MDCC 2015

2015 移动开发者大会

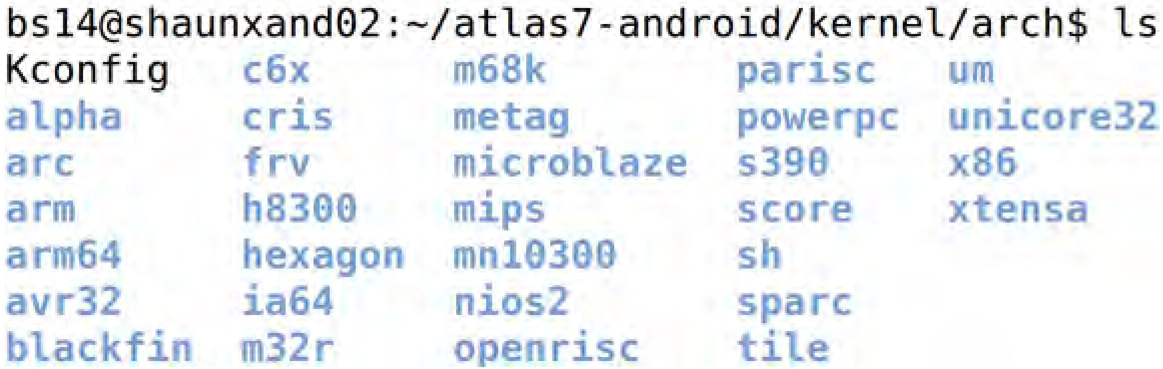
Mobile Developer Conference China 2015

Linux驱动的架构思考

宋宝华 <baohua@kernel.org>

Linux支持的体系架构

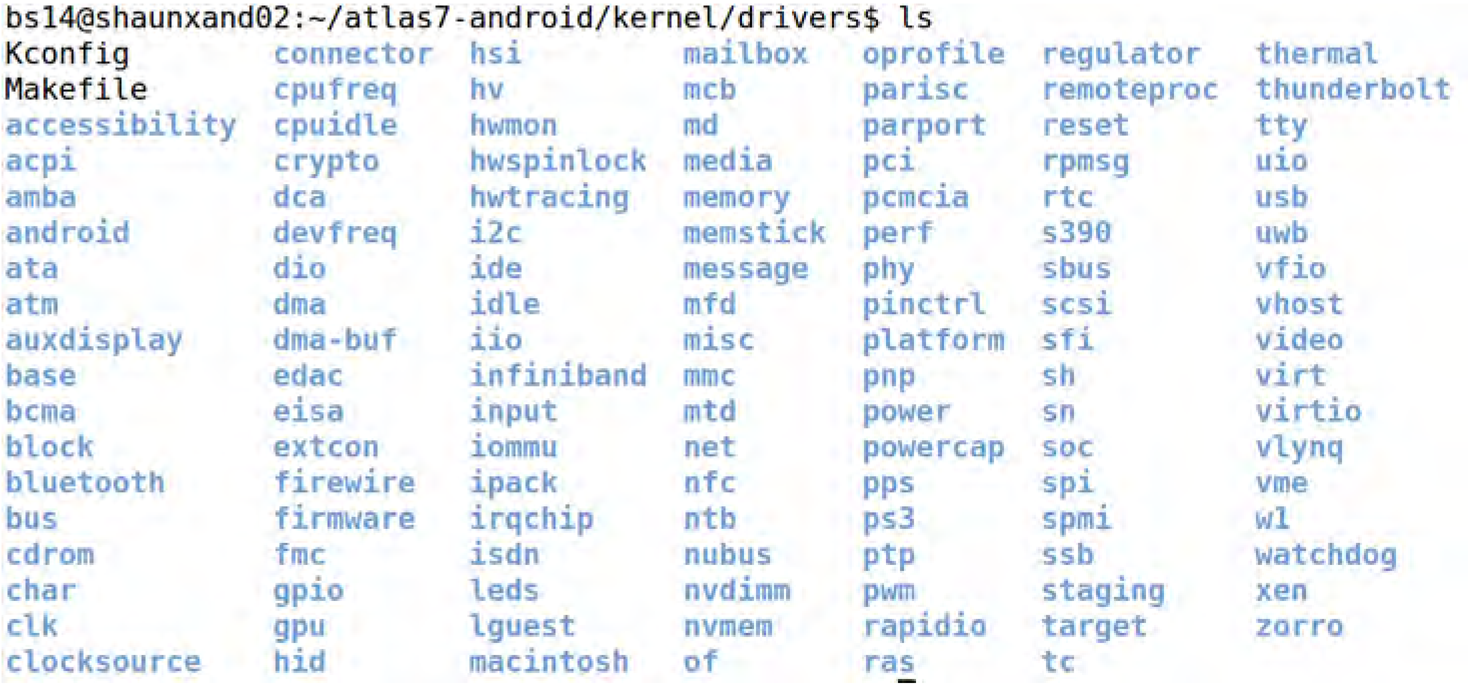
约30种体系架构



```
bs14@shaunxand02:~/atlas7-android/kernel/arch$ ls
Kconfig   c6x       m68k        parisc   um
alpha     cris      metag       powerpc  unicore32
arc       frv       microblaze  s390     x86
arm       h8300     mips        score    xtensa
arm64     hexagon   mn10300     sh
avr32     ia64      nios2       sparc
blackfin  m32r      openrisc    tile
```

Linux的驱动子系统

约130种驱动



```
bs14@shaunxand02:~/atlas7-android/kernel/drivers$ ls
Kconfig        connector  hsi         mailbox   oprofile  regulator   thermal
Makefile       cpufreq    hv          mcb       parisc    remoteproc  thunderbolt
accessibility  cpuidle    hwmon       md        parport   reset       tty
acpi           crypto     hwspinlock  media     pci       rpmsg       uio
amba           dca        hwtracing   memory    pcmcia    rtc         usb
android        devfreq    i2c         memstick  perf      s390        uwb
ata            dio        ide         message   phy       sbus        vfio
atm            dma        idle        mfd       pinctrl   scsi        vhost
auxdisplay     dma-buf    iio         misc      platform  sfi         video
base           edac       infiniband  mmc       pnp       sh          virt
bcma           eisa       input       mtd       power     sn          virtio
block          extcon     iommu       net       powercap  soc         vlynq
bluetooth      firewire   ipack       nfc       pps       spi         vme
bus            firmware   irqchip     ntb       ps3       spmi        w1
cdrom          fmc        isdn        nubus     ptp       ssb         watchdog
char           gpio       leds        nvdimm    pwm       staging     xen
clk            gpu        lguest      nvmem     rapidio   target      zorro
clocksource    hid        macintosh   of        ras       tc
```

Linux需要一个好的体系结构

- 这么多CPU，这么多外设。如果BSP不好移植，Linux不可能在嵌入式如此成功；
- 那么目标是什么？

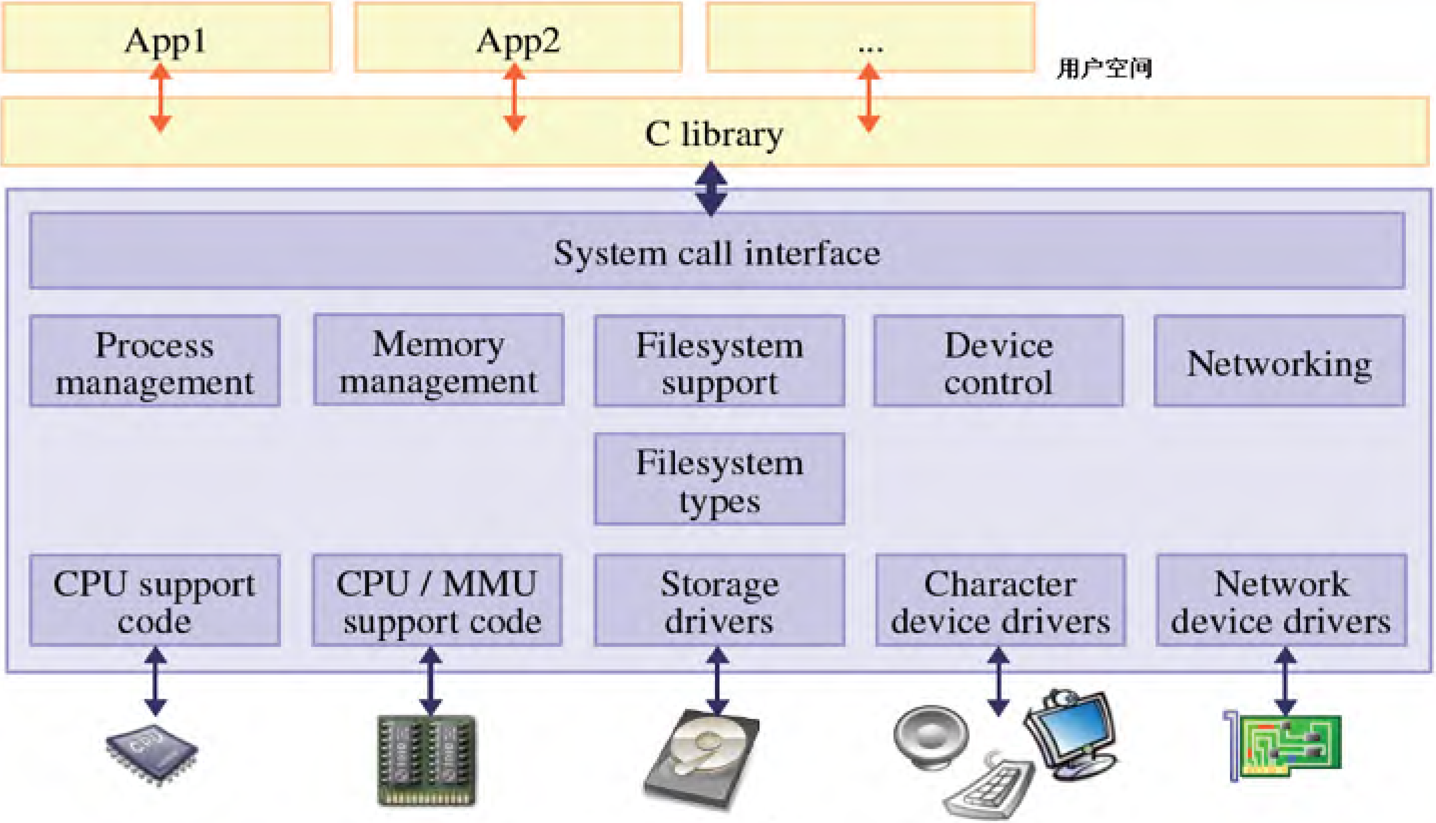
最小化移植Linux到一个新的体系架构、SoC的工作；

内核核心部分（调度器、内存管理、设备访问等）应该通用；

外设、IP驱动应该跨平台；

驱动核心层应该尽可能做更多的通用工作。

Linux软件栈

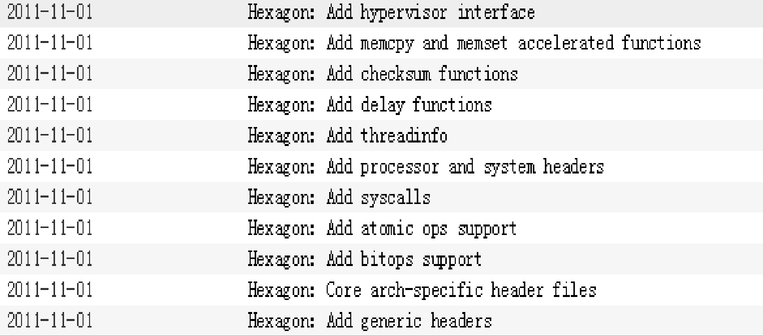

App1
App2
...
用户空间
C library
System call interface
Process management
Memory management
Filesystem support
Device control
Networking
Filesystem types
CPU support code
CPU / MMU support code
Storage drivers
Character device drivers
Network device drivers

Linux支持一个新的处理器架构需要多少行代码？

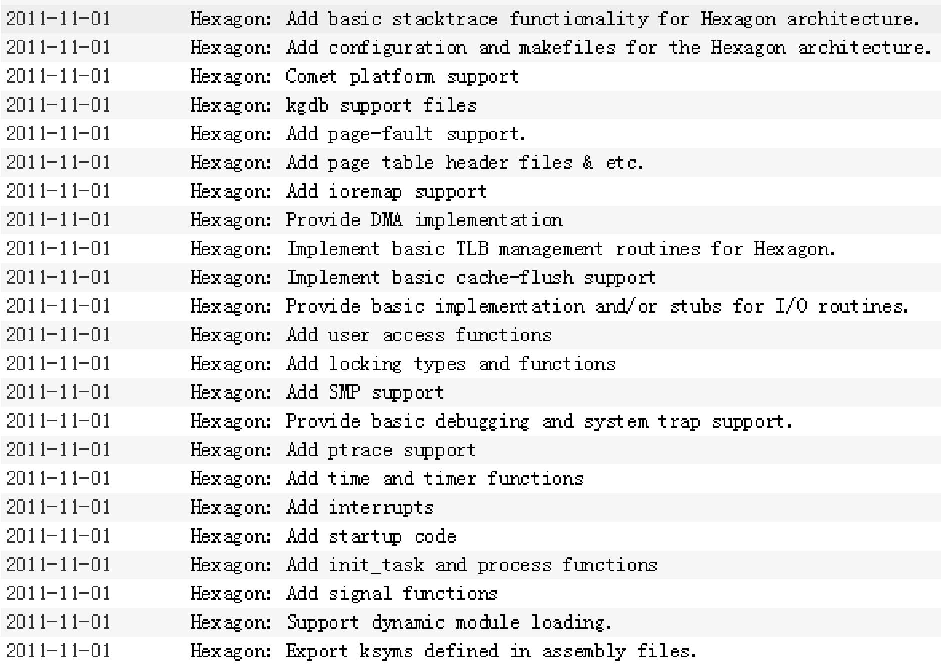
```
▪   kernel/arch/c6x$ find ./ -name "*.[c|h|S]" | xargs wc -l -- | grep total
10345 total
▪   kernel/arch/hexagon$ find ./ -name "*.[c|h|S]" | xargs wc -l -- | grep total
12045 total
▪   kernel/arch/openrisc$ find ./ -name "*.[c|h|S]" | xargs wc -l -- | grep total
  9433 total
```

一个新的处理器架构

mm, ipc, kernel, fs, net等是通用的，新的处理器架构需要对其底层进行支持。



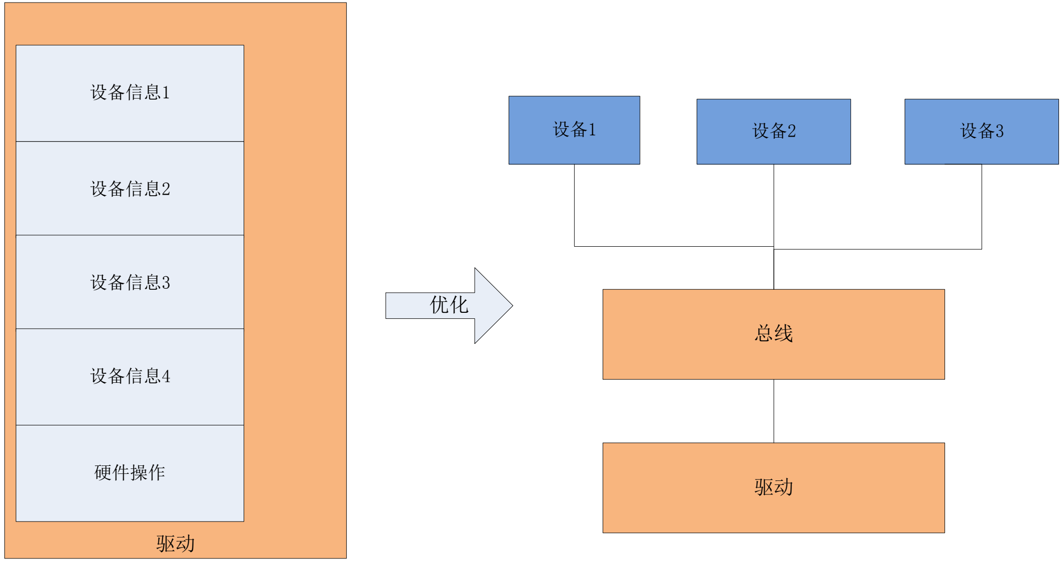
2011-11-01	Hexagon: Add hypervisor interface
2011-11-01	Hexagon: Add memcpy and memset accelerated functions
2011-11-01	Hexagon: Add checksum functions
2011-11-01	Hexagon: Add delay functions
2011-11-01	Hexagon: Add threadinfo
2011-11-01	Hexagon: Add processor and system headers
2011-11-01	Hexagon: Add syscalls
2011-11-01	Hexagon: Add atomic ops support
2011-11-01	Hexagon: Add bitops support
2011-11-01	Hexagon: Core arch-specific header files
2011-11-01	Hexagon: Add generic headers



2011-11-01	Hexagon: Add basic stacktrace functionality for Hexagon architecture.
2011-11-01	Hexagon: Add configuration and makefiles for the Hexagon architecture.
2011-11-01	Hexagon: Comet platform support
2011-11-01	Hexagon: kgdb support files
2011-11-01	Hexagon: Add page-fault support.
2011-11-01	Hexagon: Add page table header files & etc.
2011-11-01	Hexagon: Add ioremap support
2011-11-01	Hexagon: Provide DMA implementation
2011-11-01	Hexagon: Implement basic TLB management routines for Hexagon.
2011-11-01	Hexagon: Implement basic cache-flush support
2011-11-01	Hexagon: Provide basic implementation and/or stubs for I/O routines.
2011-11-01	Hexagon: Add user access functions
2011-11-01	Hexagon: Add locking types and functions
2011-11-01	Hexagon: Add SMP support
2011-11-01	Hexagon: Provide basic debugging and system trap support.
2011-11-01	Hexagon: Add ptrace support
2011-11-01	Hexagon: Add time and timer functions
2011-11-01	Hexagon: Add interrupts
2011-11-01	Hexagon: Add startup code
2011-11-01	Hexagon: Add init_task and process functions
2011-11-01	Hexagon: Add signal functions
2011-11-01	Hexagon: Support dynamic module loading.
2011-11-01	Hexagon: Export ksyms defined in assembly files.

总线、设备、驱动模型

驱动只管驱动，设备只管设备，总线则负责匹配设备和驱动，而驱动则以标准途径拿到板级信息，这样，驱动就可以放之四海而皆准了



关于设备: Linux ARM的过去

- kernel/arch/arm$ git reset --hard v2.6.39

```
find ./ -name "*.[c|h|S]" | xargs wc -l -- | grep total
699157 total
```

大量平台code：

```
static struct resource i2c_resource1[] = {
        [0] = {
                .start  = INT_I2C,
                .end    = INT_I2C,
                .flags  = IORESOURCE_IRQ,
        },
        [1] = {
                .start  = TEGRA_I2C_BASE,
                .end    = TEGRA_I2C_BASE + TEGRA_I2C_SIZE-1,
                .flags  = IORESOURCE_MEM,
        },
};
```

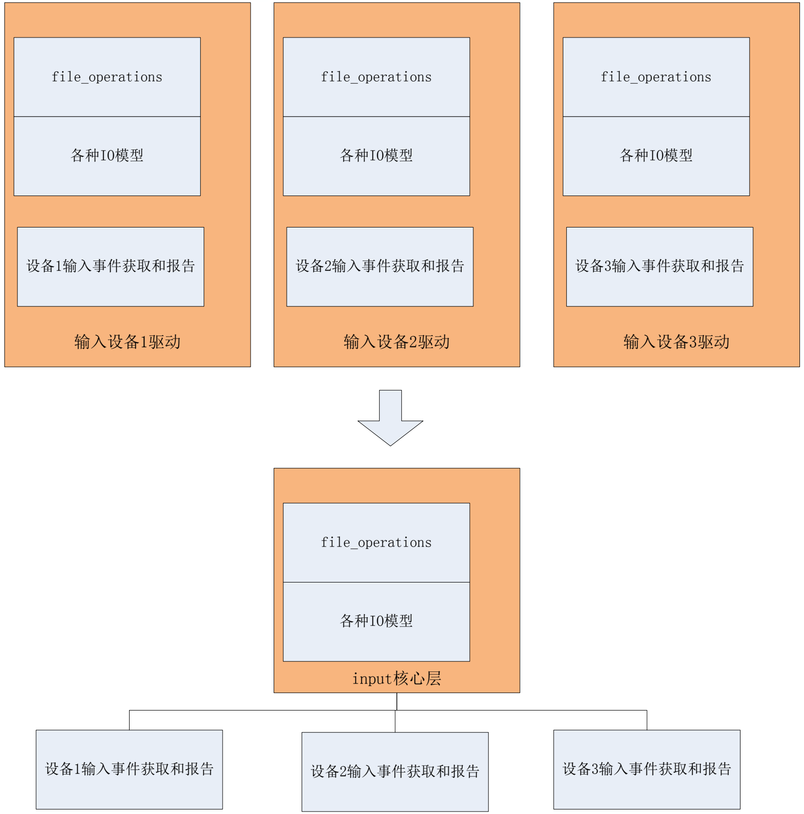
```
struct platform_device tegra_i2c_device1 = {
        .name           = "tegra-i2c",
        .id             = 0,
        .resource       = i2c_resource1,
        .num_resources  = ARRAY_SIZE(i2c_resource1),
        .dev = {
                .platform_data = 0,
        },
};
```

关于设备: Linux ARM的现在

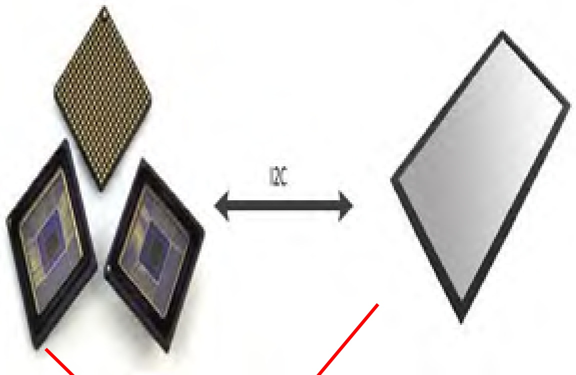
-
```
kernel/arch/arm$ git reset --hard v3.14
find ./ -name "*.[c|h|S]" | xargs wc -l -- | grep total
 592959 total
```
- 但是3.14支持的SoC远大于2.6.39, 以dts描述硬件信息
```
gic: interrupt-controller@10301000 {
        compatible = "arm,cortex-a9-gic";
        interrupt-controller;
        #interrupt-cells = <3>;
        reg = <0x10301000 0x1000>,
                  <0x10302000 0x0100>;
};
```
- arch/arm/mach-xxx的一些timer，gpio，pinmux，clock等变为drivers/

提炼驱动核心层

将软件进行分层设计：比如提炼一个input的核心层出来，把跟Linux接口的事情以及整个一套input事件的buffer机制都在这里面实现掉，**底层只管中断和报事件**

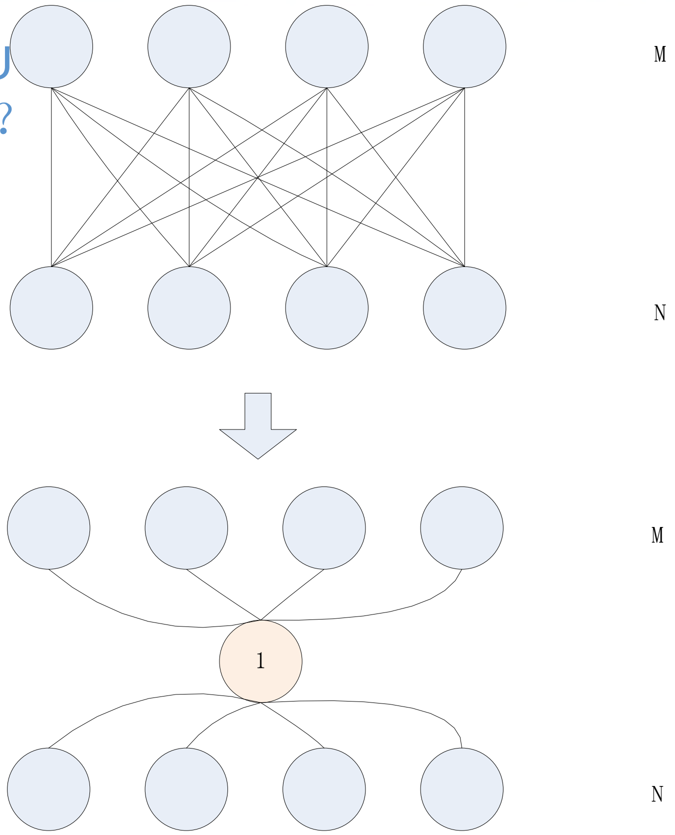


总线上接了一个外设，怎么去耦合？



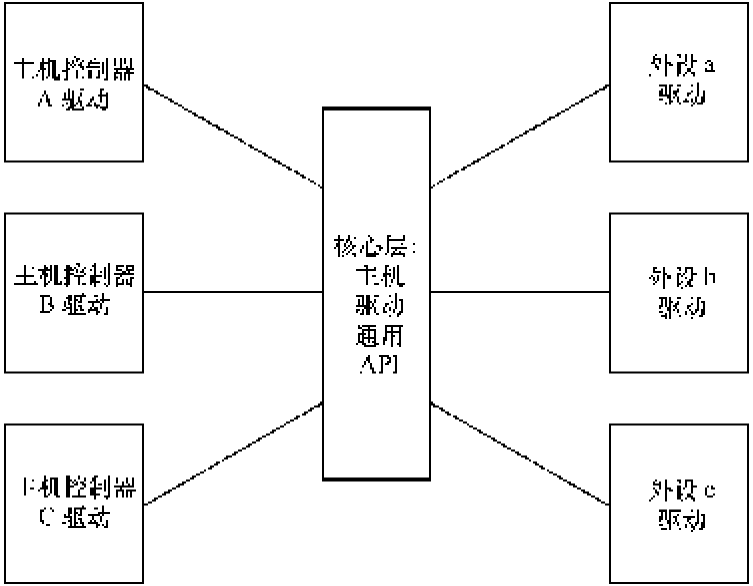
~~cpu_xxx_i2c_reg_write()~~
~~cpu_xxx_i2c_reg_read()~~
~~tp_yyy_reg1_handle()~~
~~cpu_xxx_i2c_reg_write()~~
~~cpu_xxx_i2c_reg_read()~~
~~tp_yyy_reg2_handle()~~
…

假设M个 CPU
接N个触摸屏？

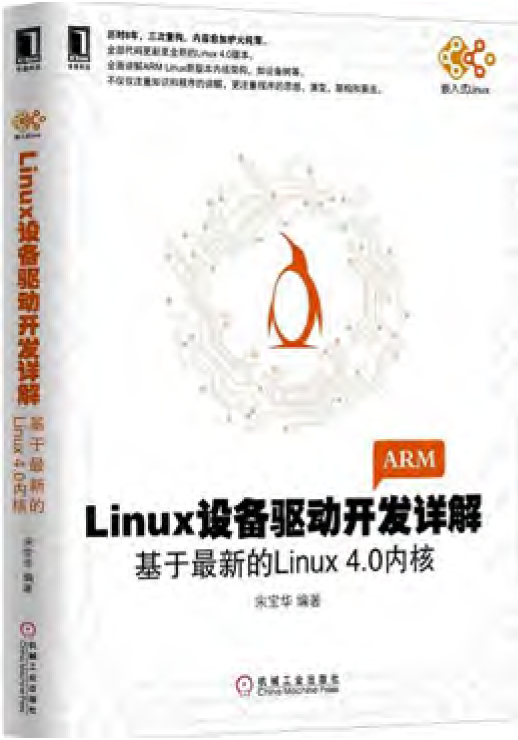


主机驱动与外设驱动分离

- 外设只是访问核心层的通用的API进行数据传输，主机和外设之间可以进行任意的组合



参考资料


ARM
Linux设备驱动开发详解
基于最新的Linux 4.0内核
宋宝华 编著
嵌入式Linux

LINUX DEVICE DRIVERS
第三版
LINUX
设备驱动程序
O'REILLY®
中国电力出版社
JONATHAN CORBET, ALESSANDRO RUBINI
& GREG KROAH-HARTMAN 著
魏永明 耿岳 钟书毅 译

Thank
you!