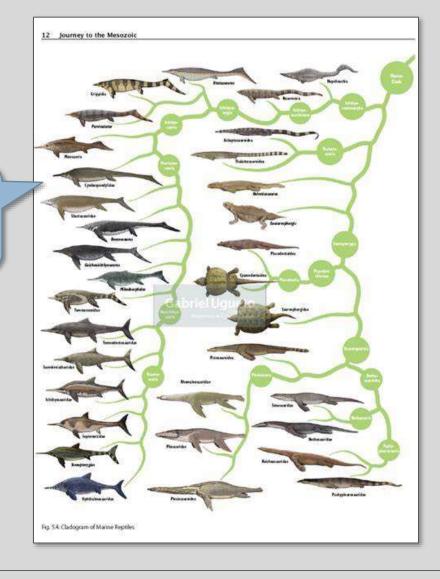


模块化教程: 项目目标

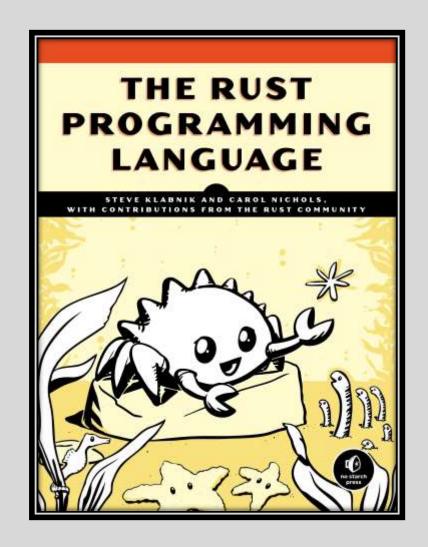
- 。当前版本设计
 - 。 从发展史过渡到功能扩充, 学习曲线平滑
 - 。 章节之间差异小但分散,容易造成困扰
 - 。 每章一个分支, 迁移管理困难
- 。优化设计
 - 。保持各章功能不变
 - 。 章节之间差异集中化,共性分离为 lib crate
 - 。 每章一个 bin crate, 在同一个分支里方便管理
- 。补充设计
 - 。 利用 qemu 和 rust 的跨平台特性,支持 windows 学习
 - 。 支持细节优化,方便学生自主学习

中生代 海爬 演化树



大纲·TOC

- 。项目目标
- 。模块设计
- 。章节实现介绍
 - 。第一章:裸机应用程序
 - 。 第一章实验:依赖注入
 - 。 第二章: 单道批处理系统
 - 。 控制流切换设计
 - 。 系统调用分发
 - 。 第三章: 多道批处理系统
 - 。 中断响应
 - 。 第四章: 地址空间
 - 。 内存分配: 可扩展伙伴分配器
 - 。 内核堆: 分配器级联
 - 。 内核地址空间:游标式页表访问



[[bin]]

[[lib]]

ch_

复用

user

接口

xtask

实现

模块设计

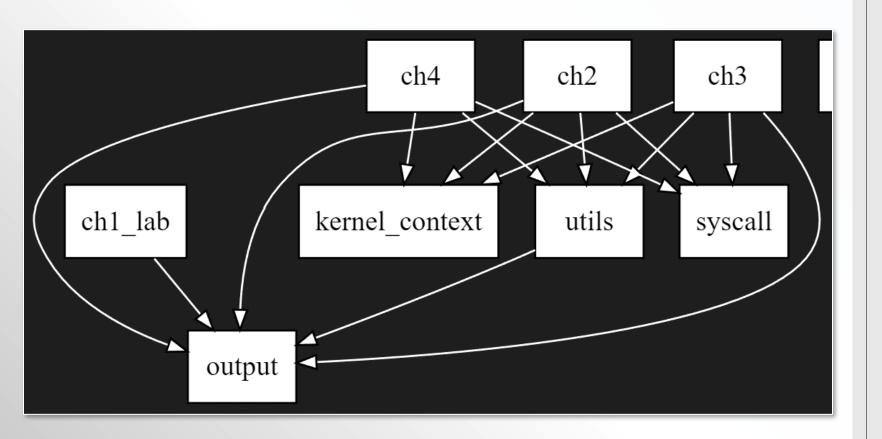
微内核≠模块化

动态加载≠模块化

它们是正交的

lib crates

7.46	संस्थान			1*** 2**					
lib	来源	接口	实现	描述					
output		\checkmark	$\overline{\checkmark}$	提供print!/println!/log!					
syscall	内部模块	\checkmark		定义并分发系统调用					
kernel-context	内叫铁坑			上下文切换					
utils				静态加载和其他杂项					
<mark>sbi-rt</mark>				提供特权二进制接口调用					
<mark>page-table</mark>	依赖库	\checkmark		页式内存管理软硬协同					
buddy-allocator		\checkmark		伙伴分配器 (重写)					
spin	第三方库			自旋锁					
log	第二 刀件			日志					



模块依赖关系

使用 rust-analyzer 绘制

裸机

环境自举

特权态

需要SEE

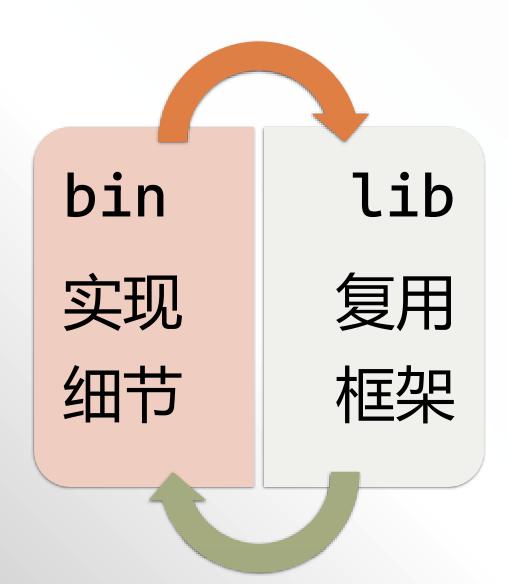
应用程序

有实际功能

SBI

第一章

特权态裸机应用程序 保持简单性



第一章实验

依赖注入

output crate

提供	要求					
print! println!						
Log::Log 的一种实现和初始化	京坝 Cancala 并供) \$大司田					
Console trait	实现 Console 并传入静态引用					
传入用户实现的函数						

```
/// 将传给 `output` 的控制台对象。
///

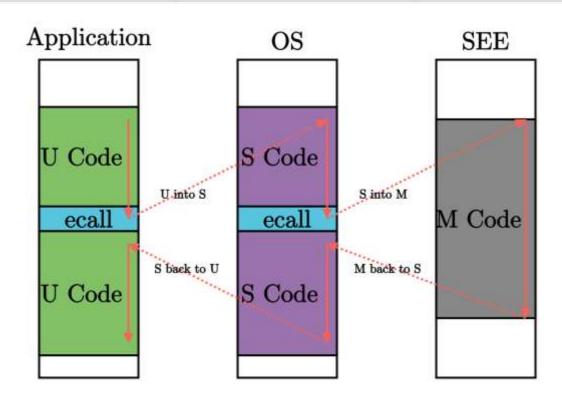
/// 这是一个 Unit struct, 它不需要空间。否则需要传一个 static 对象。
1 implementation | 2 references
struct Console;

/// 为 `Console` 实现 `output::Console` trait。
impl output::Console for Console {
    6 references
    ··· fn put_char(&self, c: u8) {
    ··· #[allow(deprecated)]
    ··· legacy::console_putchar(c as _);
}
```

```
// 初始化 console You, 3周前 · docs: 补充-
init_console(&Console);
// 设置总的日志级别
log::set_max_level(log::LevelFilter::Trace);

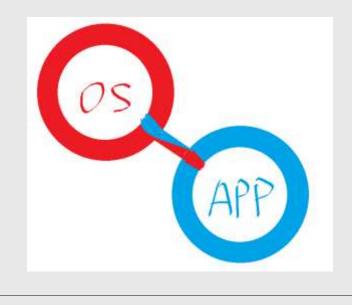
println!("[PRINT] Hello, world!");
log::trace!("Hello, world!");
log::debug!("Hello, world!");
log::info!("Hello, world!");
log::warn!("Hello, world!");
log::error!("Hello, world!");
```

第一人称视角的设计 机器态 特权态 用户态 主要控制流 调度是函数 系统调用是函数 SBI 调用是函数 系统调用是函数



第二章

线程和特权级切换 提供系统调用



内核设计: 线程切换

目标	设计						
高内聚 (形成模块)							
低耦合 (调用简洁)	线程切换封装为函数						
零开销							
可扩展	支持用户态/内核态控制流						
四17 茂	支持中断开启/屏蔽						

```
        地址空间切換

        内核固定(双页表) 集成为执行器

        内核可变(单页表) 分散
```

```
/// 线程上下文。
#[repr(C)]
1 implementation | 14 references
pub struct LocalContext {
    sctx: usize,
    x: [usize; 31],
    sepc: usize,
    /// 是否以特权态切换。
    pub supervisor: bool,
    /// 线程中断是否开启。
    pub interrupt: bool,
}
```

```
safe extern "C" fn execute_naked()
          option push
          option pic
         add1 5p, 5p, -32*8
         SAVE_ALL
         csrw stvec, to
         LOAD_ALL
     "1: csrrw sp, sscratch, sp",
         ld sp; (sp)",
         LOAD_ALL
          option pop",
     options (noreturn)
```

异界传送门

只用于内核固定式内核, 封装了隔离页表任务的切换执行

用一段汇编实现了:切换地址空间→执行任务→回到原来的地址空间

不需要两次完整的通用寄存器切换, 而是一次通用寄存器切换和一次跨 地址空间执行(即"异界传送")。 只有传送门对象需要在地址空间公 共位置,暴露更少

跨地址空间任务也不需要内核栈, 共享代码(一段可重定位汇编)和 数据都在传送门对象内部。只需要 正常的任务上下文加描述地址空间 的页表

```
#[cfg(not(feature = "coop"))]
sbi_rt::set_timer(stime_value: time::read64() + 12500);
unsafe { tcb.execute() };
use scause::*;
let finish: bool = match scause::read().cause() {
```

第三章

多道加载 任务状态扩充 系统调用扩充

虚地址空间									
应用程序地址空间	应用程序动态加载	ELF 解析							
硬件机制	动态内存分配	页表管理							
rtht	控制流/地址空间同时切换								
内核软件	内核堆								

内存分配

可扩展伙伴分配器

分配器级联

页表管理

页式内存管理

游标式访问

第四章

地址空间

内核设计: 内存分配

目标 设计 内核堆动态分配 分配器级联 (内存池) 连续页/大页分配 伙伴分配器 (弃用位图分配器,分离伙伴行和寡头行 trait) 提高性能

```
/// 伙伴分配器的一个行。
3 implementations | 8 references pub trait BuddyLine {
    /// 空集台。用于静态初始化。
    Const EMPTY: Self;

    /// 侵入式元数据的大小。
    const INTRUSIVE_META_SIZE: usize = 0;

    /// 伙件分配器可能需要集台知道自己的阶数和基序号。
    #[inline]
    5 references
    fn init(&mut self, _order: usize, _base: usize) {}

    /// 提取指定位置的元素,返回是否提取到。
    #[inline]
    3 references
    fn take(&mut self, _idx: usize) → bool {
        unimplemented!()
    }
}
```

伙伴行 (原理)

layer	order	接口	内存块															
4	N+4	寡头	0															
3	N+3	伙伴	0							1								
2	N+2	伙伴	0				1				2				3			
1	N+1	伙伴	e	0 1		2 3		3	4		5		6		7			
Θ	N	伙伴	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

- 伙伴分配器是内存块的二叉森林
- 每个层表示的内存范围是一样的
- 层和地址的关系: 每层表示一个位
- 层高对性能影响很大,尤其是大块分配
- 根的数量是无所谓的
- 根和枝操作不同
- 改进: 最高层使用不同的接口 (寡头行)

- 伙伴行的操作
 - 1. 取出一个块
 - 2. 放回一个块
 - 如果伙伴存在则两个一起取出
 - 否则放回这个块
- 寡头行的操作
 - 1. 取出连续的多个块
 - 2. 放回块 (不需要找伙伴)

伙伴行(位图)

```
You, 4天前 | 1 author (You)
      /// 用一个 usize 作为位图保存占用情况的伙伴行。
     /// - 非侵入式
     /// - 静态分配,容量有限
      5 implementations |
      pub struct UsizeBuddy {
          bits: usize,
          base: usize,
 12
 13 > impl UsizeBuddy { ...
 29 > impl BuddyLine for UsizeBuddy {...
 45 > impl OligarchyCollection for UsizeBuddy { ...
 83 > impl BuddyCollection for UsizeBuddy { ...
122 > impl fmt::Debug for UsizeBuddy [...
```

- 节点数静态有限
 - 出于性能考虑最好一个 usize, 最多一个 u128
 - 如果限制在一个整数里则性能极高
- 非侵入式,不需要寻址到内存块内部
- 实现了所有接口但非常适合寡头
- 注意:通用位图实际上是 N 层块表示 2^K 个 N-1 层块的扩展的伙伴分配器,其操作逻辑是一致的。但实际上只有位图才能实现比较有效率的发现多个连续伙伴块的存在,所以把伙伴分配器扩展出来是没意义的。而且一旦每层不能放在一个整数里,性能也高不到哪去,要找的连续块越多就会产生越多查找失败,性能就越差。想减少分配失败只能减少层间块分裂的数量,其极限就是基本的伙伴分配器。

伙伴行 (链表)

- 每行块数量是无限的
- 侵入式,需要修改所管理内存块的内容
- 最小的块是一个链表结点的大小 (一个指针)
- 提取一个任意的块可以直接第一个,很快
- 由于总是按顺序取,产生的碎片更少
- 释放块时需要找伙伴,等价于插入排序,需要遍历链表

- 优化方案:

- 针对一个伙伴行的优化不影响伙伴分配器本体
- 平衡二叉树 → 最稳定的存取开销
- 小顶堆 → 取很快, 存也不慢
- 其他更复杂、针对应用的优化

```
/// 侵入式链表伙伴行。
4 implementations | 9 references
pub struct LinkedListBuddy {
    free_list: Node,
    order: Order,
impl BuddyLine for LinkedListBuddy {
impl OligarchyCollection for LinkedListBuddy {...
impl BuddyCollection for LinkedListBuddy { ...
impl fmt::Debug for LinkedListBuddy {
You, 上周 | 1 author (You)
#[repr(transparent)]
1 implementation | 9 references
struct Node {
    next: Option<NonNull<Node>>> ,
impl Node {
```

伙伴分配器

方法

类型

```
mpl<const N: usize, O: OligarchyCollection, B: BuddyCollection> BuddyAllocator<N, O. B>
  const O_MIN_ORDER: usize = O::INTRUSIVE_META_SIZE_next_power_of_two() trailing_zeros
  const 8_MIN_ORDER: usize = 8::INTRUSIVE_META_SIZE.next_power_of_two().trailing_zeros
  #[inline]
  #[inline]
  #[inline]
  pub fn free(Sself) → usize (
  #[inline]
  const fn max_order(&self) → usize {
  m inline
  pub fn init<T>(&mut self, min_order: usize, base: NonNull<T>) [
  # inline
  pub fn allocate_type<T>(&nut self) → Result<(NonNull<T>, usize), BudoyError> (--
  pub fn allocate<T>(
impl BuddyAllocator<N, O, B>
```

```
let mut layer: usize = layer0;
let mut idx: usize = loop {

// 从真头情

if layer == Self::MAX_LAYER {

match self.oligarchy.take_any(aligner) => break idx,

None => Err(BuddyError)?,

}

// 从伙伴情

match self.buddies[layer].take_any(aligner) => break idx,

None => layer += 1,

}

you, #/KM feat: | 2 Markete
```

核心分配算法

核心释放算法

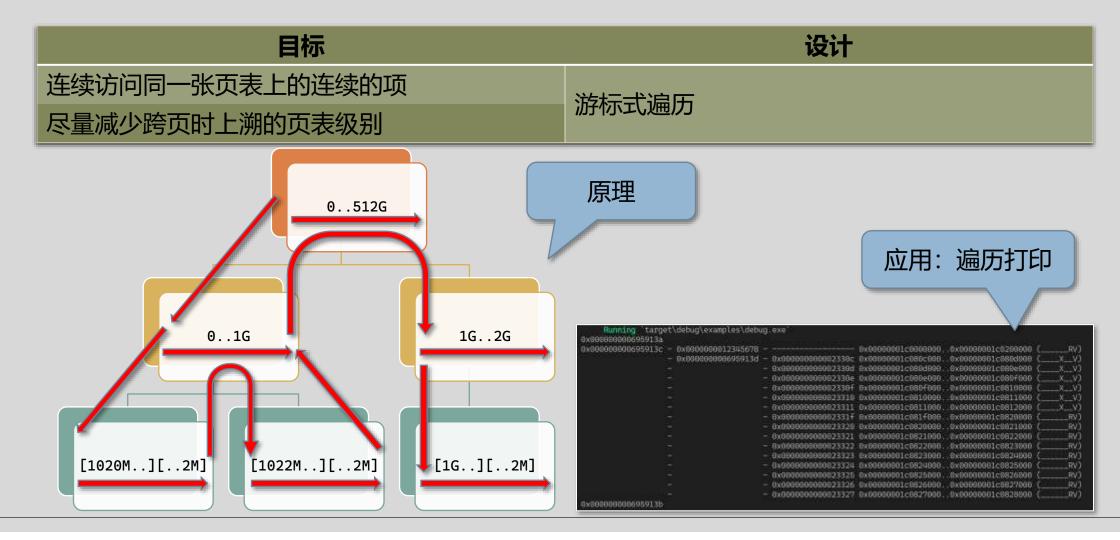
```
for layer: usize in (order - self.min_order).. {

    // 你故其人
    if layer == Self::MAX_LAYER {
        self.oligarchy.put(idx);
        break;

    // 你故代作
    match self.buddies[layer].put(idx) {
        Some(parent: usize) ⇒ idx = parent,
        None ⇒ break,

}
```

内核设计: 页表访问



页表游标:接口

用户设定游标目标,库移动游标

游标命中目标或无法移动时请求用户操作

支持只读访问或边访问边修改

```
#[derive(Clone, Copy)]

4 implementations | 28 references
pub struct Pos<Meta: VmMeta> {
    /// 目标页表项包含的一个虚页号。
    pub vpn: VPN<Meta
    /// 追历中断时的更新方案。
    /// 目标页表项的级别
    pub level: usize,
}

/// 修改目标。
1 reference
    Target(Pos<Meta>),
    /// 新建中间页表。
1 reference
Pte(Pte<Meta>, VPN<Meta>),
}
```

```
trait Visitor<Meta: VmMeta> {
   /// 出发时调用一次以设置第一个目标。
   fn start(&mut self,
    pos: Pos<Meta>
   ) → Pos<Meta>;
   fn arrive(&mut self,
            pte: Pte<Meta>,
            target_hint: Pos<Meta>
   ) → Pos<Meta>;
   fn meet(&mut self,
          level: usize,
          pte: Pte<Meta>,
          target_hint: Pos<Meta>
   ) → Pos<Meta>;
trait Visitor
```

```
ub trait Decorator<Meta: VmMeta> {
   /// 出发时调用一次以设置第一个目标。
   /// 'pos' 是页表上最高级别的第一个页的位置。
   1 reference
   fn start(&mut self,
   pos: Pos<Meta>
   ) → Pos<Meta>;
   1 reference
   fn arrive(&mut self,
            pte: &mut Pte<Meta>,
     target_hint: Pos<Meta>
   ) → Pos<Meta>;
   fn meet(&mut self,
          level: usize,
          pte: Pte<Meta>, You, 16
         target_hint: Pos<Meta>
   ) → Update<Meta>;
trait Decorator
```

页表游标: 遍历打印页表

```
struct FmtVisitor<'f1, 'f2, Meta: VmMeta> {
   f: & f1 mut fmt::Formatter< f2>,
   max_level: usize,
   new_line: bool,
   _phantom: PhantomData<Meta>,
impl<'f1, 'f2, Meta: VmMeta> Visitor<Meta> for FmtVisitor<'f1, 'f2, Meta> {
   #[inline]
   fn start(&mut self, pos: Pos<Meta>) → Pos<Meta> {
       self.max_level = pos.level;
       DOS
   fn arrive(&mut self, pte: Pte<Meta>, target_hint: Pos<Meta>) → Pos<Meta</pre>
   #[inline]
   fn meet(&mut self, _level: usize, _pte: Pte<Meta>, _target_hint: Pos<Me
impl Visitor for FmtVisitor<Meta>
```

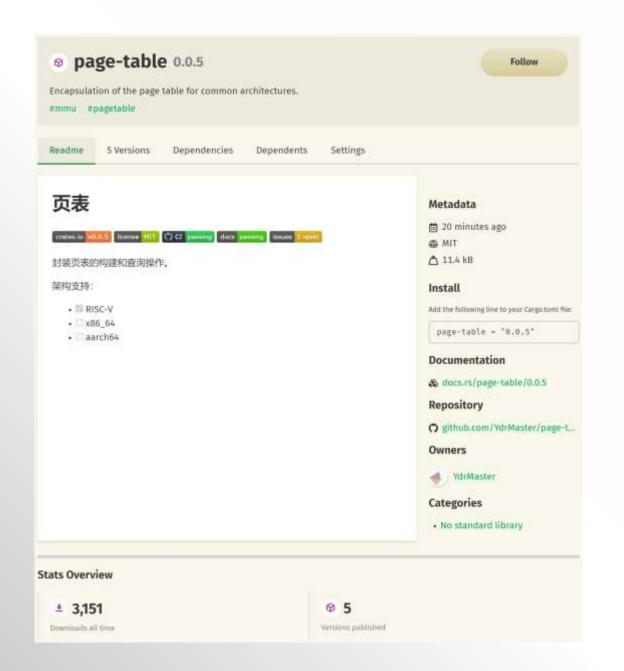
```
println!(
    "{:?}",
    PageTableShuttle {
        table: root,
        f: |ppn| VPN::new(ppn.val())
    }
); You, 现在 Uncommitted chang
```

页表游标: 内核虚存映射

impl Decorator for KernelSpaceBuilder

```
mpl Decorator Sv39> for KernelSpaceBuilder (
  #[inline]
  fn start(Gmut self, : Pos<Sv39>) → Pos<Sv39> [
     Pos::new(vpn: VAddr::new(_text as usize).floor(), level:0)
  # inline
  fn arrive(@nut self, pte: @nut Pte<5v39>, target_hint: Pos<5v39>) → Pos<5v39> 
     let addr: usize = target_hint.vpn.base().val();
     Let bits usize = IF addr < _transit m usize {
         051911 // × RV ← .text
     ] else if addr < __rodata as usize {
         961111 // XWSV ← trumpling
     } else if addr < __data as usize {
         8b9811 // _RV ← irodata
     | also if addr < __end == usize |
         8b8111 // WHY ← data → bas
         return Possistop(); // end of kernel sections
     *pte = unsafe { VeFlags::from_raw(bits) }.build_pte(ppm:PPN::new(target_hint.vpn,val()));
     target hint next()
```

```
let mut shuttle: PageTableShuttle<Sv<3>, |...| → ...> = PageTableShuttle {
   table,
   f: |ppn| VPN::new(ppn.val()),
};
shuttle.walk_mut(KernelSpaceBuilder);
println!("{shuttle:?}");
unsafe { satp::set(satp::Mode::Sv39, asid:0, ppn:kernel_root.floor().val()) };
```



欢迎试用!

#![deny(missing_docs)]

