

lab2:物理内存和页表

小组成员：

吴禹骞-2311272

谢小珂-2310422

杜泽琦-2313508

lab2:物理内存和页表

前期准备

实验目的

实验内容

老师实验视频知识点

练习

练习1：理解first-fit 连续物理内存分配算法（思考题）

前置知识点

物理内存探测

页面分配算法

first-fit算法

代码分析

default_init

default_init_memmap

default_alloc_pages

default_free_pages

default_nr_free_pages

basic_check

default_check

const struct pmm_manager default_pmm_manager

优化改进

练习2：实现 Best-Fit 连续物理内存分配算法（需要编程）

Best-Fit 算法核心思想

主要修改的代码部分

① 修改点 1： `best_fit_alloc_pages()` 查找最优块

② 修改点 2：注册管理器类型

进一步的改进空间

扩展练习Challenge：buddy system（伙伴系统）分配算法（需要编程）

1. 核心思想

2. 数据结构

3. 算法实现概述

4. 主要代码

准备工作

寻找伙伴页块

分配分裂页块

释放合并页块

5. 测试

初始化测试

分配释放最小块测试

复杂内存分配和释放测试

分配释放最大块测试

扩展练习Challenge：任意大小的内存单元slub分配算法（需要编程）

SLUB 分配算法的核心思想

总体思路

数据结构设计

1. Slab 结构体
2. Cache 结构体

核心函数逻辑分析

1. Cache 初始化
2. 创建 Slab
3. 对象分配
4. 对象释放

检查函数与验证流程

实验结果分析

可改进方向

扩展练习Challenge：硬件的可用物理内存范围的获取方法（思考题）

1. 逐块物理探测法（Memory Probing）
2. DMA 间接探测法（Using DMA Controller）
3. 利用固件接口（Firmware Table）
4. 启发式混合探测法（Hybrid Approach）

综合比较

总结

前期准备

实验二指导

我学习了vscode的连接虚拟机，[教程链接](#)

注意，Ubuntu系统可能默认启动了防火墙，需要在防火墙中打开SSH服务的端口，或者关闭防火墙。

本次 `make qemu`

```
1 root@dzq-VMware-Virtual-Platform:/home/dzq/opt/riscv/labcode/lab2# make qemu
2 + cc kern/init/entry.S
3 + cc kern/init/init.c
4 + cc kern/libs/stdio.c
5 + cc kern/debug/panic.c
6 + cc kern/driver/console.c
7 + cc kern/driver/dtb.c
8 + cc kern/mm/best_fit_pmm.c
9 + cc kern/mm/default_pmm.c
10 + cc kern/mm/pmm.c
11 + cc libs/printfmt.c
12 + cc libs/readline.c
13 + cc libs/sbi.c
14 + cc libs/string.c
15 + ld bin/kernel
16 riscv64-unknown-elf-objcopy bin/kernel --strip-all -O binary bin/ucore.img
17
18 OpensBI v0.4 (Jul  2 2019 11:53:53)
19
20 / _ _ \ / _ _ | _ \ _ |
```

```

21 | | | | _ _ _ _ _ | ( _ | | ) | | |
22 | | | | ' _ \ / _ \ ' _ \ \ _ \ | _ < | |
23 | | _ | | | ) | _ / | | | _ ) | | ) | | _
24 | \ _ _ / | . _ / \ _ | _ | | _ _ / | _ _ / _ _ |
25 | |
26 | _ |
27
28 Platform Name : QEMU Virt Machine
29 Platform HART Features : RV64ACDFIMSU
30 Platform Max HARTs : 8
31 Current Hart : 0
32 Firmware Base : 0x80000000
33 Firmware Size : 112 KB
34 Runtime SBI Version : 0.1
35
36 PMP0: 0x0000000080000000-0x000000008001ffff (A)
37 PMP1: 0x0000000000000000-0xffffffffffffffff (A,R,W,X)
38 DTB Init
39 HartID: 0
40 DTB Address: 0x82200000
41 Physical Memory from DTB:
42 Base: 0x0000000080000000
43 Size: 0x0000000080000000 (128 MB)
44 End: 0x0000000087ffffff
45 DTB init completed
46 (THU.CST) os is loading ...
47 Special kernel symbols:
48 entry 0xffffffffc02000d8 (virtual)
49 etext 0xffffffffc020172e (virtual)
50 edata 0xffffffffc0206018 (virtual)
51 end 0xffffffffc0206078 (virtual)
52 Kernel executable memory footprint: 24KB
53 memory management: default_pmm_manager
54 physcial memory map:
55 memory: 0x0000000080000000, [0x0000000080000000, 0x0000000087ffffff].
56 check_alloc_page() succeeded!
57 satp virtual address: 0xffffffffc0205000
58 satp physical address: 0x0000000080205000
59 QEMU: Terminated

```

实验目的

- 理解页表的建立和使用方法
- 理解物理内存的管理方法
- 理解页面分配算法

实验内容

实验一过后大家做出来了一个可以启动的系统，实验二主要涉及操作系统的物理内存管理。操作系统为了使用内存，还需高效地管理内存资源。本次实验我们会了解**如何发现系统中的物理内存**，然后学习**如何建立对物理内存的初步管理**，即**了解连续物理内存管理**，最后**掌握页表相关的操作**，即**如何建立页表来实现虚拟内存到物理内存之间的映射**，帮助我们对段页式内存管理机制有一个比较全面的了解。本次的实验主要是在**实验一的基础上完成物理内存管理，并建立一个最简单的页表映射**。

老师实验视频知识点

`dzq/opt/riscv/labcode/lab2/kern/init/entry.s` 系统的入口代码

`dzq/opt/riscv/labcode/lab2/tools/kernel.ld` 这是一个链接脚本，这个链接脚本用来指导内核的编译过程当中如何摆放内核得到的那一片连续的地址空间。[理解代码]

一个程序不管它是内核还是普通的应用程序，它的地址空间的布局是在它被编译和链接的时候，就已然由链接器和编译器给你确定好，然后我们只需要把指令填到那里去，填到那个对应的空间布局上就好了

c 语言在运行之前有一个默认的假设条件是栈已然就绪的，我们能看到在这个位置上有一条指令，这条指令是把 SP 寄存器给它设置了一个值，也就是这条指令是让栈寄存器指向一块可以用的内存空间。所以这样之后我们再调用 c 语言的代码就没有问题了，压栈和退栈都有的内存可以用了。

`dzq/opt/riscv/labcode/lab2/kern/mm/best_fit_pmm.c`

练习

对实验报告的要求：

- 基于markdown格式来完成，以文本方式为主
- 填写各个基本练习中要求完成的报告内容
- 列出你认为本实验中重要的知识点，以及与对应的OS原理中的知识点，并简要说明你对二者的含义，关系，差异等方面的理解（也可能出现实验中的知识点没有对应的原理知识点）
- 列出你认为OS原理中很重要，但在实验中没有对应上的知识点
- 从oslab网站上取得实验代码后，进入目录labcodes/lab2，完成实验要求的各个练习。在实验报告中回答所有练习中提出的问题。在目录labcodes/lab2下存放实验报告，推荐用**markdown**格式。每个小组建一个gitee或者github仓库，对于lab2中编程任务，完成编写之后，再通过git push命令把代码和报告上传到仓库。最后请一定提前或按时提交到git网站。

注意有“LAB2”的注释，代码中所有需要完成的地方（challenge除外）都有“LAB2”和“YOUR CODE”的注释，请在提交时特别注意保持注释，并将“YOUR CODE”替换为自己的学号，并且将所有标有对应注释的部分填上正确的代码。

Challenges是选做，完成Challenge并回答了助教问题的小组可获得本次实验的加分。

练习1：理解first-fit 连续物理内存分配算法（思考题）

first-fit 连续物理内存分配算法作为物理内存分配一个很基础的方法，需要同学们理解它的实现过程。请大家仔细阅读实验手册的教程并结合 `kern/mm/default_pmm.c` 中的相关代码，认真分析 `default_init`，`default_init_memmap`，`default_alloc_pages`，`default_free_pages`等相关函数，并描述程序在进行物理内存分配的过程以及各个函数的作用。请在实验报告中简要说明你的设计实现过程。请回答如下问题：

- 你的first fit算法是否有进一步的改进空间？

前置知识点

物理内存管理应当为虚拟内存管理提供这样的接口：

- 检查当前还有多少空闲的物理页，返回空闲的物理页数目
- 给出n，尝试分配n个物理页，可以返回一个起始地址和连续的物理页数目，也可能分配一些零散的空闲物理页，返回一个连起来的链表。
- 给出起始地址和n，释放n个连续的空闲物理页

物理内存探测

页面分配算法

first-fit算法

当需要分配或释放内存时：

- **分配**：从链表头开始，找到第一个满足大小要求的空闲块；
- **释放**：将页块按地址顺序重新插入链表，并合并前后相邻的空闲块；
- **统计**：通过全局变量 `nr_free` 记录当前空闲页总数。

代码分析

详细代码讲解我已放入文件中，这里只做核心讲解，关键点我已加粗，这部分主要是代码阅读，所以主要看检查

default_init

```
1  static void
2  default_init(void) {
3      list_init(&free_list); // 初始化空闲链表为空
4      nr_free = 0;           // 当前系统空闲页数清零
5  }
```

这里首先我们要做的就是初始化内存分配器，`free_list` 是一个**双向循环链表**，用于记录系统中所有“空闲页块”；每个空闲块在链表中用一个 `Page` 节点表示；实验刚开始时，还没有任何可分配内存页，所以 `nr_free = 0`。

default_init_memmap

对一段连续的物理页（从 `base` 开始，共 `n` 页）进行初始化，使其成为系统可管理的空闲内存块。函数首先遍历这段页区间，将每个页的标志位和引用计数清零，确保这些页处于“空闲”状态；然后将第一个页（即块首页）的 `property` 字段设为块的页数 `n`，并标记该页为空闲块的起始页；接着把这段空闲块的页数累加到全局变量 `nr_free` 中，更新系统的空闲页统计；最后，函数将该空闲块按物理地址顺序插入到全局空闲链表 `free_list` 中，以保证链表中所有空闲块的顺序性，从而方便后续的页面分配与合并操作。总体而言，它完成了从“未初始化的物理页”到“可管理的空闲块”的转换，是物理内存管理建立阶段的关键步骤。

default_alloc_pages

函数的作用是实现 **First-Fit** 页面分配算法，用于从系统空闲页链表中找到一块能够满足请求页数 `n` 的连续物理页，并将其分配出去。函数首先检查请求页数是否合法以及当前空闲页是否充足；然后从空闲页链表头开始按顺序遍历，找到第一个 `property` 值大于或等于 `n` 的空闲块（即第一个足够大的连续空闲页段）；若找到合适块，则将其从链表中移除，并在块过大的情况下对其进行**分裂**——前 `n` 页分配出去，剩余部分重新标记为空闲块并插回链表。最后更新

全局空闲页计数 `nr_free`，清除被分配块的空闲标志并返回起始页地址。整体上，这个函数完成了“在空闲块链表中查找、分割并分配物理页”的全过程，是物理内存管理中最核心的分配逻辑。

default_free_pages

将释放的 n 个连续页重新加入空闲链表，并在必要时**合并相邻空闲块**，避免碎片。这段代码的功能是将一段连续的物理页释放回系统空闲页池。首先，它遍历释放的页，将每页的标志和引用计数清零，**确保不是保留页且不是已有空闲块起点**；然后，将释放块的起始页设置为该块大小并标记为空闲块起点，同时更新总空闲页数；最后，将释放块按照物理地址从低到高的顺序插入空闲链表，以便后续分配连续页时可以高效查找并支持相邻空闲块的合并。如果释放块与前后块在物理地址上连续，则将它们合并为更大的空闲块。释放成功后 `nr_free` 增加释放的页数。

default_nr_free_pages

返回当前空闲页总数

basic_check

验证物理内存分配器是否正确工作的核心自检函数，测试了**分配正确性**，**引用计数正确性**，**地址合法性**，**空闲列表逻辑**，**释放与再分配**，**状态恢复**。

default_check

对First Fit内存页分配算法的正确性进行全面验证。它首先遍历空闲页链表，统计空闲块数量与总页数，并与系统记录的空闲页数比对，确保链表状态一致；然后调用 `basic_check()` 进行基础功能测试；接着通过一系列分配与释放操作，验证首次适应算法在不同场景下的行为是否正确，如部分释放、连续分配、块合并、无法分配时返回 `NULL` 等；最后恢复空闲链表与计数状态，并再次检查页数统计是否准确。整体而言，该函数用于自动检测内存分配器在“首次适应”策略下的页管理逻辑、链表维护及内存块合并功能是否实现正确。

const struct pmm_manager default_pmm_manager

定义默认的物理内存管理器，包含初始化、内存块初始化、页分配、页释放、空闲页计数和检查函数的实现

优化改进

是的可以优化，我从以下方向入手思考，其实下面的那些算法就相当于优化了，甚至我这两天学懂了才想起来软件安全里有讲过这个知识点，就简单提一下优化思路方向

1.减少碎片化内存，比如First-Fit 会优先使用低地址的小块以及default_alloc_pages块过大的情况下对其进行**分裂**，我想到可以就是每次遍历找到满足要求最小的内存块进行处理，但这会加大时间复杂度，那么就可以再来一个链表，记录从小到大的位置。

练习2：实现 Best-Fit 连续物理内存分配算法（需要编程）

在完成练习一后，参考kern/mm/default_pmm.c对First Fit算法的实现，编程实现Best Fit页面分配算法，算法的时空复杂度不做要求，能通过测试即可。请在实验报告中简要说明你的设计实现过程，阐述代码是如何对物理内存进行分配和释放，并回答如下问题：

- 你的 Best-Fit 算法是否有进一步的改进空间？

在完成练习一后，参考kern/mm/default_pmm.c对First Fit算法的实现，编程实现Best Fit页面分配算法，算法的时空复杂度不做要求，能通过测试即可。请在实验报告中简要说明你的设计实现过程，阐述代码是如何对物理内存进行分配和释放，并回答如下问题：

- 你的 Best-Fit 算法是否有进一步的改进空间？

Best-Fit 算法核心思想

将分区按大小顺序组织,找到的第一个适应分区是大小与要求相差最小的空闲分区

也就是说,在所有可用的空闲内存块中,找到“刚好够大”的那一块来满足请求,即剩余空间最小的空闲块。其目标是最大限度地减少“内碎片”(内部未使用空间),尽可能让每次分配都更加精确。因此,与 First-Fit (首次适应法) 相比, Best-Fit 在分配时需要遍历整个空闲链表,寻找最佳匹配,而不是遇到第一个足够大的块就立即分配。

主要修改的代码部分

Best-Fit 算法是在 `best_fit_pmm.c` 中实现的,它与 `default_pmm.c` (First-Fit) 基本框架相同,关键修改集中在页面分配函数:

① 修改点 1: `best_fit_alloc_pages()` 查找最优块

First-Fit 原逻辑:

```
1 while ((le = list_next(le)) != &free_list) {
2     struct Page *p = le2page(le, page_link);
3     if (p->property >= n) {
4         page = p;
5         break;    // 找到第一个够大的块就分配
6     }
7 }
```

Best-Fit 改进逻辑:

```
1 size_t min_diff = (size_t)-1;
2 struct Page* best = NULL;
3
4 while ((le = list_next(le)) != &free_list) {
5     struct Page* p = le2page(le, page_link);
6     if (p->property >= n && (p->property - n) < min_diff) {
7         min_diff = p->property - n;
8         best = p;    // 记录当前最接近需求的块
9     }
10 }
11 page = best;    // 使用最优块进行分配
```

区别:

- First-Fit: 遇到第一个满足条件的块立即停止;
- Best-Fit: 遍历完整个链表,找出最“贴合”的块;
- 分配后若块大于需求,还会执行同样的分裂逻辑。

② 修改点 2: 注册管理器类型

在 `pmm.c` 文件中:

```
1 // 原 First-Fit
2 // pmm_manager = &default_pmm_manager;
3
4 // 修改为 Best-Fit
5 pmm_manager = &best_fit_pmm_manager;
```

这使得系统初始化时自动加载 Best-Fit 分配器。

如同，make grade后成功通过测试：

进一步的改进空间

尽管 Best-Fit 能较好地减少单次分配的内部碎片，但整体上仍存在查找效率低、易形成外部碎片等问题。未来可在算法层面引入平衡树结构以加快搜索，在策略层面采用分级空闲链表与动态切换机制，在实现层面改进合并与碎片整理逻辑，从而在性能与内存利用率之间实现更优平衡。

扩展练习Challenge：buddy system（伙伴系统）分配算法（需要编程）

Buddy System算法把系统中的可用存储空间划分为存储块(Block)来进行管理, 每个存储块的大小必须是2的n次幂 (Pow(2, n)), 即1, 2, 4, 8, 16, 32, 64, 128...

- 参考[伙伴分配器的一个极简实现](#)，在ucore中实现buddy system分配算法，要求有比较充分的测试用例说明实现的正确性，需要有设计文档。

1. 核心思想

Buddy System (伙伴系统) 是一种动态内存分配算法，它将整个空闲内存块按照 2^n 的大小进行划分和管理。

- 分配时：如果需要的块大小 N 不等于 2^n 的空闲块 M ，则将 M 分裂成两个大小为 $M/2$ 的“伙伴”块，直到得到一个大小为 $2^k \geq N$ 的最小块，并将多余的伙伴块放入对应的空闲列表中。
- 释放时：系统尝试将释放的块与其相邻的空闲伙伴块合并，形成一个更大的块，直到无法合并或达到最大块大小。

2. 数据结构

结构/宏	描述
<code>MAX_ORDER</code>	最大阶数。设置为14(后续有解释)。
<code>BUDDY_AREA</code>	包含 Buddy System 的核心数据结构。
<code>free_area[MAX_ORDER]</code>	空闲列表数组。 <code>free_area[i]</code> 管理大小为 2^i 页的空闲块链表。
<code>buddy_nr_free</code>	整个Buddy System中所有空闲页的总数。
<code>struct Page::property</code>	在Buddy System中将记录块的阶数 i （即 2^i 页）。

3. 算法实现概述

函数	描述
buddy_init()	初始化 BUDDY_AREA 结构：清零 buddy_nr_free，并初始化所有 free_area[i].free_list 链表头。
buddy_init_memmap(base,n)	初始化内存映射。它将一个大小为 n 页的连续空闲块进行分解。首先将 n 拆分成 2^n 形式的块 (例如 $1025 = 1024 + 1$ 页)，然后将这些块插入到对应的 free_area[i] 链表中。
buddy_alloc_pages(n)	分配。 首先计算所需的最小阶数 $k = \lceil \log_2 n \rceil$ 。从 free_area[k] 开始向上搜索，找到最小的 $m \geq k$ 的空闲块。如果 $m > k$ ，则执行 分裂 操作。
buddy_free_pages(base,n)	释放和合并。 确定释放块的阶数 $k = \log_2 n$ 。尝试与伙伴块进行合并，如果伙伴块空闲且阶数相同，则合并为 $k + 1$ 阶的块，并递归向上合并。
get_order(n)	计算 n 页需要内存块的最小阶数
get_buddy_page(p,order)	计算页块 p 的伙伴页块地址
buddy_show()	可视化 buddy system 内存链表数组

4. 主要代码

准备工作

首先参考原来的空闲链表的结构体，设计了一个数据结构来存放内存块：

```
1 typedef struct {
2     free_area_t free_area[MAX_ORDER+1]; // 0 到 MAX_ORDER 阶的空闲链表
3     unsigned int nr_free;                // 总空闲页数
4 } buddy_area_t;
```

通过测试（主体代码完成后的测试）我们可知可分配内存一共有31929页（pmm.c中计算得到），14阶空闲链表就可满足，总空闲内存的起始 PPN 为 0x80347，并没有对其页数：

```
1 -----Check Begin-----
2 Total Free Pages: 31929
3 阶数 0(1 页): 空闲 1 块, 总空闲 1 页. 内存块起始PPN:
4   [ 0x87fff(00) ]
5 阶数 1(2 页): EMPTY
6 阶数 2(4 页): EMPTY
7 阶数 3(8 页): 空闲 1 块, 总空闲 8 页. 内存块起始PPN:
8   [ 0x87ff7(03) ]
9 阶数 4(16 页): 空闲 1 块, 总空闲 16 页. 内存块起始PPN:
10  [ 0x87fe7(04) ]
11 阶数 5(32 页): 空闲 1 块, 总空闲 32 页. 内存块起始PPN:
12  [ 0x87fc7(05) ]
13 阶数 6(64 页): EMPTY
14 阶数 7(128 页): 空闲 1 块, 总空闲 128 页. 内存块起始PPN:
15  [ 0x87f47(07) ]
```

```

16 阶数 8(256 页): EMPTY
17 阶数 9(512 页): EMPTY
18 阶数 10(1024 页): 空闲 1 块, 总空闲 1024 页. 内存块起始PPN:
19 [ 0x87b47(010) ]
20 阶数 11(2048 页): 空闲 1 块, 总空闲 2048 页. 内存块起始PPN:
21 [ 0x87347(011) ]
22 阶数 12(4096 页): 空闲 1 块, 总空闲 4096 页. 内存块起始PPN:
23 [ 0x86347(012) ]
24 阶数 13(8192 页): 空闲 1 块, 总空闲 8192 页. 内存块起始PPN:
25 [ 0x84347(013) ]
26 阶数 14(16384 页): 空闲 1 块, 总空闲 16384 页. 内存块起始PPN:
27 [ 0x80347(014) ]

```

由于 Buddy System 必须严格满足页地址和大小对齐来确保伙伴页块合并正确, 所以我们找到第一个能被 16384 整除的 PPN, 并从那里开始 Buddy System 的管理, 放弃前面的页。修改 pmm.c 文件:

```

1  #define MAX_ORDER_PAGES (1UL << 14)
2  ...
3  void page_init(void) {
4      // ... 原有代码 (计算 pages 数组, 设置保留位等) ...
5      // 计算当前起始 PPN
6      size_t start_ppn = PPN(mem_begin);
7      // 找到第一个满足 16384 页对齐的 PPN
8      size_t aligned_ppn = ROUNDUP(start_ppn, MAX_ORDER_PAGES);
9      // 计算新的 Buddy System 起始物理地址
10     uintptr_t new_mem_begin = aligned_ppn << PGSHIFT;
11     // 更新 mem_begin, 以新的对齐地址作为 Buddy System 的起点
12     mem_begin = new_mem_begin;
13
14     if (mem_begin < mem_end) {
15         init_memmap(pa2page(mem_begin), (mem_end - mem_begin) / PGSIZE);
16     }
17 }

```

这样Buddy System 将从 PPN 0x84000 (可以被 16384 整除) 开始管理内存, 放弃了从 0x80347000 到 0x84000000 之间的 15545 个页, 将这些页保留给内核, 不参与 Buddy System 的分配。

寻找伙伴页块

```

1  // p:页块的起始Page结构体指针    order:页块的阶数(2^order页)
2  static struct Page *get_buddy_page(struct Page *p, size_t order){
3      size_t page_num = page2ppn(p); // 获取页块物理起始地址
4      size_t buddy_page_num = page_num ^ (1 << order); // 计算伙伴页块物理起始地址
5      return pa2page(buddy_page_num << PGSHIFT); // 将伙伴页块物理起始地址转为其Page结构体指针
6  }

```

在 Buddy System 中, 任何一个大小为 2^k 页的块, 都有一个唯一的“伙伴”块, 它与原块大小相同、地址相邻。这两个块如果都空闲, 就可以合并成一个大小为 2^{k+1} 页的块。

因为 Buddy System 中所有块都必须严格满足 2^k 页的地址和大小对齐要求, 所以一个大小为 2^k 页的内存块 P , 其起始物理页号 (PPN) 最低 k 位必须全部为 0; 而与伙伴块合并后形成的 2^{k+1} 页的更大的块 M 的起始 PPN 最低 $k+1$ 位必须全部为 0。因此, 两个 2^k 的伙伴块 P 和 P' 的起始 PPN, 它们在合并后形成的 M 块的起始 PPN 最低

$k + 1$ 位中，只有第 k 位（从 0 开始计数）是不同的，我们只需翻转 P 块起始 PPN 的第 k 位，即起始 PPN 异或 2^k ，就可得到伙伴块 P' 的起始 PPN。

分配分裂页块

```
1 // 分配n个连续页
2 static struct Page *buddy_alloc_pages(size_t n){
3     assert(n > 0);
4     if(n > buddy_nr_free) return NULL; // 空闲页不足
5     // 计算所需的最小阶数 k
6     size_t min_order = get_order(n);
7     if(min_order >= MAX_ORDER) return NULL; // 太大
8     // 从第k层向上找空闲块
9     size_t order = min_order;
10    while(order <= MAX_ORDER){
11        // order阶的空闲链表是否为空，如果为空则继续向上找，反之则找到空闲块
12        if(list_empty(&(buddy_free_area[order].free_list))){
13            order++;
14        }else{
15            break;
16        }
17    }
18    if(order == MAX_ORDER+1) return NULL; // 空闲块不足
19    // 从列表中取出找到的空闲块
20    list_entry_t *target = list_next(&(buddy_free_area[order].free_list));
21    struct Page *page = 1e2page(target, page_link);
22    list_del(target);
23    // 更新buddy system 信息
24    size_t size = 1 << order;
25    buddy_free_area[order].nr_free -= size;
26    buddy_nr_free -= size;
27    // 如果阶数大于所需阶数，则分裂
28    while(order > min_order){
29        order--;
30        size >>= 1;
31        // 将后半块放入对应阶数的空闲列表
32        struct Page *back = page + size;
33        SetPageProperty(back);
34        back->property = order;
35
36        list_add(&(buddy_free_area[order].free_list), &(back->page_link));
37        buddy_free_area[order].nr_free += size;
38        buddy_nr_free += size;
39    }
40    ClearPageProperty(page); // 分配后清除头部标记
41    page->property = 0;      // 分配后清除阶数信息
42    return page;
43 }
```

释放合并页块

```
1 // 释放n个连续页
2 static void buddy_free_pages(struct Page *base, size_t n){
3     assert(n > 0);
4     // 确定释放块的阶数
5     size_t order = get_order(n);
6     assert(order <= MAX_ORDER);
7     // 设置要释放块的属性
8     struct Page *p = base;
9     for (; p != base + n; p++) {
10         assert(!PageReserved(p) && !PageProperty(p));
11         p->flags = 0; // 清除所有标志
12         set_page_ref(p, 0); // 将页块设置为空闲，可以被分配
13     }
14     // 向上合并
15     struct Page *curr_page = base;
16     while(order < MAX_ORDER){
17         struct Page *buddy = get_buddy_page(curr_page, order);
18         // 检查伙伴块是否空闲且具有相同的阶数
19         if (PageProperty(buddy) && buddy->property == order) {
20             // 从空闲链表中移除伙伴块
21             list_del(&(buddy->page_link));
22             buddy_free_area[order].nr_free -= (1 << order);
23             buddy_nr_free -= (1 << order);
24             // 确定合并后的新块起始页 (始终是地址较低的一个)
25             if (buddy < curr_page) {
26                 curr_page = buddy;
27             }
28             order++;
29         }else{
30             break;
31         }
32     }
33     // 将最终合并后的块插入到对应的空闲列表
34     curr_page->property = order;
35     SetPageProperty(curr_page);
36
37     size_t size = 1 << order;
38     list_add_before(&(buddy_free_area[order].free_list), &(curr_page->page_link));
39     buddy_free_area[order].nr_free += size;
40     buddy_nr_free += size;
41 }
```

5. 测试

测试函数如下:

```
1 static void buddy_check(void) {
2     // 基础检查
3     cprintf("-----Check Begin-----
-----\n");
```

```

4
5 struct Page *p0, *p1, *p2, *p3;
6 size_t total_free_pages = buddy_nr_free_pages();
7 buddy_show();
8 cprintf("\n-----Test 1:分配最小块(1)-----\n");
9
10 // 1. 分配最小块 (1页)
11 assert((p0 = buddy_alloc_pages(1)) != NULL);
12 assert((p1 = buddy_alloc_pages(1)) != NULL);
13 assert((p2 = buddy_alloc_pages(1)) != NULL);
14 assert(buddy_nr_free_pages() == total_free_pages - 3);
15 cprintf("分配3个最小块后...");
16 buddy_show();
17
18 // 2. 释放最小块
19 cprintf("\n-----Test 2:释放最小块(1页)-----\n");
20 buddy_free_pages(p0, 1);
21 buddy_free_pages(p1, 1);
22 buddy_free_pages(p2, 1);
23 assert(buddy_nr_free_pages() == total_free_pages);
24 cprintf("释放3个最小块后...");
25 buddy_show();
26
27 // 3. 检查大块分配和分裂
28 size_t large_size = 50;
29
30 // 分配 50 页
31 cprintf("\n-----Test 1:大块分配和分裂-----\n");
32 p0 = buddy_alloc_pages(large_size);
33 assert(p0 != NULL);
34 assert(buddy_nr_free_pages() == total_free_pages - 64);
35 cprintf("分配p0 50页后...");
36 buddy_show();
37
38 // 分配另一个 50 页块
39 p1 = buddy_alloc_pages(large_size);
40 assert(p1 != NULL);
41 assert(buddy_nr_free_pages() == total_free_pages - 2 * 64);
42 cprintf("分配p1 50页后...");
43 buddy_show();
44
45 // 分配 25 页块
46 p2 = buddy_alloc_pages(25);
47 assert(p2 != NULL);
48 assert(buddy_nr_free_pages() == total_free_pages - 2 * 64 - 32);
49 cprintf("分配p2 25页后...");
50 buddy_show();
51
52 // 4. 检查合并
53 buddy_free_pages(p2, 32);
54 cprintf("释放p2后...");
55 buddy_show();
56

```

```

57     buddy_free_pages(p0, 64);
58     cprintf("释放p0后...");
59     buddy_show();
60
61     buddy_free_pages(p1, 64);
62     cprintf("释放p1后...");
63     buddy_show();
64
65     cprintf("\n-----Test 5:最大单元操作-----\n");
66     p0 = buddy_alloc_pages(16384);
67     cprintf("分配p0 16384页后...");
68     buddy_show();
69
70     buddy_free_pages(p0, 16384);
71     cprintf("释放p0后...");
72     buddy_show();
73
74     cprintf("buddy_check() succeeded!\n");
75 }

```

结果可视化示例如图：

由于测试结果较多，我们使用文本方式呈现测试结果，使用起始 PPN 代指某空闲块。

初始化测试

```

1  -----Check Begin-----
2  Total Free Pages: 16384
3      阶数 0(1 页): EMPTY
4      阶数 1(2 页): EMPTY
5      阶数 2(4 页): EMPTY
6      阶数 3(8 页): EMPTY
7      阶数 4(16 页): EMPTY
8      阶数 5(32 页): EMPTY
9      阶数 6(64 页): EMPTY
10     阶数 7(128 页): EMPTY
11     阶数 8(256 页): EMPTY
12     阶数 9(512 页): EMPTY
13     阶数 10(1024 页): EMPTY
14     阶数 11(2048 页): EMPTY
15     阶数 12(4096 页): EMPTY
16     阶数 13(8192 页): EMPTY
17     阶数 14(16384 页): 空闲 1 块，总空闲 16384 页。内存块起始PPN:
18         [ 0x84000(014) ]
19     -----

```

可以看到初始化后检测结果正确，所有页块加起来与总页数相等。

分配释放最小块测试

```
1  -----Test 1:分配最小块(1)-----
2  分配3个最小块后...Total Free Pages: 16381
3      阶数 0(1 页): 空闲 1 块, 总空闲 1 页. 内存块起始PPN:
4      [ 0x84003(00) ]
5      阶数 1(2 页): EMPTY
6      阶数 2(4 页): 空闲 1 块, 总空闲 4 页. 内存块起始PPN:
7      [ 0x84004(02) ]
8      阶数 3(8 页): 空闲 1 块, 总空闲 8 页. 内存块起始PPN:
9      [ 0x84008(03) ]
10     阶数 4(16 页): 空闲 1 块, 总空闲 16 页. 内存块起始PPN:
11     [ 0x84010(04) ]
12     阶数 5(32 页): 空闲 1 块, 总空闲 32 页. 内存块起始PPN:
13     [ 0x84020(05) ]
14     阶数 6(64 页): 空闲 1 块, 总空闲 64 页. 内存块起始PPN:
15     [ 0x84040(06) ]
16     阶数 7(128 页): 空闲 1 块, 总空闲 128 页. 内存块起始PPN:
17     [ 0x84080(07) ]
18     阶数 8(256 页): 空闲 1 块, 总空闲 256 页. 内存块起始PPN:
19     [ 0x84100(08) ]
20     阶数 9(512 页): 空闲 1 块, 总空闲 512 页. 内存块起始PPN:
21     [ 0x84200(09) ]
22     阶数 10(1024 页): 空闲 1 块, 总空闲 1024 页. 内存块起始PPN:
23     [ 0x84400(010) ]
24     阶数 11(2048 页): 空闲 1 块, 总空闲 2048 页. 内存块起始PPN:
25     [ 0x84800(011) ]
26     阶数 12(4096 页): 空闲 1 块, 总空闲 4096 页. 内存块起始PPN:
27     [ 0x85000(012) ]
28     阶数 13(8192 页): 空闲 1 块, 总空闲 8192 页. 内存块起始PPN:
29     [ 0x86000(013) ]
30     阶数 14(16384 页): EMPTY
```

分配3个最小块结果正确，由于只有一个14阶空闲块，则将其一步一步向下分裂，首先分裂为两个13阶空闲块（0x84000 和 0x86000），前一个接着分裂为两个12阶空闲块（0x84000 和 0x85000），以此类推递归分裂，最后分裂为1阶的两块空闲块（0x84000 和 0x84002）都分裂为0阶空闲块，将前三块分配走，剩下 0x84003 块。

```
1  -----Test 2:释放最小块(1页)-----
2  释放3个最小块后...Total Free Pages: 16384
3      阶数 0(1 页): EMPTY
4      阶数 1(2 页): EMPTY
5      阶数 2(4 页): EMPTY
6      阶数 3(8 页): EMPTY
7      阶数 4(16 页): EMPTY
8      阶数 5(32 页): EMPTY
9      阶数 6(64 页): EMPTY
10     阶数 7(128 页): EMPTY
11     阶数 8(256 页): EMPTY
12     阶数 9(512 页): EMPTY
13     阶数 10(1024 页): EMPTY
14     阶数 11(2048 页): EMPTY
15     阶数 12(4096 页): EMPTY
```

```
16 | 阶数 13(8192 页): EMPTY
17 | 阶数 14(16384 页): 空闲 1 块, 总空闲 16384 页. 内存块起始PPN:
18 | [ 0x84000(014) ]
19 | -----
```

释放3个最小块后结果正确, 重新合并成一大块14阶空闲块。

复杂内存分配和释放测试

```
1 | -----Test 3:大块分配和分裂-----
2 | 分配p0 50页后...Total Free Pages: 16320
3 | 阶数 0(1 页): EMPTY
4 | 阶数 1(2 页): EMPTY
5 | 阶数 2(4 页): EMPTY
6 | 阶数 3(8 页): EMPTY
7 | 阶数 4(16 页): EMPTY
8 | 阶数 5(32 页): EMPTY
9 | 阶数 6(64 页): 空闲 1 块, 总空闲 64 页. 内存块起始PPN:
10 | [ 0x84040(06) ]
11 | 阶数 7(128 页): 空闲 1 块, 总空闲 128 页. 内存块起始PPN:
12 | [ 0x84080(07) ]
13 | 阶数 8(256 页): 空闲 1 块, 总空闲 256 页. 内存块起始PPN:
14 | [ 0x84100(08) ]
15 | 阶数 9(512 页): 空闲 1 块, 总空闲 512 页. 内存块起始PPN:
16 | [ 0x84200(09) ]
17 | 阶数 10(1024 页): 空闲 1 块, 总空闲 1024 页. 内存块起始PPN:
18 | [ 0x84400(010) ]
19 | 阶数 11(2048 页): 空闲 1 块, 总空闲 2048 页. 内存块起始PPN:
20 | [ 0x84800(011) ]
21 | 阶数 12(4096 页): 空闲 1 块, 总空闲 4096 页. 内存块起始PPN:
22 | [ 0x85000(012) ]
23 | 阶数 13(8192 页): 空闲 1 块, 总空闲 8192 页. 内存块起始PPN:
24 | [ 0x86000(013) ]
25 | 阶数 14(16384 页): EMPTY
```

分配 p0 50页后, Buddy System 实际上是分配给 p0 64 页, 过程跟 Test1 类似, 不再赘述, 结果正确。

```
1 | 分配p1 50页后...Total Free Pages: 16256
2 | 阶数 0(1 页): EMPTY
3 | 阶数 1(2 页): EMPTY
4 | 阶数 2(4 页): EMPTY
5 | 阶数 3(8 页): EMPTY
6 | 阶数 4(16 页): EMPTY
7 | 阶数 5(32 页): EMPTY
8 | 阶数 6(64 页): EMPTY
9 | 阶数 7(128 页): 空闲 1 块, 总空闲 128 页. 内存块起始PPN:
10 | [ 0x84080(07) ]
11 | 阶数 8(256 页): 空闲 1 块, 总空闲 256 页. 内存块起始PPN:
12 | [ 0x84100(08) ]
13 | 阶数 9(512 页): 空闲 1 块, 总空闲 512 页. 内存块起始PPN:
14 | [ 0x84200(09) ]
15 | 阶数 10(1024 页): 空闲 1 块, 总空闲 1024 页. 内存块起始PPN:
```



```

16      [ 0x84400(010) ]
17      阶数 11(2048 页): 空闲 1 块, 总空闲 2048 页. 内存块起始PPN:
18      [ 0x84800(011) ]
19      阶数 12(4096 页): 空闲 1 块, 总空闲 4096 页. 内存块起始PPN:
20      [ 0x85000(012) ]
21      阶数 13(8192 页): 空闲 1 块, 总空闲 8192 页. 内存块起始PPN:
22      [ 0x86000(013) ]
23      阶数 14(16384 页): EMPTY

```

再分配 p1 50 页后, 将6阶空闲链表中的剩下那一块取走, 结果正确。

```

1      分配p2 25页后...Total Free Pages: 16224
2      阶数 0(1 页): EMPTY
3      阶数 1(2 页): EMPTY
4      阶数 2(4 页): EMPTY
5      阶数 3(8 页): EMPTY
6      阶数 4(16 页): EMPTY
7      阶数 5(32 页): 空闲 1 块, 总空闲 32 页. 内存块起始PPN:
8      [ 0x840a0(05) ]
9      阶数 6(64 页): 空闲 1 块, 总空闲 64 页. 内存块起始PPN:
10     [ 0x840c0(06) ]
11     阶数 7(128 页): EMPTY
12     阶数 8(256 页): 空闲 1 块, 总空闲 256 页. 内存块起始PPN:
13     [ 0x84100(08) ]
14     阶数 9(512 页): 空闲 1 块, 总空闲 512 页. 内存块起始PPN:
15     [ 0x84200(09) ]
16     阶数 10(1024 页): 空闲 1 块, 总空闲 1024 页. 内存块起始PPN:
17     [ 0x84400(010) ]
18     阶数 11(2048 页): 空闲 1 块, 总空闲 2048 页. 内存块起始PPN:
19     [ 0x84800(011) ]
20     阶数 12(4096 页): 空闲 1 块, 总空闲 4096 页. 内存块起始PPN:
21     [ 0x85000(012) ]
22     阶数 13(8192 页): 空闲 1 块, 总空闲 8192 页. 内存块起始PPN:
23     [ 0x86000(013) ]
24     阶数 14(16384 页): EMPTY

```

接着分配 p2 25页, 实际是分配 32 页, 继续分裂之后分配, 取走分裂出来的前一块。

```

1      -----Test 4:大块释放和合并-----
2      释放p0后...Total Free Pages: 16288
3      阶数 0(1 页): EMPTY
4      阶数 1(2 页): EMPTY
5      阶数 2(4 页): EMPTY
6      阶数 3(8 页): EMPTY
7      阶数 4(16 页): EMPTY
8      阶数 5(32 页): 空闲 1 块, 总空闲 32 页. 内存块起始PPN:
9      [ 0x840a0(05) ]
10     阶数 6(64 页): 空闲 2 块, 总空闲 128 页. 内存块起始PPN:
11     [ 0x840c0(06) 0x84000(06) ]
12     阶数 7(128 页): EMPTY
13     阶数 8(256 页): 空闲 1 块, 总空闲 256 页. 内存块起始PPN:
14     [ 0x84100(08) ]

```

```

15 阶数 9(512 页): 空闲 1 块, 总空闲 512 页. 内存块起始PPN:
16   [ 0x84200(09) ]
17 阶数 10(1024 页): 空闲 1 块, 总空闲 1024 页. 内存块起始PPN:
18   [ 0x84400(010) ]
19 阶数 11(2048 页): 空闲 1 块, 总空闲 2048 页. 内存块起始PPN:
20   [ 0x84800(011) ]
21 阶数 12(4096 页): 空闲 1 块, 总空闲 4096 页. 内存块起始PPN:
22   [ 0x85000(012) ]
23 阶数 13(8192 页): 空闲 1 块, 总空闲 8192 页. 内存块起始PPN:
24   [ 0x86000(013) ]
25 阶数 14(16384 页): EMPTY

```

释放 p0 后, 将分配给 p0 的页块 (0x84000) 重新插入对应阶数 (6阶) 的空闲链表中, 此时6阶空闲链表中有两块空闲块, 但它们并不相邻, 所以不执行合并。

```

1  释放p1后...Total Free Pages: 16352
2  阶数 0(1 页): EMPTY
3  阶数 1(2 页): EMPTY
4  阶数 2(4 页): EMPTY
5  阶数 3(8 页): EMPTY
6  阶数 4(16 页): EMPTY
7  阶数 5(32 页): 空闲 1 块, 总空闲 32 页. 内存块起始PPN:
8   [ 0x840a0(05) ]
9  阶数 6(64 页): 空闲 1 块, 总空闲 64 页. 内存块起始PPN:
10  [ 0x840c0(06) ]
11 阶数 7(128 页): 空闲 1 块, 总空闲 128 页. 内存块起始PPN:
12  [ 0x84000(07) ]
13 阶数 8(256 页): 空闲 1 块, 总空闲 256 页. 内存块起始PPN:
14  [ 0x84100(08) ]
15 阶数 9(512 页): 空闲 1 块, 总空闲 512 页. 内存块起始PPN:
16  [ 0x84200(09) ]
17 阶数 10(1024 页): 空闲 1 块, 总空闲 1024 页. 内存块起始PPN:
18  [ 0x84400(010) ]
19 阶数 11(2048 页): 空闲 1 块, 总空闲 2048 页. 内存块起始PPN:
20  [ 0x84800(011) ]
21 阶数 12(4096 页): 空闲 1 块, 总空闲 4096 页. 内存块起始PPN:
22  [ 0x85000(012) ]
23 阶数 13(8192 页): 空闲 1 块, 总空闲 8192 页. 内存块起始PPN:
24  [ 0x86000(013) ]
25 阶数 14(16384 页): EMPTY

```

释放 p1 后, 将分配给 p1 的页块 (0x84040) 重新插入对应阶数 (6阶) 的空闲链表中, 此时6阶空闲链表中有两块相邻的页块, 执行合并操作, 合并为一块7阶空闲块 (0x84000)。

```

1  释放p2后...Total Free Pages: 16384
2  阶数 0(1 页): EMPTY
3  阶数 1(2 页): EMPTY
4  阶数 2(4 页): EMPTY
5  阶数 3(8 页): EMPTY
6  阶数 4(16 页): EMPTY
7  阶数 5(32 页): EMPTY
8  阶数 6(64 页): EMPTY

```

```

9      阶数 7(128 页): EMPTY
10     阶数 8(256 页): EMPTY
11     阶数 9(512 页): EMPTY
12     阶数 10(1024 页): EMPTY
13     阶数 11(2048 页): EMPTY
14     阶数 12(4096 页): EMPTY
15     阶数 13(8192 页): EMPTY
16     阶数 14(16384 页): 空闲 1 块, 总空闲 16384 页. 内存块起始PPN:
17     [ 0x84000(014) ]

```

最后释放 p2, 依次执行合并操作, 最终恢复为一块14阶空闲块。

分配释放最大块测试

```

1      -----Test 5:最大单元操作-----
2      分配p0 16384页后...Total Free Pages: 0
3      阶数 0(1 页): EMPTY
4      阶数 1(2 页): EMPTY
5      阶数 2(4 页): EMPTY
6      阶数 3(8 页): EMPTY
7      阶数 4(16 页): EMPTY
8      阶数 5(32 页): EMPTY
9      阶数 6(64 页): EMPTY
10     阶数 7(128 页): EMPTY
11     阶数 8(256 页): EMPTY
12     阶数 9(512 页): EMPTY
13     阶数 10(1024 页): EMPTY
14     阶数 11(2048 页): EMPTY
15     阶数 12(4096 页): EMPTY
16     阶数 13(8192 页): EMPTY
17     阶数 14(16384 页): EMPTY
18     -----
19     释放p0后...Total Free Pages: 16384
20     ... (此处省略)
21     阶数 14(16384 页): 空闲 1 块, 总空闲 16384 页. 内存块起始PPN:
22     [ 0x84000(014) ]
23

```

最大块分配与释放结果也都正确。

扩展练习Challenge: 任意大小的内存单元slub分配算法（需要编程）

slub算法, 实现两层架构的高效内存单元分配, 第一层是基于页大小的内存分配, 第二层是在第一层基础上实现基于任意大小的内存分配。可简化实现, 能够体现其主体思想即可。

- 参考[linux的slub分配算法/](#), 在ucore中实现slub分配算法。要求有比较充分的测试用例说明实现的正确性, 需要有设计文档。

SLUB 分配算法的核心思想

SLUB (The SLAB allocator) 是一种面向小对象的内存分配算法。它的核心思想是：

在传统页分配算法（如 First-Fit）的基础上，构建一层“对象缓存（Cache）”机制。每个页被划分为多个固定大小的对象，从而让小对象的分配与释放更加快速、减少碎片、增强缓存局部性，并且在多核环境下减少锁竞争。

非常好，这一版代码的设计逻辑已经完全符合 SLUB 算法的核心思想和实验要求，可以直接作为实验报告中“Challenge2：任意大小内存单元的 SLUB 分配算法实现”的分析对象。下面是结合操作系统原理与代码的详细分析说明。

总体思路

该实现遵循 两层内存管理架构：

- 1. 页级层 (Page Level)
底层仍依赖 uCore 的页分配器（如 default_pmm_manager 或 best_fit_pmm_manager），负责分配整页。
- 2. 对象层 (Object Level)
在页级基础上实现小对象分配器 SLUB。每个页被划分为多个固定大小的对象（object），由 slab 管理。这样在频繁分配小对象（如内核结构体）时，可避免内存碎片和频繁页分配带来的性能损耗。

数据结构设计

1. Slab 结构体

```
1 typedef struct slab {
2     struct Page* page;
3     unsigned char bitmap[OBJ_PER_SLAB];
4     int free_cnt;
5     list_entry_t slab_link;
6 } slab;
```

- 每个 slab 对应一页物理内存。
- `bitmap` 记录对象是否被占用（1 表示已分配）。
- `free_cnt` 表示剩余空闲对象数量。
- `slab_link` 用于将多个 slab 连接入 cache 的链表。

2. Cache 结构体

```
1 typedef struct cache {
2     size_t obj_size;
3     list_entry_t slabs_full;
4     list_entry_t slabs_partial;
5     list_entry_t slabs_free;
6 } cache;
```

每个 cache 管理一种固定大小对象的所有 slab，并通过三个链表分类：

- `slabs_free`：全空闲；
- `slabs_partial`：部分占用；

- `slabs_full`: 已满。

这种三态管理模型直接继承自 Linux 内核的 SLUB 思想，使查找和回收更高效。

核心函数逻辑分析

1. Cache 初始化

```
1 static void cache_init(Cache* cache, size_t obj_size) {
2     cache->obj_size = obj_size;
3     list_init(&cache->slabs_full);
4     list_init(&cache->slabs_partial);
5     list_init(&cache->slabs_free);
6 }
```

完成 cache 的基础结构初始化，并为后续 slab 管理做好链表准备。

2. 创建 Slab

```
1 static Slab* slab_create(void) {
2     struct Page* page = alloc_page();
3     if (page == NULL) return NULL;
4
5     Slab* slab = (Slab*)page2kva(page);
6     memset(slab->bitmap, 0, sizeof(slab->bitmap));
7     slab->page = page;
8     slab->free_cnt = OBJ_PER_SLAB;
9
10    list_add(&kmem_cache.slabs_free, &(slab->slab_link));
11    return slab;
12 }
```

- 通过底层页分配器获得物理页；
- 在页内初始化 Slab 控制结构与位图；
- 链接入 `slabs_free` 列表。

这一步相当于“预热内存池”，为后续对象分配提供资源基础。

3. 对象分配

```
1 static void* cache_alloc(Cache* cache) {
2     Slab* slab = NULL;
3     if (!list_empty(&cache->slabs_partial)) {
4         slab = le2slab(list_next(&cache->slabs_partial), slab_link);
5     } else if (!list_empty(&cache->slabs_free)) {
6         slab = le2slab(list_next(&cache->slabs_free), slab_link);
7         list_del(&(slab->slab_link));
8         list_add(&cache->slabs_partial, &(slab->slab_link));
9     } else {
```

```

10     slab = slab_create();
11     list_add(&cache->slabs_partial, &(slab->slab_link));
12 }
13 ...
14 }

```

- 优先在 **部分占用 slab** 中寻找空位；
- 若无部分占用 slab，则从 **空 slab** 提取；
- 若仍无空 slab，则创建新 slab；
- 分配对象后，若该 slab 被占满，则迁移至 `slabs_full`。

这种逻辑兼顾了时间效率（避免遍历所有 slab）与空间利用率。

4. 对象释放

```

1  static void cache_free(Cache* cache, void* obj) {
2      // 遍历 FULL / PARTIAL 两类列表
3      ...
4      if (slab->free_cnt == OBJ_PER_SLAB) {
5          list_del(&(slab->slab_link));
6          list_add(&cache->slabs_free, &(slab->slab_link));
7      }
8  }

```

- 找到对象所在 slab；
- 清空 bitmap 对应位置；
- 若 slab 完全空闲，则转入 `slabs_free`；
- 若释放无效对象，系统打印警告。

检查函数与验证流程

以上就是我们根据Slub的主要思想设计的相应算法，下面是测试的流程。`slob_check()` 模拟真实使用场景，对算法正确性进行验证：

```

1  static void slob_check(void) {
2      void* objs[100];
3      ...
4      // ① 批量分配对象
5      // ② 部分释放（1/3）
6      // ③ 再次分配以测试复用
7      // ④ 全部释放
8      // ⑤ 异常释放检测
9  }

```

输出如下结构：

```
1 [Check1] Allocated 50 objects successfully.
2 [Check2] Released 1/3 objects, testing partial reuse...
3 [Check3] Additional allocations complete.
4 [Check4] All objects freed, slabs recycled correctly.
5 [Check5] Exception handling verified.
6 ===== SLUB Check PASSED =====
```

验证内容包括：

- 内存是否复用；
- 是否正确回收；
- 是否能检测非法释放；
- 各阶段对象数量是否匹配预期。

测试结果如图：

说明成功通过测试。

实验结果分析

- 所有分配/释放操作均输出 `[SLUB]` 调试信息；
- 检查函数完整通过；
- 系统未发生断言错误；
- 物理内存状态与虚拟页映射保持一致。

说明该 SLUB 算法实现逻辑正确、行为稳定。

可改进方向

1. 多 Cache 支持：

支持多种对象大小（如 32B、64B、128B）缓存，提高通用性。

2. 跨页 Slab：

允许 slab 跨越多页，用于分配较大对象。

3. NUMA 感知优化：

实现多 CPU 结构下的本地 cache，加快访问速度。

4. 延迟释放策略：

结合 LRU 队列，减少频繁的页释放与分配。

扩展练习Challenge：硬件的可用物理内存范围的获取方法（思考题）

- 如果 OS 无法提前知道当前硬件的可用物理内存范围，请问你有什么办法让 OS 获取可用物理内存范围？

1. 逐块物理探测法（Memory Probing）

一种直接但有效的方式是——从低地址开始逐页访问物理内存，验证哪些地址可读写。实现上可以按照如下逻辑：

```
1  for (addr = 0x100000; addr < MAX_MEM; addr += PAGE_SIZE) {
2      volatile uint64_t *ptr = (uint64_t *)addr;
3      uint64_t old = *ptr;
4      *ptr = 0xAA55AA55AA55AA55;
5      if (*ptr != 0xAA55AA55AA55AA55)
6          break; // 访问失败，说明超出物理内存范围
7      *ptr = old; // 恢复原数据
8  }
```

当访问无效内存时，CPU 会产生 **总线错误（Bus Error）** 或 **页错误（Page Fault）**，内核捕获该异常，即可推断最大可访问物理地址。

优点： 不依赖外部描述信息；实现简单，适合嵌入式或裸机环境。

缺点： 访问越界风险高，需结合异常处理机制以防系统崩溃；部分内存区域可能被设备映射，导致误判。

2. DMA 间接探测法（Using DMA Controller）

通过配置 **DMA 控制器** 让外设尝试将数据写入不同的物理地址区间。如果 DMA 传输在某个地址失败（例如返回超时或错误码），则说明该地址并未映射到有效内存。

该方法具有两大特点：

- 不依赖 CPU 直接访问内存；
- 能避免 CPU 触发保护异常。

典型流程如下：

1. 选择安全的 DMA 设备（如网卡或磁盘控制器）；
2. 从低地址开始逐段设置 DMA 目标地址；
3. 检测 DMA 完成状态寄存器，判断该段是否有效。

优点： 安全性高，不会触发 CPU 级异常。

缺点： 实现复杂，需硬件层支持；探测精度受设备 DMA 步长限制。

3. 利用固件接口（Firmware Table）

在现代系统中，更常见的是通过固件提供的接口自动获取内存布局信息。例如：

- **BIOS 模式：** 调用中断 `INT 0x15`，`EAX=0xE820` 获取内存映射表；
- **UEFI 模式：** 调用 `GetMemoryMap()` 接口；
- **RISC-V / ARM 平台：** 解析设备树（`.dtb` 文件）中 `/memory` 节点。

该方法由固件在引导时提供精确的内存段描述，包括：

- 可用（RAM）；
- 保留（reserved）；
- 设备内存（device memory）；
- ACPI 或 MMIO 区域。

内核只需解析这些结构即可完成物理内存探测。

例如 uCore 中常见的 `dtb.c` 文件中就通过 `dtb_scan_memory()` 函数解析 `/memory` 节点，从而得到 `mem_start` 与 `mem_end`。

优点： 安全、准确、通用性强；不依赖试探操作。

缺点： 需要系统固件支持，在极简或裸机系统中可能不可行。

4. 启发式混合探测法（Hybrid Approach）

在某些轻量系统中，可以将逐页探测与固件描述相结合：

- 优先读取设备树/BIOS 提供的初始范围；
- 若信息缺失，则采用逐块测试扩展边界；
- 最后利用 DMA 或 I/O 设备验证可访问性。

这种方法既可提升安全性，又能保证通用性。例如在嵌入式 RISC-V SoC 中，Bootloader 可能只报告部分 DRAM 区域，而剩余部分可通过探测法进一步确认。

综合比较

方法	精度	安全性	实现复杂度	适用场景
逐块物理探测	中	低（需异常保护）	简单	裸机 / 无固件系统
DMA 间接探测	中	高	较复杂	嵌入式系统
固件接口解析	高	高	简单	通用现代硬件
混合探测法	高	高	中	异构或自研硬件平台

总结

若 OS 无法提前获知硬件的物理内存范围，可通过**多层机制协同探测**实现自适应内存识别：

- 优先解析固件/设备树提供的内存表；
- 必要时启用主动探测机制，逐页写读或 DMA 验证；
- 结合异常捕获机制确保系统安全；
- 最终在内核页分配初始化（如 `pmm_init`）阶段形成 `memmap` 表，用于后续物理页管理。