进行评论，以及对发现的问题提出修改建议。
- **合并主线**：如果一切顺利，那么代码将会被合并入主线。若该补丁在合并主线后，被发现具有较大的问题，那么它可能会被回退。重新进入前面的阶段，进行修改。
- **长期维护**：虽然说，代码被合并之后，原来的开发人员可能会在很久一段时间后，不再理会这些代码，但是这种行为可能会给其他开发者们留下不好的印象。其实，当补丁被合并入主线后，其他开发人员会尝试继续维护这些代码，这样能够很大程度的减轻您的维护负担。但是，如果想要这些代码能够长期的被保留下来，持续的发光发热，那么离不开原始开发人员的支持（不然的话，后来的人可能难以理解、维护这些代码），这是一件很有意义的事情。

对于没有参与过开源项目的同学来说，他们可能会想当然的，简单的把上述流程简化成**合并主线**这一个步骤，这样是不可取的。因为这样会导致很多问题，包括但不限于“写了代码但是效果很差”、“写的东西由于无法满足项目的需求，而不能被合并”。

2.3. 开发工具

从上面的描述可以看出，为了让开发人员们能高效地进行协作，那么必须使用版本控制工具来管理这个过程。目前，DragonOS 使用 Git 来进行源代码管理。它是一个非常优秀的工具，这是不必多说的。对于每个开发者而言，Git 的使用是一项必备的技能；哪怕您只是想学习 DragonOS 的源代码，您也需要 git 来获取、同步最新的代码。虽然 Git 的使用，对于新手来说，有些困难，但是经过一些时间的学习后，还是可以掌握的。

git 的官方网站为<https://git-scm.com/>

2.4. 沟通交流

DragonOS 的主要开发工作是通过飞书以及 bbs 进行的。对于正准备参与的开发者而言，您可以加入我们的交流讨论 QQ 群，具体的群号可以在[与社区建立联系](#)一文中找到。

何时使用即时通讯软件？我们在飞书上创建了工作群，为提高即时沟通的效率，这个群仅供那些真正有意愿、且正在进行或准备进行（能够保证会进行）代码开发的开发者们加入。

何时使用 BBS？对于一些正式的、需要大家广泛参与，或者是能够帮助尚未参与开发的同学了解当前的开发进度的主题，请您在<https://bbs.DragonOS.org>上，使用类似于写信件一样的，正式的语言，完整地描述清楚您想表达的内容。这样有助于更多的人快速明白您要表达的是什么，也能提高整体的沟通效率。并且，bbs 能够长期保存以往的帖子，这样后来者能更方便的了解“当初开发的时候，人们究竟是怎么想的”。

关于交流讨论会除由于法定节假日放假，或特殊情况以外，我们每周末都会召开线上交流讨论会，同步介绍每周的进展。社区成员可以在这里分享自己的方案设计或是一些操作系统相关的知识（分享的话，需要提前跟会议主持人提出，以便妥善安排）。

如何提问？下面这些建议能够帮助您与他人开展高效率的对话：

- **对于具有主题性的问题，在 bbs 上发帖进行讨论。**这样能够让讨论更具有目标性。当谈论到新的主题的时候，请开一个新的帖子，并在原来的帖子中，添加对特定的子问题所在的主题链接。
- **请礼貌的交流。**文明的语言能够减少不必要的冲突。技术意见上的冲突是思维的碰撞，但是如果涉及到了不文明的语言，或者在非技术层面，对他人进行攻击，这将破坏和谐讨论的氛围，这是我们反对的。如果有人试图激怒你，请忽略他的消息，别理他就好了。
- **在提问之前，请确保您已经搜索了 bbs 以及互联网上的解决方案，并描述清楚您的问题的上下文情景、您的思考以及网络上的解决方案。**一些开发人员会对“明显没有进行认真思考”的问题，表现出不耐烦的态度（因为未经思考的问题会浪费他们大量的时间）。
- **当别人向您提问时，**请您耐心听他人的问题。如果您认为对方的问题过于简单或是未经思考，还请您为对方指个路，告诉对方，在哪些地方，使用什么样的方式搜索，能够得到对解决问题有帮助的资料。有时候，**新手需要的是一个指路人**，他会非常感谢您的！

2.5. 如何入门开发？

DragonOS 原采用 C 语言进行开发，目前正在用 Rust 重构原有代码、开发新的模块，也就是说，除非您要进行对 C 语言代码的 BUG 修复，否则，其余的开发工作，我们都建议您通过 Rust 完成。因为，它从语言层面解决那些让我们头疼的内存安全问题。从长期来看，能够提升开发效率以及软件质量。

如何开发第一个补丁，是一个非常常见的问题。可以理解的是，个人开发者面对这样一个项目，常常会不知道从哪个地方开始入手。这是一件很正常的事情，因此我们建议您通过上文提到的方式，与社区建立联系，了解目前社区正在做什么，以及需要什么。

对于一个新的参与者来说，我们建议您从这样一个步骤开始：

阅读文档，编译、运行 DragonOS，并且尝试使用它目前已有的功能。

然后，您可以通过查看 DragonOS 的 GitHub 仓库的 project 面板，看看目前仍有哪些待解决的问题。可以肯定的是，永远不会缺少待解决的问题，您在解决这些问题的过程中，能够获得一些宝贵的经验。

17.1.3 3. 早期设计

对于软件开发而言，写代码永远不是第一步。在编写代码之前，进行一些必要的设计（提出架构、技术方案），是项目成功的基础工作。在新的补丁开发的早期，花费一些时间进行计划和沟通，往往能够在接下来的阶段节省更多的时间。

3.1. 定义我们要解决的问题

与大多数的工程项目一样，在 DragonOS 中进行开发，首先需要清晰的描述要解决的问题。只有精准的定义了问题，才能找到正确的解决方案。有时候，我们能很轻易的定义问题，比如“编写串口驱动程序，使得它能把屏幕上打印的字符，输出到串口”。

但是，有时候，**我们容易将真正的问题与某个解决方案相混淆**，并且还没意识到这一点。

在 2022 年 10 月时，我发现，在真机调试的时候，需要频繁的拔插硬盘（先连接到电脑，待数据拷贝完毕后，再连接到测试机）。我对这一过程非常的不满，因为很浪费时间。我的直觉想法是：“有没有一种设备，能够一头连接电脑，另一头连接测试机的 SATA 接口。从测试机的角度来看，这是一块硬盘；测试机对这块磁盘的操作，都转换为了对我的电脑上面的一个磁盘镜像文件的操作。”我的想法就是：“购买/定制一款设备，能够实现上面的这个功能，那就能解决频繁拔插硬盘的烦恼了！”然后我为了寻找这样的设备，询问了一些硬件公司，他们的开价都在 2 万元左右。

我在上面的这个过程中，就犯了一个错误：将真正的问题与某个解决方案相混淆了。真正的问题是：“解决需要频繁的拔插硬盘”，但是，在我的思考的过程中，不知不觉间，将问题转换成了“如何实现一款具有硬盘模拟功能的设备”。后面这个问题，只是某个解决方案下，需要解决的问题，并不是我们要解决的根本问题。

对于要解决的根本问题，事实上有更好的解决方案：“制作一个类似于开关一样的转换器，当数据从电脑拷贝到磁盘后，把开关拨向另一边，使得电路与测试机接通”。这个方案的成本估摸着就十几二十块钱。

上面的这个故事，告诉我们的是，**在早期设计阶段，我们应当关注的是问题本身——而不是特定的解决方案**。

我们需要关注系统的稳定性、长期的可维护性，解决方案必须考虑到这两点。由于系统比较复杂，因此，请您在开始编码之前，与社区的小伙伴讨论您的设计方案，以便您的方案能充分地，从全局的角度，考虑到系统的稳定性、可维护性。

因此，在开发的早期，我们应当对以下三个问题，拥有一个答案：

- 要解决的本质问题是什么？
- 这个问题会影响哪些方面/哪些用户？提出的解决方案应当解决哪些用例、场景？
- DragonOS 目前在解决该问题的方面，具有哪些不足/问题？

只有考虑清楚了上面三个问题，讨论的解决方案才是有意义的。这是一个架构设计的过程，需要进行仔细的思考。尽管我们目前提倡敏捷开发，但是前期的架构设计仍然是非常重要的。

3.2. 早期讨论

在实施开发之前，与社区的成员们进行讨论是非常有意义的。这能够通过多种方式节省您的时间，并减少许多不必要的麻烦：

- DragonOS 可能以您不知道、不理解的方式，已经解决了相关的问题。DragonOS 里面的一些特性、功能细节不是很明显，他们不一定会被写进文档。也许这些细节只是在某个不起眼的注释里面提到了，因此您很难注意到这些。这种细节可能只有一些开发人员知道。因此，与社区沟通能够避免您进行重复的工作。
- 您提出的解决方案中，可能会有一些东西，由于一些原因（比如方案中的一些设计会在将来造成问题、方案的架构设计具有明显缺陷），无法合入主线。
- 其他的开发人员可能已经考虑过这个问题；他们可能有更好的解决方案，或者是更好的想法。并且，他们可能愿意帮助你一起完善你的解决方案。

Linux 文档中提到：闭门造车式的设计和开发，所产生的代码总会有问题，这些问题只有在发布到社区里面的时候才会暴露出来。因此，我们必须吸取前人之鉴，通过与社区开发人员进行早期讨论，从而避免大量的痛苦和额外的工作。

3.3. 在何时发布帖子？

如果可能的话，在开发的早期阶段发布您的计划、设计，会是一个不错的选择。发帖的时候，您可以描述您正在解决的问题，以及已经制定的一些计划。包括但不限于：如何将设计付诸实施。您在社区内发布帖子，不仅能够帮助您获得一些有用的建议，也能帮助整个 DragonOS 社区提供有用的信息，使得社区沟通更高效。

在这个阶段，可能您发布的帖子并不会引来很多评论，这并不一定意味着您做的不好，或者大家对您所做的工作不感兴趣。当然，也不能就此认为您的设计、想法不存在问题。可能只是因为大家比较忙，看了您的帖子之后，了解到了您的工作，但是大家并不一定有时间进行回复。（但是事实上您发布的信息对他人来说是有用的）

在这种情况下，请不要气馁，您最好的做法就是，继续您的工作，并且不时地在您的帖子下分享您的工作，这样能够让社区的成员们随时了解到您的最新进展。

3.4. 获得您所在的组织的支持

如果您对 DragonOS 的开发工作，是在您的公司内完成的。那么，很显然，在您把计划、代码发布到社区论坛之前，您必须取得您的经理或上级的许可。

同时，请注意，根据我们的授权许可，基于 DragonOS 操作系统的内核、官方开源的用户库而开发的代码，或者为 DragonOS 操作系统本身而开发的代码，根据开源授权许可，必须同样以 GPLv2 协议进行授权发布。如果您所在的组织，违背了 GPLv2 协议中的要求，以除 GPLv2 以外的协议开放源代码，或者是进行“闭源使用”，那么 DragonOS 社区对您的公司/组织所授予的使用 DragonOS 源代码的授权，将会被自动撤销。这将会面临一系列的法律问题。因此，在这个问题上，公司的管理人员、法律人员如果能越早地就公司要在 DragonOS 中开发的软件项目达成一致，将能促进您的公司在该项目上的进展。

如果您的公司的项目/或者是您研究的项目根据您所在组织的保密规定，不能将其进行过早的披露，那也没有问题。只要您的公司能够确保这部分代码，在其编译而成的二进制产品被发布之时，按照 GPLv2 协议进行开源，并向开源社区贡献这部分代码即可。

17.1.4 4. 如何正确的编写代码

17.1.5 5. 发起 Pull Request

17.1.6 6. 后期跟进

17.1.7 7. 另外的一些话题

17.1.8 8. 更多信息

17.1.9 9. 结语

17.2 C 语言代码风格

这份文档将会简要的介绍 DragonOS 的 C 语言代码风格。每个人的代码风格都各不相同，这是一件非常正常的事情。但是，对于一个开源项目的可维护性而言，我们希望制定一些代码规范，以便包括您在内的每个开发者都能在看代码的时候更加舒服。一个充斥着各种不同代码风格的项目，是难以维护的。

我们在这里提出一些建议，希望您能够尽量遵循这些建议。这些建议与 Linux 的代码规范相似，但又略有不同。在变量命名上，DragonOS 采用 Linux 的风格；对于缩进，DragonOS 采用 Microsoft 风格。

17.2.1 0. 代码格式化工具

在提出下面的建议之前，我们建议您在开发的时候使用 Visual Studio Code 的 C/C++ Extension Pack 插件作为代码格式化工具。这些插件能为您提供较好自动格式化功能，使得您的代码的基本格式符合 DragonOS 的要求。

当您在编码时，经常性的按下 `Ctrl+shift+I` 或您设置的代码格式化快捷键，能帮助您始终保持良好的代码格式。

17.2.2 1. 缩进

一个制表符的宽度等于 4 个空格。代码的缩进是按照制表符宽度 (在多数编辑器上为 4 个字符) 进行缩进的。

这样能够使得您的代码变得更加容易阅读, 也能更好的看出代码的控制结构。这样能避免很多不必要的麻烦!

举个例子: 在 `switch` 语句中, 将 `switch` 和 `case` 放置在同一缩进级别。并且将每个 `case` 的代码往右推进一个 `tab`。这样能让代码可读性变得更好。

```
switch (cmd)
{
case AHCI_CMD_READ_DMA_EXT:
    pack->blk_pak.end_handler = NULL;
    pack->blk_pak.cmd = AHCI_CMD_READ_DMA_EXT;
    break;
case AHCI_CMD_WRITE_DMA_EXT:
    pack->blk_pak.end_handler = NULL;
    pack->blk_pak.cmd = AHCI_CMD_WRITE_DMA_EXT;
    break;
default:
    pack->blk_pak.end_handler = NULL;
    pack->blk_pak.cmd = cmd;
    break;
}
```

17.2.3 2. 分行

我们建议, 每行不要超过 120 个字符。如果超过了, 除非有必要的理由, 否则应当将其分为两行。

在分行时, 我们需要从被分出来的第二行开始, 比第一行的起始部分向右进行一个缩进, 以表明这是一个子行。使用代码格式化的快捷键能让你快速完成这件事。

对于一些日志字符串而言, 为了能方便的检索到他们, 我们不建议对其进行分行。

对于代码的分行, 请不要试图通过以下方式将几个语句放置在同一行中, 这样对于代码可读性没有任何好处:

```
// 错误示范 (1)
if(a) return 1;

// 错误示范 (2)
if(b)
    do_a(), do_b();
```

17.2.4 3. 大括号和空格

3.1 大括号

大括号的放置位置的选择是因人而异的，主要是习惯原因，而不是技术原因。我们推荐将开始括号和结束括号都放置在一个新的行首。如下所示：

```
while(i<10)
{
    ++i;
}
```

这种规定适用于所有的代码块。

这么选择的原因是，在一些编辑器上，这样放置括号，编辑器上将会出现辅助的半透明竖线，且竖线两端均为括号。这样能帮助开发者更好的定位代码块的层次关系。

下面通过一些例子来演示：

在下面这个代码块中，我们需要注意的是，else if 语句需要另起一行，而不是跟在上一个} 后方。这是因为我们规定 { 必须在每行的起始位置，并且还要保持缩进级别的缘故。

```
if (*fmt == ' ')
{
    ++fmt;
}
else if (is_digit(*fmt))
{
    field_width = skip_and_atoi(&fmt);
}
```

当循环中有多个简单的语句的时候，需要使用大括号。

```
while (condition)
{
    if (test)
        do_something();
}
```

当语句只有 1 个简单的子句时，我们不必使用大括号。

```
if(a)
    return 1;
```

3.2 空格

对于大部分关键字，我们需要在其后添加空格，以提高代码的可读性。

请您在所有这些关键字后面输入一个空格：

```
if, switch, case, for, do, while
```

关键字 `sizeof`、`typeof`、`alignof`、`__attribute__` 的后面则不需要添加空格，因为使用他们的时候，就像是使用函数一样。

对于指针类型的变量，`*` 号要贴近变量名而不是贴近类型名。如下所示：

```
char *a;
void *func(char* s, int **p);
```

在大多数二元和三元运算符周围（在每一侧）使用一个空格，如下所示：

```
= + - < > * / % | & ^ <= >= == != ? :
```

这些一元运算符后方没有空格

```
& * + - ~ ! sizeof typeof alignof __attribute__ defined
```

特殊的例子，以下运算符的前后都不需要空格：

```
++ -- . ->
```

17.2.5 4. 命名

DragonOS 中的命名规范不使用诸如 `TempValue` 这样的驼峰命名法的函数名，而是使用 `tmp` 这样言简意赅的命名。

注意，这里指的是我们在整个项目内都不希望使用驼峰命名法。并不意味着程序员可以随便的使用一些难以理解的缩写来作为变量名。

对于全局变量或者全局可见的函数、结构体而言，我们需要遵循以下的命名规定：

- 名称需要易于理解，且不具有歧义。如：对于一个计算文件夹大小的函数而言，我们建议使用 `count_folder_size()` 来命名，而不是 `cntfs()` 这样令其他人头大的命名。
- 全局的，非 `static` 的名称，除非有特别的必要，命名时需要遵循以下格式：模块名缩写前缀 _ 函数/变量名。这样的命名能便于别人区分这个名称位于哪个模块内，也减少了由于命名冲突所导致的麻烦。
- 不需要让其他代码文件可见的全局名称，必须添加 `static` 修饰符。

对于函数内的局部变量而言，命名规范则是需要言简意赅。过长的名称在局部变量中没有太大的意义。

【文档未完成，待进一步完善】

17.3 Rust 语言代码风格

这篇文档将会介绍 DragonOS 中的 Rust 语言代码风格。随着开发的进行，这些风格可能会发生变化，但是我们会尽量保持风格的一致性。

17.3.1 1. 命名

这部分基于 Rust 语言圣经中的命名规范进行修改，本文未提及的部分，请参考 Rust 语言圣经中的命名规范。

17.3.2 2. 格式

2.1 缩进

请在提交代码之前，使用 `cargo fmt` 命令对代码进行格式化。

2.2 函数返回值

尽管 Rust 可以返回函数的最后一行的语句的值，但是，这种方式会使代码的可读性变差。因此，我们推荐您在函数的最后一行使用 `return` 语句，而不是直接返回值。

```
// 不推荐
fn foo() -> i32 {
    1 + 2
}

// 推荐
fn foo() -> i32 {
    return 1 + 2;
}
```

2.3 错误处理

DragonOS 采用返回 Posix 错误码作为模块间错误处理的方式。为了确保在模块之间，错误处理代码的一致性，我们推荐在发生错误的时候，返回 `Err(i32)`，并且 `posix` 错误码的值应当是负数。这样做的优点尤其体现在跨模块调用函数时，可以直接返回错误码，从而降低错误处理代码的耦合度。

```
// 函数跨越模块边界时（由其他模块调用当前函数），不推荐
fn foo() -> Result<(), CustomErr> {
    if 1 + 2 == 3 {
        return Ok(());
    }
}
```

(下页继续)

(续上页)

```

    } else {
        return Err(CustomErr::error);
    }
}

// 函数跨越模块边界时（由其他模块调用当前函数），推荐
fn foo() -> Result<(), i32> {
    if 1 + 2 == 3 {
        return Ok(());
    } else {
        return Err(-(EINVAL as i32));
    }
}

```

在模块内部，您既可以采用返回自定义错误 `enum` 的方式，也可以采用返回 `Posix` 错误码的方式。但是，我们推荐您在模块内部，采用返回自定义错误 `enum` 的方式，这样可以使错误处理代码更加清晰。

17.3.3 3. 注释

DragonOS 的注释风格与 Rust 官方的不太一样。但是，我们仍然推荐您在代码中加入尽可能多的有效注释，以便于其他人理解您的代码。并且，变量、函数等声明，遵守第一节中提到的命名规范，使其能够“自注释”。

3.1 函数注释

函数注释应该包含以下内容：

- 函数的功能
- 函数的参数
- 函数的返回值
- 函数的错误处理
- 函数的副作用或者其他的需要说明的内容

函数注释的格式如下：

```

/// @brief 函数的功能
///
/// 函数的详细描述
///
/// @param 参数1 参数1的说明
/// @param 参数2 参数2的说明

```

(下页继续)

(续上页)

```
///  
/// @return 返回值的说明
```

如果函数的返回值是 `Result` 类型，那么返回值应当这样进行解释：

```
/// @return Ok(返回值类型) 返回值的说明  
/// @return Err(错误值类型) 错误的说明
```


CHAPTER 18

与社区建立联系

社区公共邮箱: contact@DragonOS.org

DragonOS 负责人: longjin

工作邮箱: longjin@DragonOS.org

开发交流 QQ 群: 115763565

DragonOS 官网: <https://DragonOS.org>

CHAPTER 19

发行日志

这里是 DragonOS 的发行日志，会记录 DragonOS 的每一个版本的更新内容。

19.1 V0.1.5

备注： 本文作者：龙进 longjin@RinGoTek.cn

2023 年 3 月 13 日

19.1.1 贡献者名单

DragonOS V0.1.5 版本由以下小伙伴贡献代码：

- 龙进 longjin@DragonOS.org
- 关锦权 guanjinquan@DragonOS.org
- 苏匡韬 sujintao@DragonOS.org
- 孔维超 kongweichao@DragonOS.org
- 侯嘉滢 houjiaying@DragonOS.org
- 吴宇健 wuyujian@DragonOS.org
- 周于匡 zhouyuzhe@DragonOS.org

- [Satin Wuker 74630829+SatinWuker@users.noreply.github.com](#)

19.1.2 赞助者名单

感谢以下同学的赞赏，我们将不断努力！

- 万晓兰
- David Wen
- [Seele.Clover](#)
- [FindWangHao](#)
- 叶锦毅
- 林
- Albert
- [TerryLeeSCUT](#) · [GitHub](#)
- slientbard
- 悟

19.1.3 更新内容-内核

- scheduler: doc: 实时进程调度器文档 (#163)
- scheduler: rt: RTQueue 改用双向链表存储 (#174)
- scheduler: load balance: 多核负载均衡 (#193)
- Semaphore: new: 新增了 rust 实现的信号量 (#183)
- mm: refactor: 重构了 MMIO 地址分配器 (#184)
- RwLock: new: 新增了 rust 实现的读写锁 (#186)
- driver: update: 完善 pci 的功能 (#194)
- driver: new: VirtIO 网卡驱动（仍存在问题） (#194)
- driver: refactor: Rust 版本的 AHCI 驱动 (#198)
- block io: delete: 移除 Block IO 调度器. (#196)
- filesystem: refactor: 新版的 VFS (#198)
- filesystem: refactor: 新版的 ProcFS (#198)
- filesystem: refactor: 新版的 DevS (#198)
- filesystem: new: RamFS 内存文件系统 (#198)

- filesystem: new: FAT12/FAT16/FAT32 文件系统 (#198)
- filesystem: new: 新的设备、块设备抽象 (#198)

19.1.4 更新内容-用户环境

- libc: 调整，将所有的 app 直接链接到 libc.a 中，而不是都执行一遍”搜索.o”的过程 (#171)

19.1.5 更新内容-其他

- bootstrap: 解决 ubuntu2210 版本无法正确编译 grub，以及正确安装 qemu 的问题 (#176)
- toolchain: 添加 rust 的 bare bone 工具链 (#197)

19.1.6 更新内容-软件移植

无

19.1.7 源码、发布版镜像下载

您可以通过以下方式获得源代码:

通过 Git 获取

- 您可以访问<https://github.com/DragonOS-Community/DragonOS/releases>下载发布版的代码，以及编译好的，可运行的磁盘镜像。
- 我们在 gitee 上也有镜像仓库可供下载：<https://gitee.com/DragonOS/DragonOS>

通过 DragonOS 软件镜像站获取

为解决国内访问 GitHub 慢、不稳定的问题，同时为了方便开发者们下载 DragonOS 的每个版本的代码，我们特意搭建了镜像站，您可以通过以下地址访问镜像站：

您可以通过镜像站获取到 DragonOS 的代码压缩包，以及编译好的可运行的磁盘镜像。

- <https://mirrors.DragonOS.org>
- <https://mirrors.DragonOS.org.cn>
- 国内镜像加速：[\[https://mirrors.ringotek.cn/\]](https://mirrors.ringotek.cn/) (<https://mirrors.ringotek.cn/>)

19.1.8 开放源代码声明

备注：为促进 DragonOS 项目的健康发展，DragonOS 以 GPLv2 开源协议进行发布。所有能获得到 DragonOS 源代码以及相应的软件制品（包括但不限于二进制副本、文档）的人，都能享有我们通过 GPLv2 协议授予您的权利，同时您也需要遵守协议中规定的义务。

这是一个相当严格的，保护开源软件健康发展，不被侵占的协议。

对于大部分的善意的人们而言，您不会违反我们的开源协议。

我们鼓励 DragonOS 的自由传播、推广，但是请确保所有行为没有侵犯他人的合法权益，也没有违反 GPLv2 协议。

请特别注意，对于违反开源协议的，尤其是**商业闭源使用以及任何剽窃、学术不端行为将会受到严肃的追责**。（这是最容易违反我们的开源协议的场景）。

并且，请注意，按照 GPLv2 协议的要求，基于 DragonOS 修改或二次开发的软件，必须同样采用 GPLv2 协议开源，并标明其基于 DragonOS 进行了修改。亦需保证这些修改版本的用户能方便的获取到 DragonOS 的原始版本。

您必须使得 DragonOS 的开发者们，能够以同样的方式，从公开渠道获取到您二次开发的版本的源代码，否则您将违反 GPLv2 协议。

关于协议详细内容，还敬请您请阅读项目根目录下的 **LICENSE** 文件。请注意，按照 GPLv2 协议的要求，**只有英文原版才具有法律效力**。任何翻译版本都仅供参考。

开源软件使用情况

DragonOS 在开发的过程中，参考了一些开源项目的设计，或者引入了他们的部分代码，亦或是受到了他们的启发。现将他们列在下面。我们对这些开源项目的贡献者们致以最衷心的感谢！

格式：< 项目名 > - < 链接 > - < 开源协议 >

- Linux - <https://git.kernel.org/> - GPLv2
- skiftOS - <https://github.com/skiftOS/skift> - MIT
- FYSOS - <https://github.com/fysnet/FYSOS> - [FYSOS' License](#)
- LemonOS - <https://github.com/LemonOSProject/LemonOS.git> - BSD 2-Clause License
- LZ4 - <https://github.com/lz4/lz4> - BSD 2-Clause license
- SerenityOS - <https://github.com/SerenityOS/serenity.git> - BSD 2-Clause license
- MINE - 《一个 64 位操作系统的设计与实现》田宇; 人民邮电出版社
- chcore - 《现代操作系统：设计与实现》陈海波，夏虞斌; 机械工业出版社
- SimpleKernel - <https://github.com/Simple-XX/SimpleKernel> - MIT

- rcore-fs - <https://github.com/rcore-os/rcore-fs.git> - MIT
- redox - <https://gitlab.redox-os.org/redox-os/redox> - MIT

19.1.9 当前版本的所有提交记录

```
commit 84407d360511c7699938a0f245ae33ff76f16b17
```

```
Author: login <longjin@DragonOS.org>
```

```
Date: Mon Mar 13 00:26:04 2023 +0800
```

bugfix: 解决touch命令失败的问题 (#199)

* bug fix : 解决touch命令失败的问题

```
commit 004e86ff19727df303c23b42c7a271b9214c6898
```

```
Author: login <longjin@DragonOS.org>
```

```
Date: Sun Mar 12 22:36:11 2023 +0800
```

新版文件系统重构完成 (#198)

1. 重构: VFS
2. 重构: ProcFS
3. 重构: DevFS
4. 重构: FAT32
5. 重构: AHCI 驱动
6. 新增: RamFS
7. 新增: MountFS
8. 新增: FAT12
9. 新增: FAT16
10. 重构: 设备抽象

Co-authored-by: guanjinquan <1666320330@qq.com>

Co-authored-by: DaJiYuQia <88259094+DaJiYuQia@users.noreply.github.com>

```
commit 17041e0e307eaf9e8d8ddbdfaf186cd1f10f1bc0
```

```
Author: login <longjin@DragonOS.org>
```

```
Date: Sun Mar 12 21:04:37 2023 +0800
```

添加rust的bare bone工具链 (#197)

```
commit 26d84a31393c50063ff416bc509316e8d342028c
```

```
Author: YJwu2023 <119829947+YJwu2023@users.noreply.github.com>
```

```
Date: Sat Mar 11 21:09:50 2023 +0800
```

(下页继续)

(续上页)

新增VirtIO网卡驱动 (#194)

- * 修复内存bug与grub安装脚本的错误
- * 修改小bug
- * PCI增加功能与virtio-net驱动
- * little fix
- * virtio-net小修改

commit 1d48996375149279a721777b2c600e1b5c3ee1b5

Author: kong <45937622+kkkkkkong@users.noreply.github.com>

Date: Sat Mar 11 18:17:35 2023 +0800

多核负载均衡 (#193)

- * feat(sched): CPU负载检测初步实现
- * fix(smp):调整smp中的apic的头文件声明
- * fix(smp):简单的负载均衡算法实现
- * fix(sched):抽离负载均衡方法
- * fix(sched):修改rt中的运行队列bug, 调整负载均衡逻辑
- * fix(process):移除无用测试代码
- * reformat code

commit ef9f9732b09f78d7192f1d0dd3b41be655fb0914

Author: houmkh <100781004+houmkh@users.noreply.github.com>

Date: Thu Mar 9 23:31:25 2023 +0800

修复了mmio buddy的bug (#189)

- * 修改buddy_query

commit c1396d277115b371d09ad6d39a1c419f9224ffd0

Author: Gou Ngai <sujintao@dragonos.org>

Date: Mon Mar 6 11:28:32 2023 +0800

(下页继续)

(续上页)

Rwlock 文档 (#186)

* Rwlock 文档

```
commit a7eb62a47a8d701b90a14f83cc9028cfed07c268
Author: houmkh <100781004+houmkh@users.noreply.github.com>
Date: Mon Mar 6 11:21:29 2023 +0800
```

修改mmio - buddy代码结构和函数名 (#184)

* 修改mmio-buddy结构和函数名

```
commit c2481452f81750ec02adec627ab2edbc93d9cd9c
Author: houmkh <100781004+houmkh@users.noreply.github.com>
Date: Sat Mar 4 18:36:55 2023 +0800
```

rust 重构mmio_buddy和mmio (#178)

* rust 重构mmio_buddy和mmio

* mmio-buddy文档

Co-authored-by: longjin <longjin@RinGoTek.cn>

```
commit f1284c35717a2f9f8cee7cefc835ba1d23a1161
Author: Gou Ngai <sujintao@dragonos.org>
Date: Sat Mar 4 17:47:17 2023 +0800
```

新增了rust实现的信号量 (#181)

* 新增了rust实现的信号量

Co-authored-by: longjin <longjin@RinGoTek.cn>

```
commit 83b9512c1c1e828900084adcafddebee6a23f16
Author: Gou Ngai <sujintao@dragonos.org>
Date: Sat Mar 4 16:54:42 2023 +0800
```

(下页继续)

(续上页)

新增了rust实现的信号量 (#183)

* 新增了rust实现的信号量

```
commit e532a536a0b244f4590e6eb7910084bd63049704
Author: login <longjin@ringotek.cn>
Date: Thu Mar 2 22:50:07 2023 +0800
```

添加赞助者: FengWangHao (#179)

```
commit b66beefd4e9ead61ee55f335246eb8277d3011
Author: login <longjin@ringotek.cn>
Date: Mon Feb 27 01:00:35 2023 +0800
```

解决ubuntu2210版本无法正确编译grub, 以及正确安装qemu的问题 (#176)

```
commit 4177d0327c3eacdc606f0b22f99f208fd48cfff3
Author: kong <45937622+kkkkkkong@users.noreply.github.com>
Date: Mon Feb 20 17:03:37 2023 +0800
```

RTQueue改用双向链表存储 (#174)

* RTQueue改用双向链表存储

```
commit 2bf5ee0e3cac3a91dee6a13b71c86a9477c07d9b
Author: login <longjin@ringotek.cn>
Date: Sat Feb 11 13:04:24 2023 +0800
```

修改libc的编译相关内容 (#171)

1. 将libc的include文件夹分为export和internal
2. 将所有app都直接链接libc.a, 而不是都执行一遍"搜索.o"的过程

```
commit 90b077f9d3ecd48ca46f8bbb32363620db6ddbe6
Author: kong <45937622+kkkkkkong@users.noreply.github.com>
Date: Thu Feb 9 15:24:37 2023 +0800
```

Sched rt doc (#163)

* update

* 完善调度器文档

(下页继续)

(续上页)

* 更新RT调度器文档

* 更新实时调度文档

commit 009f92d50fe2e52e425bce397801d3fa204daecd

Author: Satin Wuker <74630829+SatinWuker@users.noreply.github.com>

Date: Tue Feb 7 19:29:09 2023 -0800

fix typos 改正README_EN的错别字和语法错误 (#167)

19.2 V0.1.4

备注：本文作者：龙进 longjin@RinGoTek.cn

2023 年 2 月 4 日

19.2.1 贡献者名单

DragonOS V0.1.4 版本由以下小伙伴贡献代码：

- 龙进 longjin@RinGoTek.cn
- Gou Ngai sujintao@DragonOS.org
- 孔维超 kongweichao@DragonOS.org
- 侯嘉滢 houjiaying@DragonOS.org

19.2.2 赞助者名单

感谢以下同学的赞赏，我们将不断努力！

- David Wen (人民币 2000 元)
- Seele.Clover (人民币 500 元)
- 叶锦毅 (人民币 100 元)
- 林 (人民币 50 元)
- Albert (人民币 9.99 元)
- TerryLeeSCUT (人民币 6.66 元)
- slientbard (人民币 6.66 元)

- 悟 (人民币 2.00 元)
- 【其他匿名的热心人士】 (人民币 1.00 元)

19.2.3 更新内容-内核

- Spinlock: new: 新增具有守卫的自旋锁 SpinLock, 支持编译期对锁的使用进行检查。(#148)
- Spinlock: feature: Raw spin lock 增加 lock_irqsave、unlock_irqrestore(#151)
- Mutex: new: Rust 版本的 Mutex (#157)
- doc: new: Rust 代码风格文档 (#161)
- WaitQueue: new: Rust 版本的 WaitQueue (#162)
- WaitQueue: update: C 版本的 wait_queue 的唤醒, 改为立即唤醒 (#158)
- block io: new: Block IO 调度器. 当具有多核时, io 调度器在核心 1 上运行。(#158)
- smp: bugfix: 为 AP 核启动 apic_timer, 使其能够运行调度 (#158)
- smp: new: 增加 kick_cpu 功能, 支持让某个特定核心立即运行调度器 (#158)
- smp: new: 增加进程在核心间迁移的功能 (#158)
- scheduler: new: 增加实时进程调度器 (支持 FIFO、RR 策略) (#139)
- scheduler: update: CFS 调度器为每个核心设置单独的 IDLE 进程 pcb (pid 均为 0) (#158)
- scheduler: bugfix: process_wakeup 时, 对 cfs 的进程, 重设虚拟运行时间。解决由于休眠的进程, 其虚拟运行时间过小, 导致其他进程饥饿的问题 (#158)
- process: new: pcb 中增加 migrate_to 字段 (#158)

19.2.4 更新内容-用户环境

无

19.2.5 更新内容-其他

无

19.2.6 更新内容-软件移植

无

19.2.7 源码、发布版镜像下载

您可以通过以下方式获得源代码:

通过 Git 获取

- 您可以访问<https://github.com/fslongjin/DragonOS/releases>下载发布版的代码，以及编译好的，可运行的磁盘镜像。
- 我们在 gitee 上也有镜像仓库可供下载：<https://gitee.com/DragonOS/DragonOS>

通过 DragonOS 软件镜像站获取

为解决国内访问 GitHub 慢、不稳定的问题，同时为了方便开发者们下载 DragonOS 的每个版本的代码，我们特意搭建了镜像站，您可以通过以下地址访问镜像站：

您可以通过镜像站获取到 DragonOS 的代码压缩包，以及编译好的可运行的磁盘镜像。

- <https://mirrors.DragonOS.org>
- <https://mirrors.DragonOS.org.cn>
- 国内镜像加速：[\[https://mirrors.ringotek.cn/\]](https://mirrors.ringotek.cn/) (<https://mirrors.ringotek.cn/>)

19.2.8 开放源代码声明

备注：为促进 DragonOS 项目的健康发展，DragonOS 以 GPLv2 开源协议进行发布。所有能获得到 DragonOS 源代码以及相应的软件制品（包括但不限于二进制副本、文档）的人，都能享有我们通过 GPLv2 协议授予您的权利，同时您也需要遵守协议中规定的义务。

这是一个相当严格的，保护开源软件健康发展，不被侵占的协议。

对于大部分的善意的人们而言，您不会违反我们的开源协议。

我们鼓励 DragonOS 的自由传播、推广，但是请确保所有行为没有侵犯他人的合法权益，也没有违反 GPLv2 协议。

请特别注意，对于违反开源协议的，尤其是商业闭源使用以及任何剽窃、学术不端行为将会受到严肃的追责。（这是最容易违反我们的开源协议的场景）。

并且, 请注意, 按照 GPLv2 协议的要求, 基于 DragonOS 修改或二次开发的软件, 必须同样采用 GPLv2 协议开源, 并标明其基于 DragonOS 进行了修改。亦需保证这些修改版本的用户能方便的获取到 DragonOS 的原始版本。

您必须使得 DragonOS 的开发者们, 能够以同样的方式, 从公开渠道获取到您二次开发的版本的源代码, 否则您将违反 GPLv2 协议。

关于协议详细内容, 还敬请您阅读项目根目录下的 **LICENSE** 文件。请注意, 按照 GPLv2 协议的要求, **只有英文原版才具有法律效力**。任何翻译版本都仅供参考。

开源软件使用情况

DragonOS 在开发的过程中, 参考了一些开源项目的设计, 或者引入了他们的部分代码, 亦或是受到了他们的启发。现将他们列在下面。我们对这些开源项目的贡献者们致以最衷心的感谢!

格式: < 项目名 > - < 链接 > - < 开源协议 >

- Linux - <https://git.kernel.org/> - GPLv2
- skiftOS - <https://github.com/skiftOS/skift> - MIT
- FYSOS - <https://github.com/fysnet/FYSOS> - [FYSOS' License](#)
- LemonOS - <https://github.com/LemonOSProject/LemonOS.git> - BSD 2-Clause License
- LZ4 - <https://github.com/lz4/lz4> - BSD 2-Clause license
- SerenityOS - <https://github.com/SerenityOS/serenity.git> - BSD 2-Clause license
- MINE - 《一个 64 位操作系统的设计与实现》田宇; 人民邮电出版社
- chcore - 《现代操作系统: 设计与实现》陈海波, 夏虞斌; 机械工业出版社
- SimpleKernel - <https://github.com/Simple-XX/SimpleKernel> - MIT
- rcore-fs - <https://github.com/rcore-os/rcore-fs.git> - MIT

19.2.9 当前版本的所有提交记录

```
commit f6ba114bb0420e848ef7fc844c96c0d7a0552d93
Author: houmkh <100781004+houmkh@users.noreply.github.com>
Date: Sat Feb 4 12:31:15 2023 +0800
```

```
Block IO Scheduler (#158)
```

```
* Block io调度器
```

```
* process_
```

→wakeup时, 对cfs的进程, 重设虚拟运行时间。解决由于休眠的进程, 其虚拟运行时间过小, 导致其他进程饥饿的

(下页继续)

(续上页)

- * 1、为AP核启动apic_timer,使其能够运行调度
- 2、增加kick_cpu功能,支持让某个特定核心立即运行调度器
- 3、wait_queue的唤醒,改为立即唤醒。
- 4、增加进程在核心间迁移的功能
- 5、CFS调度器为每个核心设置单独的IDLE进程pcb (pid均为0)
- 6、pcb中增加migrate_to字段
- 7、当具有多核时,io调度器在核心1上运行。

* io调度器文件位置修改

* 修改io的makefile

* 更新makefile中的变量名

* 修改io调度器函数名

Co-authored-by: login <longjin@ringotek.cn>

commit 151251b50b7ed55596edd32ffec49a4041010e2a

Author: login <longjin@ringotek.cn>

Date: Tue Jan 31 19:27:02 2023 +0800

Patch add rust waitqueue (#162)

* new: rust版本的waitqueue

* new:等待队列的文档

commit 3c369b1430e8d571bcc74a8ef7fefc1c4cae5dd2

Author: login <longjin@ringotek.cn>

Date: Mon Jan 30 15:43:42 2023 +0800

new:新增rust代码风格 (#161)

commit c28bd540ac856cd9d8d5597852af8f2588a660e4

Author: login <longjin@ringotek.cn>

Date: Mon Jan 30 15:10:24 2023 +0800

更新赞助者名单 (#160)

(下页继续)

(续上页)

* 更新赞赏者列表

commit 935f40ec174fec217aed4553d45996327443bc0e

Author: login <longjin@ringotek.cn>

Date: Tue Jan 17 21:30:16 2023 +0800

new: Rust 版本的 Mutex (#157)

commit d8a064128a8a06b90ff4c7b87c193518d9572641

Author: Gou Ngai <94795048+AlbertSanoe@users.noreply.github.com>

Date: Mon Jan 16 19:58:50 2023 +0800

Raw spin lock 增加lock_irqsave、unlock_irqrestore(#151)

Raw spin lock 增加lock_irqsave、unlock_irqrestore

commit 06b09f34ed64a006a80ae8df383e3c8b176f02e0

Author: kong <45937622+kkkkkkong@users.noreply.github.com>

Date: Sat Jan 14 22:38:05 2023 +0800

Patch sched rust (#139)

* update

* 添加rt调度器的rust初步实现

* 完善rt调度逻辑

* 调试rt调度器

* 修改sched的返回值

* cargo fmt 格式化

* 删除无用代码，修补rt bug

* 删除无用的代码，和重复的逻辑

* 软中断bugfix

* 删除一些代码

* 添加kthread_run_rt文档

(下页继续)

(续上页)

* 解决sphinx警告_static目录不存在的问题

Co-authored-by: longjin <longjin@RinGoTek.cn>

commit ec53d23ed03347854189d92b7e175f309779321b

Author: login <longjin@ringotek.cn>

Date: Sat Jan 14 10:35:49 2023 +0800

new: 新增具有守卫的自旋锁SpinLock, 支持编译期对锁的使用进行检查。 (#148)

commit 41474ba3df99b6822ce452dc94dc53a4da62cba1

Author: login <longjin@ringotek.cn>

Date: Tue Jan 10 22:07:41 2023 +0800

更新Readme中关于DragonOS的介绍部分 (#146)

commit 8ad2e358fd3b05eed2919de50640682e51687fb5

Author: login <longjin@ringotek.cn>

Date: Sun Jan 8 15:51:59 2023 +0800

更新about app中的版本号 (#145)

* 更新about app中的版本号

commit a8b621c8d1fe77251b8e4eafe258dc0ee7366dd5

Author: login <longjin@ringotek.cn>

Date: Sun Jan 8 15:47:44 2023 +0800

修正由于libc中具有crti.S和crtn.S, 造成的与x86_64-elf-gcc不兼容的问题 (#144)

commit 9358ff0f6f7daa18d6fab4497de025736b3d6725

Author: login <longjin@ringotek.cn>

Date: Sun Jan 8 15:06:52 2023 +0800

Add v0.1.3 changelog (#143)

* new: 0.1.3发行日志

* 新增输出指定时间范围内的贡献者名单的脚本

* 更新bootloader文档

(下页继续)

(续上页)

- * update: 简介文档
- * new: 镜像站文档
- * update: 功能特性文档

19.3 V0.1.3

备注: 本文作者: 龙进 longjin@RinGoTek.cn

2023 年 1 月 8 日

19.3.1 贡献者名单

DragonOS V0.1.3 版本由以下小伙伴贡献代码:

- 龙进 longjin@RinGoTek.cn
- 吴宇健 wuyujian@DragonOS.org
- 关锦权 guanjinquan@DragonOS.org
- Gou Ngai sujintao@DragonOS.org

19.3.2 赞助者名单

感谢以下同学的赞赏, 我们将不断努力!

- David Wen
- [Seele.Clover](#)
- TerryLeeSCUT
- 悟
- slientbard

19.3.3 更新内容-内核

- syscall: new: 增加 getpid 系统调用 (#120)
- signal: update: 对于除了 SIGKILL 以外的信号，也将他们加入 SigQueue (#120)
- rtc: refactor: 使用 Rust 重构 RTC 驱动 (#118)
- doc: new: 新增 signal 的文档 (#126)
- Spinlock: new: 新增 rust 写的 RawSpinlock (#127)
- arch: update: 在 lib.rs 中，将 arch 模块的路径进行更改，使得其他模块使用 arch 的代码时，不需要指定 arch::x86_64 (#128)
- mm: bugfix: 修复页面分配器在初始化时，ZONE_NORMAL_INDEX 始终为 0 的 bug (#129)
- scheduler: new: 使用 Rust 重构 CFS 调度器 (#131)
- smp: 删除已经在 smp 中废弃的 HPET 中断转发函数 (#131)
- process: bugfix: 修复 init 进程忘记设定 fs gs 寄存器的问题。 (#132)
- vfs: update: 将 VFS 文件夹重命名为 vfs (#133)
- lockref: new: 新增 rust 版本的 lockref (#135)
- cpu: new: new:Rust 封装 cpu_relax(), 通过 pause 指令，让 cpu 休息一会儿, 降低空转功耗. (#135)
- 使用 rust 重构 softirq 机制 (#138)

19.3.4 更新内容-用户环境

- libc: bugfix: 注册信号处理函数时，总是注册 sigkill 的问题 (#120)
- libc: new: 增加了 raise、kill、abort (#120)
- libc: new: 新增 arch 文件夹，在下面新增 crt0 crti crtn 文件 (#134)
- libc: new: 新增 fflush(), fprintf(), stdin, stdout, stderr, ferror(), fopen(), fclose(), putchar(), puts() (#136)
- libc: new: 简单添加了 fopen() 对 mode 参数的处理。请注意，它没有完全遵循 posix，也与 Linux 的不一致，将来使用 Rust 的时候完善它。 (#141)
- 移植: new: 新增了 gmp, mpfr, mpc 的移植构建脚本 (#136)
- 移植: new: 新增了 gcc、binutils 的交叉编译构建脚本以及 gcc-11.3.0, binutils-2.38 的补丁（在 DragonOS-community 下的仓库中） (#136)
- compile: update: 更改编译器的 Include 路径，使得 include 时不需要加 <libc/src/include/> 前缀 (#124)

19.3.5 更新内容-其他

- bugfix: 修复 docker 安装时异常退出的 bug (#116)
- new: 新增目标为 x86_64-elf 的 GCC 裸机编译器, 并使用它来编译 DragonOS (#111)
- update: 更新 Docker 编译镜像至版本 dragonos/dragonos-dev:v1.2, 并支持从 Dockerfile 构建这个编译镜像 (#111)
- bugfix: 修复 MBR 磁盘镜像未设置启动标志的 bug (#111)
- update: 更新 github workflow, 增加 cache, 加快 build check 的速度
- bugfix: 修复下载 grub2.06 时的提示错误 (#125)

19.3.6 更新内容-软件移植

- new: gcc 11.3.0 userland cross-compiler, 提交: 64a5b1cbf28e3305560e166c1b6624e99745c720, 仓库: <https://github.com/DragonOS-Community/gcc>
- new: binutils 2.38 cross-compile tools, 提交: a0ae560e0065862a9867b9e1f8364749ef38d99e, 仓库: <https://github.com/DragonOS-Community/binutils>
- new: gmp 6.2.1, 提交: dd9eee5778fb6027fafa4fe850aff21b1a71c18e, 仓库: <https://github.com/DragonOS-Community/gmp-6.2.1>
- new: mpfr 4.1.1, 提交: fa8e30cdc2e838fdd82b60fec31fcfc5e118aad6, 仓库: <https://github.com/DragonOS-Community/mpfr>
- new: mpc 1.2.1, (无需打补丁即可移植), 仓库: <https://github.com/DragonOS-Community/mpc>

19.3.7 源码、发布版镜像下载

您可以通过以下方式获得源代码:

通过 Git 获取

- 您可以访问<https://github.com/fslongjin/DragonOS/releases>下载发布版的代码, 以及编译好的, 可运行的磁盘镜像。
- 我们在 gitee 上也有镜像仓库可供下载: <https://gitee.com/DragonOS/DragonOS>

通过 DragonOS 软件镜像站获取

为解决国内访问 GitHub 慢、不稳定的问题，同时为了方便开发者们下载 DragonOS 的每个版本的代码，我们特意搭建了镜像站，您可以通过以下地址访问镜像站：

您可以通过镜像站获取到 DragonOS 的代码压缩包，以及编译好的可运行的磁盘镜像。

- <https://mirrors.DragonOS.org>
- <https://mirrors.DragonOS.org.cn>
- 国内镜像加速: [<https://mirrors.ringotek.cn/>] (<https://mirrors.ringotek.cn/>)

19.3.8 开放源代码声明

备注：为促进 DragonOS 项目的健康发展，DragonOS 以 GPLv2 开源协议进行发布。所有能获得到 DragonOS 源代码以及相应的软件制品（包括但不限于二进制副本、文档）的人，都能享有我们通过 GPLv2 协议授予您的权利，同时您也需要遵守协议中规定的义务。

这是一个相当严格的，保护开源软件健康发展，不被侵占的协议。

对于大部分的善意的人们而言，您不会违反我们的开源协议。

我们鼓励 DragonOS 的自由传播、推广，但是请确保所有行为没有侵犯他人的合法权益，也没有违反 GPLv2 协议。

请特别注意，对于违反开源协议的，尤其是**商业闭源使用以及任何剽窃、学术不端行为将会受到严肃的追责**。（这是最容易违反我们的开源协议的场景）。

并且，请注意，按照 GPLv2 协议的要求，基于 DragonOS 修改或二次开发的软件，必须同样采用 GPLv2 协议开源，并标明其基于 DragonOS 进行了修改。亦需保证这些修改版本的用户能方便的获取到 DragonOS 的原始版本。

您必须使得 DragonOS 的开发者们，能够以同样的方式，从公开渠道获取到您二次开发的版本的源代码，否则您将违反 GPLv2 协议。

关于协议详细内容，还请您阅读项目根目录下的 **LICENSE** 文件。请注意，按照 GPLv2 协议的要求，**只有英文原版才具有法律效力**。任何翻译版本都仅供参考。

开源软件使用情况

DragonOS 在开发的过程中，参考了一些开源项目的设计，或者引入了他们的部分代码，亦或是受到了他们的启发。现将他们列在下面。我们对这些开源项目的贡献者们致以最衷心的感谢！

格式：< 项目名 > - < 链接 > - < 开源协议 >

- Linux - <https://git.kernel.org/> - GPLv2
- skiftOS - <https://github.com/skiftOS/skift> - MIT
- FYSOS - <https://github.com/fysnet/FYSOS> - [FYSOS' License](#)
- LemonOS - <https://github.com/LemonOSProject/LemonOS.git> - BSD 2-Clause License
- LZ4 - <https://github.com/lz4/lz4> - BSD 2-Clause license
- SerenityOS - <https://github.com/SerenityOS/serenity.git> - BSD 2-Clause license
- MINE - 《一个 64 位操作系统的设计与实现》田宇; 人民邮电出版社
- chcore - 《现代操作系统：设计与实现》陈海波, 夏虞斌; 机械工业出版社
- SimpleKernel - <https://github.com/Simple-XX/SimpleKernel> - MIT
- rcore-fs - <https://github.com/rcore-os/rcore-fs.git> - MIT

19.3.9 当前版本的所有提交记录

```
commit a8b621c8d1fe77251b8e4eafe258dc0ee7366dd5
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sun Jan 8 15:47:44 2023 +0800
```

修正由于libc中具有crti.S和crti.S, 造成的与x86_64-elf-gcc不兼容的问题 (#144)

```
commit 9358ff0f6f7daa18d6fab4497de025736b3d6725
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sun Jan 8 15:06:52 2023 +0800
```

Add v0.1.3 changelog (#143)

* new: 0.1.3发行日志

* 新增输出指定时间范围内的贡献者名单的脚本

* 更新bootloader文档

* update: 简介文档

(下页继续)

(续上页)

- * new: 镜像站文档

- * update: 功能特性文档

```
commit fd91905f022b3ceaa59e666d1ff42d91fb8d40ef
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sun Jan 8 11:38:59 2023 +0800
```

解决编译gcc、binutils的脚本中，变量名称错误的问题 (#142)

```
commit 62e4613978193aaf5d949a331df0398f2d085a30
```

```
Author: Gou Ngai <94795048+AlbertSanoe@users.noreply.github.com>
```

```
Date: Sat Jan 7 23:15:37 2023 +0800
```

使用rust重构softirq机制；解决Rtc驱动的编译警告问题 (#138)

- * 使用rust重构softirq机制

- * 解决Rtc驱动的编译警告问题

```
Co-authored-by: longjin <longjin@RinGoTek.cn>
```

```
commit e9fdc57bf878f1bc5cc5743dfaceaef743439291
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sat Jan 7 22:36:49 2023 +0800
```

└─

→简单添加了fopen对mode参数的处理。请注意，它没有完全遵循posix，也与Linux的不一致，将来使用Rust的时候
→ (#141)

```
commit 2224c93ea968bc74621f7e124b4aca04875b3e6a
```

```
Author: guanjinquan <1666320330@qq.com>
```

```
Date: Fri Jan 6 21:29:23 2023 +0800
```

完善libc，构建了OS-specific工具链，编译了基于gcc-11.3.0的DragonOS userland
→compiler，移植了mpfr,gmp,mpc库 (#134)

- * 修改include路径

- * 添加了创建libsysapi.a和/bin/sysroot/usr/include/+lib/的代码

- * 修补.gitignore

- * 删除多余项

(下页继续)

(续上页)

- * 优化脚本可读性
- * 新增 `crt0 crt1 crtn`
- * 编译 `binutils` 所需的东西
- * `fflush()` 和 `fprintf()` 的简单实现
- * 应用程序启动前，调用初始化 `libc` 的函数
- * 自动创建 `sysroot`
- * 添加了 `stderr` 的初始化
- * 修改了 `stderr` 的初始化
- * 内核添加对 `stdio` 的简略处理
- * 格式化代码
- * 修正打开 `stdio` 文件描述符的问题
- * `bugfix`: 修复 `fprintf` 忘记释放 `buf` 的问题
- * 修复 `shell` 错误地把入口设置为 `main` 而不是 `_start` 的问题
- * 新增 `__cxa_atexit` (gcc 要求 `libc` 提供这个)
- * 增加 `putchar puts`
- * 更新写入磁盘镜像的脚本，默认无参数时，使用 `legacy` 方式安装
- * 更新编译脚本
- * `stdio` 增加 `eof` 的定义
- * 新增 `extern cplusplus`
- * `mpfr gmp mpc` 构建脚本
- * 更新 `libsysapi.a` 为 `libc.a`

(下页继续)

(续上页)

- * 加上ferror fopen fclose

- * 更新移植的软件的构建脚本

- * 更改build_gcc_toolchain.sh中的-save参数名为-save-cache

Co-authored-by: longjin <longjin@RinGoTek.cn>

commit 61de2cdc3f29cdc6c441f128119e01e003e6f3ca

Author: login <longjin@ringotek.cn>

Date: Tue Jan 3 23:09:25 2023 +0800

新增rust版本的lockref (#135)

- * new: Rust封装cpu_relax(), 通过pause指令, 让cpu休息一会儿。降低空转功耗

- * new: Rust版本的lockref

- * Rust的RawSpinlock新增is_locked() 和set_value() 方法。

- * lockref文档

commit 2726f101b4cc787bbd36a69afffb0112f3a6567f

Author: login <longjin@ringotek.cn>

Date: Tue Jan 3 21:01:56 2023 +0800

删除无用的cfs.h (#136)

commit 587086d3f299f7394559d547c828191be20cfc11

Author: login <longjin@ringotek.cn>

Date: Sun Jan 1 16:53:57 2023 +0800

1、在文件系统目录下增加mod.rs 2、将VFS的路径改为vfs (#133)

2、将VFS的路径改为vfs

commit 843e442971a47693f37a5f8d3452c383f7325359

Author: login <longjin@ringotek.cn>

Date: Sat Dec 31 18:43:05 2022 +0800

修复init进程忘记设定fs gs寄存器的问题。 (#132)

commit 74bde36e014ff501241bf40dd83653db47a2c8e4

(下页继续)

(续上页)

Author: guanjinquan <1666320330@qq.com>

Date: Sat Dec 31 17:35:39 2022 +0800

Patch porting gcc v2 (#124)

* 更改编译器的Include路径, 使得include时不需要加`<libc/src/include/>`前缀

* 修改include路径

Co-authored-by: longjin <longjin@RinGoTek.cn>

commit d4f3de93a23e4bd4f000a3663768d47d094bf188

Author: login <longjin@ringotek.cn>

Date: Sat Dec 31 17:26:12 2022 +0800

使用Rust重构CFS调度器 (#131)

* 新建调度器的文件

* 把softirq vector移动到c文件中 (原来在.h)

* 将进程切换方式改为“中断返回时切换”

* new:使用rust重构CFS

* 删除已经在smp中废弃的HPET中断转发函数

* 代码格式化

* 删除多余的dunce依赖

commit 156949680c83f2d7e3b21ed68b11698b88eaf396

Author: login <longjin@ringotek.cn>

Date: Sat Dec 31 13:47:49 2022 +0800

└─

↪bugfix:修复当使用sched()运行调度器, 在切换进程的时候, 由于不在中断上下文内, 导致当前进程的上下文丢失 (#130)

└─

↪bugfix:修复当使用sched()运行调度器, 在切换进程的时候, 由于不在中断上下文内, 导致当前进程的上下文丢失

└─

↪bugfix:修复切换进程的宏的汇编代码的损坏部分, 未声明rax寄存器, 从而导致的编译器未定义行为问题。

(下页继续)

(续上页)

```
commit 882f0b7e7498dbff8de527b2b9159b7f6e6359c9
Author: YJwu2023 <119829947+YJwu2023@users.noreply.github.com>
Date:   Wed Dec 28 19:35:17 2022 +0800
```

修复内存bug与grub安装脚本的错误 (#129)

- * 修复内存bug与grub安装脚本的错误

- * 修改小bug

```
commit adc1846b06fb862caed049f435fc0061488a6ff9
Author: login <longjin@ringotek.cn>
Date:   Mon Dec 26 13:13:12 2022 +0800
```

内核：在lib.

→rs中，将arch模块的路径进行更改，使得其他模块使用arch的代码时，不需要指定arch::x86_64 (#128)

```
commit ac643d420b22f9d454ecefcccd51ed34a9664586b
Author: login <longjin@ringotek.cn>
Date:   Sun Dec 25 23:53:35 2022 +0800
```

new:新增rust写的RawSpinlock (#127)

```
commit 998390210549b47e6bdcc3fdab49eff4086ad18b
Author: login <longjin@ringotek.cn>
Date:   Sat Dec 24 23:30:26 2022 +0800
```

新增signal文档 (#126)

- * 新增signal文档

```
commit a7f5ca7b67160557abf84a1169dd60093220aeb0
Author: YJwu2023 <119829947+YJwu2023@users.noreply.github.com>
Date:   Sat Dec 24 23:29:36 2022 +0800
```

修复下载grub2.06时的提示错误 (#125)

- * 修复grub下载显示提示显示错误

```
commit 82762007da41148e1ed1df465211eb5c8ba2c15e
Author: login <longjin@ringotek.cn>
```

(下页继续)

(续上页)

Date: Fri Dec 23 18:11:47 2022 +0800

Update makefile.yml

commit b975025ec8854ca232152f4ee44cc2226891a34c

Author: login <longjin@ringotek.cn>

Date: Fri Dec 23 11:45:19 2022 +0800

Update makefile.yml

commit ad2bb74d949bfc2935e43ac7b261d7ecce23389

Author: login <longjin@ringotek.cn>

Date: Fri Dec 23 11:21:22 2022 +0800

Update makefile.yml

commit 6b7776d189ab5f19fbab20d6c5c9ed3ab20c7ab6

Author: login <longjin@ringotek.cn>

Date: Fri Dec 23 10:59:15 2022 +0800

修正 smp 的 makefile 中没有替换 AS 的问题

commit beb12a188b6c6bc4196796ac2ae1ecd7d8ed8223

Author: login <longjin@ringotek.cn>

Date: Fri Dec 23 10:57:39 2022 +0800

Update makefile.yml

commit d65c527730e5c8a75f6dad0f996c093040699ee3

Author: login <longjin@ringotek.cn>

Date: Thu Dec 22 22:58:28 2022 +0800

Update makefile.yml (#121)

commit 5ed4cd460200cb19aae8c3c67dfd77e1e9f0e105

Author: guanjinquan <75822481+guanjinquan@users.noreply.github.com>

Date: Thu Dec 22 21:09:12 2022 +0800

Patch gcc toolchain (#111)

* 添加了 GCC_cross_compile——tool_chain

* - 解决环境变量路径拼接时, 多了 `/` 的问题

(下页继续)

(续上页)

- apt安装时增加-y,不需用户确认
- * 解决添加环境变量的命令有误的问题
- * 修正编译错误时,还会执行下一步的问题
- * new: 编译完成后清理临时文件
- * 更新makefile
- * 调整: 把grub安装在 \$HOME/opt/dragonos-grub下
- * new: 新增dockerfile
- * 将镜像源换成中科大的(原因是清华的总是ban掉用于构建镜像的服务器的ip)
- * 修改为基于debian bullseye构建
- * 取消指定版本
- * 修复MBR磁盘镜像未设置启动标志的bug
- * 取消在docker中安装grub
- * 安装grub的过程改到客户机上进行
- * bootstrap.sh 添加--no-docker
- * 使用新版的docker编译镜像
- * 修补, 添加了一些关于gcc的检查

Co-authored-by: longjin <longjin@RinGoTek.cn>

commit ba0d93d8b26034abc54bcaf3f0ff04863bbd076e

Author: Gou Ngai <94795048+AlbertSanoe@users.noreply.github.com>

Date: Mon Dec 19 15:04:37 2022 +0800

refactor rtc module in rust (#118)

- * 用rust重构rtc模块

- * refactor the rtc module by rust

(下页继续)

- * rtc-updated

- * rtc-updated-4

- * rtc

commit c588d6f77f4b38939701b946228218ea81a7c8dc

Author: login <longjin@ringotek.cn>

Date: Mon Dec 19 15:03:44 2022 +0800

Patch add abort func (#120)

- * 对于除了sigkill以外的信号，也加入队列

- * bugfix:libc中，注册信号处理函数时，总是注册sigkill的问题

- * 增加getpid系统调用

- * 增加了raise、kill、abort

commit 47f0d12a1f1a1aa11be8e751ecdbf76f0cb596d9

Author: YJwu2023 <119829947+YJwu2023@users.noreply.github.com>

Date: Mon Dec 19 14:53:51 2022 +0800

修复docker安装时异常退出的bug (#119)

- * 修复docker安装时异常退出的bug

- * 修复grub编译脚本的小bug

commit 978043e47d1143ca2d5cf22b20793f032e8eb5a5

Author: login <longjin@ringotek.cn>

Date: Sun Dec 18 15:09:15 2022 +0800

修复当系统中不存在dosfstools时，无法正确格式化磁盘镜像的问题 (#117)

- * 修复当系统中不存在dosfstools时，无法正确格式化磁盘镜像的问题

commit f9127772dc372a2e607388fdd6818d3f9c4c6d28

Author: YJwu2023 <119829947+YJwu2023@users.noreply.github.com>

Date: Sat Dec 17 23:43:23 2022 +0800

修复docker安装时异常退出的bug (#116)

19.4 V0.1.2

备注：本文作者：龙进 longjin@RinGoTek.cn

2022 年 12 月 17 日

19.4.1 贡献者名单

DragonOS V0.1.2 版本由以下小伙伴贡献代码：

- 龙进 longjin@ringotek.cn
- 吴宇健 wuyujian@DragonOS.org
- Gou Ngai sujintao@DragonOS.org
- 黄厅 huangting@DragonOS.org
- 王文聪 1297389017@qq.com

19.4.2 赞助者名单

感谢以下同学的赞赏，我们将不断努力！

- David Wen
- Seele.Clover
- TerryLeeSCUT
- 悟
- slientbard

其中，非常感谢 **Seele.Clover** 给予 DragonOS 项目人民币 500 元的赞助与支持！我们对于每一笔赞助款项，将仔细登记，并确保其能被妥善的使用。

19.4.3 更新内容-内核

- 删除 rust_helloworld 文件 (#113)
- Signal: 允许用户注册信号处理函数，能够进入自定义的 handler。(#112)
 - 支持 kill 命令
 - 允许用户自定义信号处理函数
 - 新增 2 个系统调用：SYS_SIGACTION, SYS_RT_SIGRETURN

- libc 增加 `signal()`, `sigaction()` 函数。
- 暂时只支持旧版的 `sighandler`, 即: 只有 1 个参数的 `void handler(int signum)` 类型的信号处理函数。对于另一种信号处理函数 `void handler(int signum, siginfo_t *info, void* data)`, 尚不支持传递第三个参数。
- 在内核代码中加入自定义的 `stdint.h` 文件 (#109)
- 调整编译 `grub` 的脚本的部分 (#108)
- 新增 32、64 位 `uefi` 启动 (#105)(#101)
- 使用编译安装的 `grub-2.06`, 解决客户机上 `grub` 版本不对导致的编译无法运行的问题。
- 增加了 `timekeeping` 模块 (#106)
- `bugfix`: 修复 `rtc` 时钟对 `BCD` 码进行转换的时候, 忘了处理 `day` 字段的问题 (#104)
- `new`: 开发过程文档 (完成了一半)
- `bootstrap.sh` 解决下载 `rust` 慢的问题
- 更新“构建系统”文档
- `procfs->status` 增加显示 `preempt` 和虚拟运行时间 (#100)
- `ffz` 函数: 获取 `u64` 中的第一个值为 0 的 `bit` (#100)
- 解决由于编译器优化导致 `local_irq_restore` 无法获取到正确的 `rflags` 的值的值的问题
- 使用 `Rust` 重构串口驱动 (#99)

19.4.4 更新内容-用户环境

- `about app`: 显示当前构建的 `git commit sha1` 以及构建时间 (#114)
- `shell`: 修复 `shell` 的 `exec` 命令对绝对路径的拼接错误问题 (#114)
- `shell`: `exec` 命令增加“&”后台运行选项 (#100)
- `new`: 测试 `signal` 用的 `app`
- 将 `libc` 目录进行调整, 加入 `cargo` 作为 `rust` 的包管理器

19.4.5 源码、发布版镜像下载

您可以通过以下方式获得源代码:

通过 Git 获取

- 您可以访问<https://github.com/fslongjin/DragonOS/releases>下载发布版的代码，以及编译好的，可运行的磁盘镜像。
- 我们在 gitee 上也有镜像仓库可供下载：<https://gitee.com/DragonOS/DragonOS>

通过 DragonOS 软件镜像站获取

为解决国内访问 GitHub 慢、不稳定的问题，同时为了方便开发者们下载 DragonOS 的每个版本的代码，我们特意搭建了镜像站，您可以通过以下地址访问镜像站：

您可以通过镜像站获取到 DragonOS 的代码压缩包，以及编译好的可运行的磁盘镜像。

- <https://mirrors.DragonOS.org>
- <https://mirrors.DragonOS.org.cn>

19.4.6 开放源代码声明

备注：为促进 DragonOS 项目的健康发展，DragonOS 以 GPLv2 开源协议进行发布。所有能获得到 DragonOS 源代码以及相应的软件制品（包括但不限于二进制副本、文档）的人，都能享有我们通过 GPLv2 协议授予您的权利，同时您也需要遵守协议中规定的义务。

这是一个相当严格的，保护开源软件健康发展，不被侵占的协议。

对于大部分的善意的人们而言，您不会违反我们的开源协议。

我们鼓励 DragonOS 的自由传播、推广，但是请确保所有行为没有侵犯他人的合法权益，也没有违反 GPLv2 协议。

请特别注意，对于违反开源协议的，尤其是**商业闭源使用以及任何剽窃、学术不端行为将会受到严肃的追责**。（这是最容易违反我们的开源协议的场景）。

并且，请注意，按照 GPLv2 协议的要求，基于 DragonOS 修改或二次开发的软件，必须同样采用 GPLv2 协议开源，并标明其基于 DragonOS 进行了修改。亦需保证这些修改版本的用户能方便的获取到 DragonOS 的原始版本。

您必须使得 DragonOS 的开发者们，能够以同样的方式，从公开渠道获取到您二次开发的版本的源代码，否则您将违反 GPLv2 协议。

关于协议详细内容，还请您请阅读项目根目录下的 **LICENSE** 文件。请注意，按照 GPLv2 协议的要求，**只有英文原版才具有法律效力**。任何翻译版本都仅供参考。

开源软件使用情况

DragonOS 在开发的过程中, 参考了一些开源项目的设计, 或者引入了他们的部分代码, 亦或是受到了他们的启发。现将他们列在下面。我们对这些开源项目的贡献者们致以最衷心的感谢!

格式: < 项目名 > - < 链接 > - < 开源协议 >

- Linux - <https://git.kernel.org/> - GPLv2
- skiftOS - <https://github.com/skiftOS/skift> - MIT
- FYSOS - <https://github.com/fysnet/FYSOS> - [FYSOS' License](#)
- LemonOS - <https://github.com/LemonOSProject/LemonOS.git> - BSD 2-Clause License
- LZ4 - <https://github.com/lz4/lz4> - BSD 2-Clause license
- SerenityOS - <https://github.com/SerenityOS/serenity.git> - BSD 2-Clause license
- MINE - 《一个 64 位操作系统的设计与实现》田宇; 人民邮电出版社
- chcore - 《现代操作系统: 设计与实现》陈海波, 夏虞斌; 机械工业出版社
- SimpleKernel - <https://github.com/Simple-XX/SimpleKernel> - MIT

19.4.7 当前版本的所有提交记录

```
commit 7a818da88a1c7a1760de7671141b0ce1ca4e3dde
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sat Dec 17 17:49:12 2022 +0800
```

```
    Patch about auto gen version string (#114)
```

```
    * new: about app 中, 显示当前构建的 git commit sha1 以及构建时间
```

```
    * bugfix: 修复 shell 的 exec 命令对绝对路径的拼接错误问题
```

```
commit 83a7aaa46bbc411c43d4fc099c6c8884efbe4771
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sat Dec 17 16:31:50 2022 +0800
```

```
    删除 rust_helloworld 文件 (#113)
```

```
commit 6efd4740336205c9bfdd8b164e667cee2f38781e
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sat Dec 17 16:27:50 2022 +0800
```

```
    允许用户自定义信号处理函数 (#112)
```

(下页继续)

(续上页)

- * new: 用户注册信号处理函数, 能够进入自定义的handler

- * 修复忘了传信号的数字给用户的处理函数的bug

- * new:sigreturn

- * 删除注释

```
commit 0e0c187484281768391e131495f0655e40d70cf7
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Fri Dec 16 16:20:09 2022 +0800
```

在内核代码中加入自定义的stdint.h文件 (#109)

```
commit d02e6ea4112ad520aa4090ff73cdf592e14c0a82
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Wed Dec 14 20:01:55 2022 +0800
```

调整编译grub的脚本的部分 (#108)

- 1、bugfix: 修复编译grub的脚本的部分错误

- 2、将grub下载源替换为tuna

- 3、优化写入磁盘镜像的脚本

- 4、将bios文件夹改名为legacy

```
commit 38b341b8aa671f75ac26d05059aa2e9a09e653b7
```

```
Author: YJwu2023 <119829947+YJwu2023@users.noreply.github.com>
```

```
Date: Wed Dec 14 16:58:49 2022 +0800
```

新增32位uefi启动 (#105)

- * 新增32位uefi启动

- * 修复小bug

- * 增加grub本地编译安装

- * 增加本地grub编译安装脚本

- * 修正小错误

- * 修复空文件夹不上传的bug

(下页继续)

(续上页)

```
commit 01876902fbf6ed43992cc7d153bd8c505cb5224b
Author: Gou Ngai <94795048+AlbertSanoe@users.noreply.github.com>
Date:   Wed Dec 14 15:13:54 2022 +0800
```

增加了timekeeping模块 (#106)

- * 增加了timekeeping模块

- * 格式化文档和细节更改

Co-authored-by: longjin <longjin@RinGoTek.cn>

```
commit 728aca308917a7d4d0ba10fe8174e9408d77a9a6
Author: login <longjin@ringotek.cn>
Date:   Sun Dec 11 22:59:47 2022 +0800
```

bugfix: 修复rtc时钟对BCD码进行转换的时候, 忘了处理day字段的问题 (#104)

```
commit 237e95c6ddce72d72ae7fedfeca412fab82b3622
Author: wwc-15172310230 <78997674+wwc-15172310230@users.noreply.github.com>
Date:   Sun Dec 11 22:22:10 2022 +0800
```

调整user下libs的libc目录结构 (#103)

- * 调整user下libs的libc目录结构

- * 修正.gitignore文件的问题

- * 修复无法编译的问题

Co-authored-by: longjin <longjin@RinGoTek.cn>

```
commit 2291ffdece1dc5a703602f79f74df8a4854d215b
Author: login <longjin@ringotek.cn>
Date:   Sun Dec 11 20:09:58 2022 +0800
```

文档更新 (#102)

- * new: 开发过程文档 (完成了一半)

- * bootstrap.sh解决下载rust慢的问题

(下页继续)

(续上页)

* 更新“构建系统”文档

```
commit 7f439c5ddbd2ecffc112149d16983975f523052c
Author: YJwu2023 <119829947+YJwu2023@users.noreply.github.com>
Date: Fri Dec 9 16:08:54 2022 +0800
```

增加uefi启动 (#101)

* 增加uefi启动

* 修改脚本

* uefi修改

* 删除错误的注释

* 修正写入磁盘镜像的脚本

* 修改X86_64为x86_64

Co-authored-by: longjin <longjin@RinGoTek.cn>

```
commit 1a2eaa402f05f82aaeebe1e03824534a0a425d4d
Author: login <longjin@ringotek.cn>
Date: Thu Dec 8 22:59:51 2022 +0800
```

signal的处理(kill命令)以及一些其他的改进 (#100)

* 将entry.S中冗余的ret_from_syscall代码删除, 改为jmp Restore_all

* new: 增加判断pt_regs是否来自用户态的函数

* new: rust的cli和sti封装

* 将原有的判断pt_regs是否来自用户态的代码, 统一改为调用user_mode函数

* ffz函数: 获取u64中的第一个值为0的bit

* spinlock增加 spinlock irq spin_unlock_irq

* 临时解决显示刷新线程迟迟不运行的问题

* 更改ffi_convert的生命周期标签

(下页继续)

(续上页)

- * new: 测试signal用的app
- * 解决由于编译器优化导致local_irq_restore无法获取到正确的rflags的值的问题
- * new: exec命令增加"&"后台运行选项
- * procfs->status增加显示preempt和虚拟运行时间
- * 更改引用计数的FFIBind2Rust trait中的生命周期标签
- * new: signal处理(kill)
- * 更正在review中发现的一些细节问题

```
commit f8b55f6d3fcbf152a1cb6d6fc722bf1607418b28
Author: TingHuang <92705854+TingSHub@users.noreply.github.com>
Date: Tue Dec 6 22:15:03 2022 +0800
```

Patch uart (#99)

- * 添加UART驱动相关文件
- * 添加驱动核心文件, 将rust编写的驱动代码加入Package中
- * 添加glib.h文件生成rust代码, 添加uart驱动代码
- * 添加串口发送及接收相关代码
- * 添加字符串发送函数, 未实现具体功能
- * 为调用uart驱动的代码添加rust接口
- * 添加字符串发送函数, 修改C语言调用接口
- * 添加rust串口驱动
- * 添加uart.h头文件, 将串口端口类型改为enum
- * 添加注释, 规范代码

```
commit 036acc52ce9d0fb9e7d92768ff74939a29c07f32
Author: login <longjin@ringotek.cn>
```

(下页继续)

(续上页)

Date: Tue Nov 29 21:46:13 2022 +0800

将entry.S中冗余的ret_from_syscall代码删除, 改为jmp Restore_all (#98)

* 将entry.S中冗余的ret_from_syscall代码删除, 改为jmp Restore_all

19.5 V0.1.1

备注: 本文作者: 龙进 longjin@RinGoTek.cn

2022 年 11 月 27 日

19.5.1 贡献者名单

DragonOS V0.1.1 版本由以下小伙伴贡献代码:

- 龙进 longjin@RinGoTek.cn
- 周于[Ⓔ] zhouyuzhe@DragonOS.org

19.5.2 赞助者名单

感谢以下同学的赞赏, 我们将不断努力!

- David Wen
- TerryLeeSCUT
- 悟
- slientbard

其中, 非常感谢 **David Wen** 给予 DragonOS 项目人民币 1000 元的赞助与支持! 我们对于每一笔赞助款项, 将仔细登记, 并确保其能被妥善的使用。

19.5.3 更新内容-内核

- 新增 rust ffi (#77)
- port kmalloc and printk to rust
- rust 下的 kdebug kinfo kwarn kBUG kerror 宏
- bugfix: 修复进程 pcb 被回收时, 未将其从链表中删除的问题
- 目录结构优化: 移动 asm.h 和 cmpxchg.h
- signal 的发送
- procfs: 查看进程的状态
- 解决第一次编译时磁盘镜像权限错误的问题
- 将 fork 相关代码移动到 fork.c

19.5.4 更新内容-用户环境

- shell: 增加 kill 命令, 可向目标进程发送信号。但由于仍未完善 signal 机制, 因此目标进程暂时不能响应这个信号。

19.5.5 源码、发布版镜像下载

您可以通过以下方式获得源代码:

通过 Git 获取

- 您可以访问<https://github.com/fslongjin/DragonOS/releases>下载发布版的代码, 以及编译好的, 可运行的磁盘镜像。
- 我们在 gitee 上也有镜像仓库可供下载: <https://gitee.com/DragonOS/DragonOS>

通过 DragonOS 软件镜像站获取

为解决国内访问 GitHub 慢、不稳定的问题, 同时为了方便开发者们下载 DragonOS 的每个版本的代码, 我们特意搭建了镜像站, 您可以通过以下地址访问镜像站:

您可以通过镜像站获取到 DragonOS 的代码压缩包, 以及编译好的可运行的磁盘镜像。

- <https://mirrors.DragonOS.org>
- <https://mirrors.DragonOS.org.cn>

19.5.6 开放源代码声明

备注：为促进 DragonOS 项目的健康发展，DragonOS 以 GPLv2 开源协议进行发布。所有能获得到 DragonOS 源代码以及相应的软件制品（包括但不限于二进制副本、文档）的人，都能享有我们通过 GPLv2 协议授予您的权利，同时您也需要遵守协议中规定的义务。

这是一个相当严格的，保护开源软件健康发展，不被侵占的协议。

对于大部分的善意的人们而言，您不会违反我们的开源协议。

我们鼓励 DragonOS 的自由传播、推广，但是请确保所有行为没有侵犯他人的合法权益，也没有违反 GPLv2 协议。

请特别注意，对于违反开源协议的，尤其是**商业闭源使用以及任何剽窃、学术不端行为将会受到严肃的追责**。（这是最容易违反我们的开源协议的场景）。

并且，请注意，按照 GPLv2 协议的要求，基于 DragonOS 修改或二次开发的软件，必须同样采用 GPLv2 协议开源，并标明其基于 DragonOS 进行了修改。亦需保证这些修改版本的用户能方便的获取到 DragonOS 的原始版本。

您必须使得 DragonOS 的开发者们，能够以同样的方式，从公开渠道获取到您二次开发的版本的源代码，否则您将违反 GPLv2 协议。

关于协议详细内容，还敬请您请阅读项目根目录下的 **LICENSE** 文件。请注意，按照 GPLv2 协议的要求，**只有英文原版才具有法律效力**。任何翻译版本都仅供参考。

开源软件使用情况

DragonOS 在开发的过程中，参考了一些开源项目的设计，或者引入了他们的部分代码，亦或是受到了他们的启发。现将他们列在下面。我们对这些开源项目的贡献者们致以最衷心的感谢！

格式：< 项目名 > - < 链接 > - < 开源协议 >

- Linux - <https://git.kernel.org/> - GPLv2
- skiftOS - <https://github.com/skiftOS/skift> - MIT
- FYSOS - <https://github.com/fysnet/FYSOS> - [FYSOS' License](#)
- LemonOS - <https://github.com/LemonOSProject/LemonOS.git> - BSD 2-Clause License
- LZ4 - <https://github.com/lz4/lz4> - BSD 2-Clause license
- SerenityOS - <https://github.com/SerenityOS/serenity.git> - BSD 2-Clause license
- MINE - 《一个 64 位操作系统的设计与实现》田宇; 人民邮电出版社
- chcore - 《现代操作系统：设计与实现》陈海波，夏虞斌; 机械工业出版社
- SimpleKernel - <https://github.com/Simple-XX/SimpleKernel> - MIT

19.5.7 当前版本的所有提交记录

```
commit d65ade9c5909076747bd00966a398fe27fbd290d
Author: DaJiYuQia <88259094+DaJiYuQia@users.noreply.github.com>
Date: Sun Nov 27 14:21:31 2022 +0800
```

Patch procfs (#95)

- * debug color problem

Co-authored-by: longjin <longjin@RinGoTek.cn>

```
commit 6cb769c423b09e88fea1763210200a716477be0a
Author: login <longjin@ringotek.cn>
Date: Sun Nov 27 14:17:36 2022 +0800
```

将include目录下的rust代码转移到他们应当属于的模块中 (#96)

- * 将include目录下的rust代码转移到他们应当属于的模块下。

```
commit 27a97abd2474b03ad09b562e5ed11e1fdae8eb32
Author: DaJiYuQia <88259094+DaJiYuQia@users.noreply.github.com>
Date: Sat Nov 26 17:34:00 2022 +0800
```

Patch procfs (#90)

- * 1234

- * 123

- * 合并master

- * procfs

- * 1

- * procfs展示进程基本信息

- * modified code

- * 恢复权限

- * 恢复权限

(下页继续)

(续上页)

```

#恢复权限

* modify permission

* 删除run.sh

* 解决第一次编译时磁盘镜像权限错误的问题

* 恢复.vscode/c_cpp_properties.json

* 删除process.c中错误的do_fork

* remake procfs

* 修改一些变量名

* 修改类型

* modified

* data_puts缓冲区溢出后return

Co-authored-by: longjin <longjin@RinGoTek.cn>

commit ad23fcddf893d7f92d2bf3efdb66e969416d2852
Author: login <longjin@ringotek.cn>
Date:   Wed Nov 23 21:34:35 2022 +0800

    bugfix: 修复进程退出时未释放signal和sighand && 增加赞赏者名单: David Wen (#93)

    * bugfix: 修复进程退出时未释放signal和sighand的bug

    * 增加赞赏者名单: David Wen

commit 0274cd6eeec01885232e7418a501857cb76da69e
Author: login <longjin@ringotek.cn>
Date:   Wed Nov 23 20:43:18 2022 +0800

    修正drop signal结构体的box对象的的问题 (#92)

    * fix: exit signal and exit sighand

commit c8025a88798dc57ecc5d7f20ad69de695445638f

```

(下页继续)

(续上页)

Author: login <longjin@ringotek.cn>

Date: Wed Nov 23 20:18:22 2022 +0800

new: 在fork时拷贝signal和sighand (#91)

* refcount初始化

* new: 实现copy_sighand

del: 删除sighand_struct的wqh, 待将来有需要时, 替换成rust版本的

* new: 拷贝signal

bugfix: 解决拷贝sighand时的uaf问题

commit 66f67c6a95b8aad85cfd2146a86e5e3e6a3568e7

Author: login <longjin@ringotek.cn>

Date: Wed Nov 23 11:38:20 2022 +0800

signal的发送 (暂时父子进程之间共享信号及相应的结构体) (#89)

* 解决由于spinlock.h中包含preempt_enable()带来的循环include问题

* new: 初步实现signal的数据结构

* new: signal相关数据结构

* fix: 解决bindings.rs报一堆警告的问题

* new: rust下的kdebug kinfo kwarn kBUG kerror宏

* 移动asm.h和cmpxchg.h

* new: signal的发送 (暂时只支持父子进程共享信号及处理函数)

commit 3d729e2069e01ee07525ff83167566dac5322a40

Author: login <longjin@ringotek.cn>

Date: Fri Nov 18 17:59:33 2022 +0800

bugfix: 修复进程pcb被回收时, 未将其从链表中删除的问题 (#87)

* bugfix: 修复进程pcb被回收时, 未将其从链表中删除的问题

new: pcb相关api文档

* 将文档加入目录

(下页继续)

(续上页)

```
commit 0bfe94f46be9bdde1ade81a20e803aa2aafd2964
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Fri Nov 18 16:32:15 2022 +0800
```

```
new: rust下的kdebug kinfo kwarn kBUG kerror宏 (#86)
```

```
* new: rust下的kdebug kinfo kwarn kBUG kerror宏
```

```
commit c6174797dcf3427f38bfa0f4bd3e039c319f7c5b
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Thu Nov 17 20:29:29 2022 +0800
```

```
fix: 解决bindings.rs报了很多警告的问题 (#85)
```

```
* fix: 解决bindings.rs报一堆警告的问题
```

```
commit cffd7144fbed84f9775e89d7b99602c6ccc5a510
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Wed Nov 16 15:18:03 2022 +0800
```

```
signal相关数据结构&代码结构优化 (#84)
```

```
* 解决由于spinlock.h中包含preempt_enable()带来的循环include问题
```

```
* new: 初步实现signal的数据结构
```

```
commit fb6c29d01d4cf92368efec08c01e419c2a941f7d
```

```
Author: login <longjin@ringotek.cn>
```

```
Date: Sun Nov 13 16:43:58 2022 +0800
```

```
port kmalloc and printk to rust (#83)
```

```
* 暂时移除cbindgen
```

```
* 将lib文件夹更名为libs文件夹 (解决rust的冲突)
```

```
* 实现了全局的allocator
```

```
* 实现了printk宏
```

```
* new: 完善了printk的颜色
```

(下页继续)

(续上页)

```
commit 82d2e446a401e7eee57a847f48a6d162931170c3
Author: login <longjin@ringotek.cn>
Date: Sat Nov 12 15:25:54 2022 +0800
```

new: 暂时移除cbindgen (#82)

```
commit 2aaf7808efe44ecfaadd51ae4f8892e667108578
Author: login <longjin@ringotek.cn>
Date: Fri Nov 11 22:21:44 2022 +0800
```

在内核中引入cbindgen, 生成rust-C的FFI (#81)

* 解决codeql失败问题

* new: 为内核引入cbindgen

```
commit 2813126e3190c9b3c1a836a647b259a7adbe0cf3
Author: login <longjin@ringotek.cn>
Date: Fri Nov 11 15:35:37 2022 +0800
```

新增rust ffi (#77)

* 引入cargo

* 取消对Cargo.lock的跟踪

* 解决vscode报错问题

* new: rust的代码能够调用c语言的printf_color

* 1、将原本run.sh的工作拆解, 变为几个不同的make命令
2、在docker镜像中编译rust

* 更改workflow

* update workflow

* new: 解决workflow无法通过编译的问题

```
commit 5e023cf7911333eb05bfe65704dce4b01fa4d0a7
Author: login <longjin@ringotek.cn>
Date: Fri Nov 11 15:21:45 2022 +0800
```

(下页继续)

(续上页)

```
Update makefile.yml

commit e44795008f7e34d2068cf28dcedbcb91f5ccd66b
Author: login <longjin@ringotek.cn>
Date:   Fri Nov 11 15:18:13 2022 +0800

Update makefile.yml (#80)

commit ec5fb84b61c313824cc2199ab64e3af4b7e5f895
Author: login <longjin@ringotek.cn>
Date:   Fri Nov 11 15:08:09 2022 +0800

Update makefile.yml

commit 6d9dff5f1ff347ea780a0249e54eef356cdcaaea
Author: login <longjin@ringotek.cn>
Date:   Fri Nov 11 15:07:48 2022 +0800

Revert "Update makefile.yml (#78)" (#79)

This reverts commit badc7d238f2341e844a90be3e357e5dd77a447fc.

commit badc7d238f2341e844a90be3e357e5dd77a447fc
Author: login <longjin@ringotek.cn>
Date:   Fri Nov 11 15:05:52 2022 +0800

Update makefile.yml (#78)
```

19.6 V0.1.0

备注：本文作者：龙进 longjin@RinGoTek.cn

2022 年 11 月 6 日

19.6.1 前言

DragonOS 从 2022 年 1 月 15 日开始开发，到如今已经经历了将近 300 天。在这么多个日夜里，已经数不清到底花了多少时间在 DragonOS 的开发之中，我基本上把所有的空闲时间都给了 DragonOS，保守估计总工时已经在 1000 小时以上。能够发布第一个版本，我感到十分有成就感。

在 2022 年 7 月以来，陆陆续续的，有来自 6 所高校或企业的小伙伴/大佬加入了 DragonOS 的开发。我当时非常的欣喜，我想，也许在大家的一同努力下，我们能创造出一个真正具有实用性的操作系统呢！我们累计召开了 14 次交流讨论会。我相信，在大家的共同努力下，将来，我们一定能创造出真正独立自主的、开放的、面向服务器领域应用的开源操作系统，并在生产环境中得到应用。

尽管 DragonOS 目前只是一个玩具水平的操作系统，只是“比本科生毕业设计难度略高的”操作系统。但是，请不要小看它，它的内在的架构设计，瞄准了 Linux 5.18 及以后的发行版，虽尚未能达到 Linux 的水平，但我们正努力追赶。得益于 Linux 的相关资料，DragonOS 在架构设计之时，学习了 Linux 的很多设计思想，相关组件都尽量考虑了可扩展性与可移植性。

千里之行，始于足下。DragonOS V0.1.0 版本的发布，是一个全新的开始。希望在未来的十年里，我们能与众多伙伴们一同努力，在 2032 年，将 DragonOS 建设成为具有实用意义的，能够在服务器领域取得广泛应用的开源操作系统！

百舸争流，奋楫者先；中流击水，勇进者胜。我相信，在接下来的时间里，在社区开发者们的不断努力下，我们的目标，终将成为现实！

19.6.2 特别鸣谢

在 DragonOS V0.1.0 版本的发布之际，我想对我的老师、前辈以及学校表示衷心的感谢！

- **佛山市南海区大沥镇中心小学姚志城老师**：您是带领我接触计算机，学会编程的领路人。十年前，与您交谈时，您说过：“我们国家目前还没有自主的、成熟的操作系统”。这句话，为我的梦想埋下了种子。您培养了我对计算机的热爱，因此我选择了软件工程这个专业。感谢当年您的教导，师恩难忘！
- **佛山市南海区石门实验学校**：在石实就读的三年里，非常感谢石实的“扬长教育”理念，在老师们的培养下，让我充分发挥了自己的个性和特长，也取得了不错的成绩。在石实的三年里，我学会了 C++、Java 以及简单的算法，也自己开发了几个安卓 app，积累了将近 6 千行的代码量。
- **佛山市南海区石门中学**：“任重道远，毋忘奋斗”是石中的校训，我想，这句校训，也应当成为我们每个新时代青年人的座右铭。在石门中学的三年，家国情怀教育对我产生了很大的影响。我想，我们作为新时代的青年，更应当肩负起时代的重任，奋勇拼搏，为祖国的发展，为民族的自强，为人类的未来，努力奋斗！
- **华南理工大学**：“博学慎思，明辨笃行”，在华工，我得到了进一步的学习与发展。开拓了自己的视野，学会了跟很多人打交道。并且，在软件学院，我遇到了一群认真负责的老师。非常感谢学院对我的支持，支持我们成立项目组。我相信，在学院的支持下，能让 DragonOS 取得更好的发展，走的更远！
- **华南理工大学软件学院王国华老师**：王老师是我《操作系统》课程的老师，在她的指导下，我对操作系统的原理有了更深的理解，并参加了“泛珠三角+大学生计算机作品赛”，在 2022 年 6 月的广东省选拔赛中，DragonOS 取得了一等奖、最佳创新奖的好成绩。

- **华南理工大学软件学院汤峰老师**：汤老师是我们在校内的项目组的指导老师。在她的悉心指导下，我们将不断前行，保持我们前进的方向，持续构建开源社区。我由衷地感谢汤老师的指导！
- **Yaotian Feng**：在 Bilibili 上认识了这位非常厉害的老哥，解答了我的很多问题，很多次在我毫无头绪的 debug 了几天之后，几句话点醒了我，让我找到解决问题的路径。并且，他也跟我分享了容易踩坑的地方，让我在将要踩坑的时候能够有个心理预期，不至于那么难受哈哈哈哈哈。

19.6.3 贡献者名单

DragonOS V0.1.0 版本的发布，离不开以下小伙伴们共同努力：

- 龙进 longjin@RinGoTek.cn
- zzy666-hw zzy666@mail.ustc.edu.cn
- 关锦权 guanjinquan@DragonOS.org
- 周于[㊦] zhouyuzhe@DragonOS.org
- kkkkkong kongweichao@DragonOS.org
- houmkh jiaying.hou@qq.com
- wang904 1234366@qq.com
- Liric Mechan i@liric.cn
- Mustang handsomepd@qq.com
- Eugene caimal2138@foxmail.com
- kun 1582068144@qq.com
- zhujikuan 1335289286@qq.com
- Alloc Alice 1548742234@qq.com

19.6.4 赞助者名单

感谢以下同学的赞赏，我们将不断努力！

- TerryLeeSCUT
- 悟
- slientbard

19.6.5 内核

遵循的一些标准规范

- 启动引导：Multiboot2
- 系统接口：posix 2008

硬件架构

- 目前支持在 x86-64 架构的处理器上运行

Bootloader

- 使用 Grub 2.06 作为 bootloader

内存管理

- 实现了基于 bitmap 的页分配器
- 实现了 slab 分配器，用来分配小块的、具有对齐要求的内存
- 抽象出 VMA（虚拟内存区域）
- 实现 VMA 反向映射机制
- 实现 MMIO 地址空间自动映射机制

多核

- 支持多核引导。也就是说，在 DragonOS 启动后，将会启动 AP 处理器。但是，为了简化其他内核模块的实现，目前 AP 处理器上，暂时没有任务在运行。
- 粗略实现了 IPI（处理器核间通信）框架

进程管理

- 支持进程的创建、回收
- 内核线程
- kthread 机制
- 用户态、内核态进程/线程的 fork/vfork（注意，用户态的 fork 和内核态的有一定的区别，内核态的 fork 更复杂）
- exec 让进程去执行一个新的可执行文件

- 进程的定时睡眠 (sleep) (支持 spin/rdtsc 高精度睡眠、支持上下文切换方式的睡眠)

同步原语

- spinlock 自旋锁
- mutex 互斥量
- atomic 原子变量
- wait_queue 等待队列
- semaphore 信号量

调度相关

- CFS 调度器
- 单核调度 (暂时不支持多核负载均衡)
- completion “完成” 机制, 让一个进程能等待某个任务的完成。

IPC 进程间通信

- 匿名管道

文件系统

- VFS 虚拟文件系统的基本功能
- FAT32 文件系统 (尚不支持删除文件夹)
- devfs 设备文件系统。目前只将键盘文件注册到其中。
- rootfs 根文件系统, 在真实的磁盘文件系统被挂载前, 为其他的伪文件系统提供支持。
- 挂载点抽象。目前实现了文件系统的挂载, 使用类似于栈的方式管理所有的挂载点。(将来需要优化这部分)

异常及中断处理

- 处理器异常的捕获
- 对 APIC 的支持
- softirq 软中断机制
- 能够对内核栈进行 traceback

内核数据结构

- 普通的二叉树
- kfifo 先进先出缓冲区
- 循环链表
- IDR 映射数据结构
- IDA ID 分配数据组件

屏幕显示

- VESA VBE 显示芯片驱动
- 实现了屏幕管理器，支持多个显示框架注册到屏幕管理器中。
- 实现了 TextUI 文本界面框架，能够渲染文本到屏幕上。并且预留了上下滚动翻页、多显示窗口的支持。
- printk

内核实用库

- 字符串操作库
- ELF 可执行文件支持组件
- 基础数学库
- CRC 函数库

软件移植

- 移植了 LZ4 压缩库 (V1.9.3)，为将来实现页面压缩机制打下基础。

内核测试

- ktest 单元测试框架
- 支持使用串口 (COM1) 输出屏幕内容到文件之中。

驱动程序支持

- IDE 硬盘
- AHCI 硬盘 (SATA Native)
- ACPI 高级电源配置模块
- PCI 总线驱动
- XHCI 主机控制器驱动 (usb3.0)
- ps/2 键盘
- ps/2 鼠标
- HPET 高精度定时器
- RTC 时钟
- local APIC 定时器
- UART 串口 (支持 RS-232)
- VBE 显示
- 虚拟 tty 设备

系统调用

DragonOS 目前一共有 22 个有效的系统调用。

- SYS_PUT_STRING 往屏幕上打印字符
- SYS_OPEN 打开文件
- SYS_CLOSE 关闭文件
- SYS_READ 读取文件
- SYS_WRITE 写入文件
- SYS_LSEEK 调整文件指针
- SYS_FORK fork 系统调用
- SYS_VFORK vfork 系统调用
- SYS_BRK 调整堆大小为指定值
- SYS_SBRK 调整堆大小为相对值
- SYS_REBOOT 重启 (将来 sysfs 完善后, 将删除这个系统调用, 请勿过度依赖这个系统调用)
- SYS_CHDIR 切换进程的工作目录
- SYS_GET_DENTS 获取目录中的目录项的元数据

- SYS_EXECVE 让当前进程执行新的程序文件
- SYS_WAIT4 等待进程退出
- SYS_EXIT 退出当前进程
- SYS_MKDIR 创建文件夹
- SYS_NANOSLEEP 纳秒级睡眠（最长 1 秒）在小于 500ns 时，能够进行高精度睡眠
- SYS_CLOCK 获取当前 cpu 时间
- SYS_PIPE 创建管道
- SYS_MSTAT 获取系统当前的内存状态信息
- SYS_UNLINK_AT 删除文件夹或删除文件链接

Rust 支持

- 实现了一个简单的 rust 语言的 hello world，计划在接下来的版本中，逐步转向使用 rust 进行开发。

19.6.6 用户环境

LibC

LibC 是应用程序与操作系统交互的纽带。DragonOS 的 LibC 实现了一些简单的功能。

- malloc 堆内存分配器
- 基础数学库
- 简单的几个与文件相关的函数
- pipe
- fork/vfork
- clock
- sleep
- printf

Shell 命令程序

- 基于简单的字符串匹配的解析（不是通过编译课程学的的那一套东西做的，因此比较简单，粗暴）
- 支持的命令：ls,cd,mkdir,exec,about,rmdir,rm,cat,touch,reboot

用户态驱动程序

- 用户态键盘驱动程序

19.6.7 源码、发布版镜像下载

您可以通过以下方式获得源代码：

通过 Git 获取

- 您可以访问<https://github.com/fslongjin/DragonOS/releases>下载发布版的代码，以及编译好的，可运行的磁盘镜像。
- 我们在 gitee 上也有镜像仓库可供下载：<https://gitee.com/DragonOS/DragonOS>

通过 DragonOS 软件镜像站获取

为解决国内访问 GitHub 慢、不稳定的问题，同时为了方便开发者们下载 DragonOS 的每个版本的代码，我们特意搭建了镜像站，您可以通过以下地址访问镜像站：

您可以通过镜像站获取到 DragonOS 的代码压缩包，以及编译好的可运行的磁盘镜像。

- <https://mirrors.DragonOS.org>
- <https://mirrors.DragonOS.org.cn>

19.6.8 开放源代码声明

备注：为促进 DragonOS 项目的健康发展，DragonOS 以 GPLv2 开源协议进行发布。所有能获得到 DragonOS 源代码以及相应的软件制品（包括但不限于二进制副本、文档）的人，都能享有我们通过 GPLv2 协议授予您的权利，同时您也需要遵守协议中规定的义务。

这是一个相当严格的，保护开源软件健康发展，不被侵占的协议。

对于大部分的善意的人们而言，您不会违反我们的开源协议。

我们鼓励 DragonOS 的自由传播、推广，但是请确保所有行为没有侵犯他人的合法权益，也没有违反 GPLv2 协议。

请特别注意，对于违反开源协议的，尤其是**商业闭源使用以及任何剽窃、学术不端行为将会受到严肃的追责**。（这是最容易违反我们的开源协议的场景）。

并且，请注意，按照 GPLv2 协议的要求，基于 DragonOS 修改或二次开发的软件，必须同样采用 GPLv2 协议开源，并标明其基于 DragonOS 进行了修改。亦需保证这些修改版本的用户能方便的获取到 DragonOS 的原始版本。

您必须使得 DragonOS 的开发者们，能够以同样的方式，从公开渠道获取到您二次开发的版本的源代码，否则您将违反 GPLv2 协议。

关于协议详细内容，还敬请您请阅读项目根目录下的 **LICENSE** 文件。请注意，按照 GPLv2 协议的要求，**只有英文原版才具有法律效力**。任何翻译版本都仅供参考。

开源软件使用情况

DragonOS 在开发的过程中，参考了一些开源项目的设计，或者引入了他们的部分代码，亦或是受到了他们的启发。现将他们列在下面。我们对这些开源项目的贡献者们致以最衷心的感谢！

格式：< 项目名 > - < 链接 > - < 开源协议 >

- Linux - <https://git.kernel.org/> - GPLv2
- skiftOS - <https://github.com/skiftOS/skift> - MIT
- FYSOS - <https://github.com/fysnet/FYSOS> - [FYSOS' License](#)
- LemonOS - <https://github.com/LemonOSProject/LemonOS.git> - BSD 2-Clause License
- LZ4 - <https://github.com/lz4/lz4> - BSD 2-Clause license
- SerenityOS - <https://github.com/SerenityOS/serenity.git> - BSD 2-Clause license
- MINE - 《一个 64 位操作系统的设计与实现》田宇; 人民邮电出版社
- chcore - 《现代操作系统：设计与实现》陈海波，夏虞斌; 机械工业出版社
- SimpleKernel - <https://github.com/Simple-XX/SimpleKernel> - MIT

CHAPTER 20

Indices and tables

- `genindex`
- `modindex`
- `search`