Linux设备模型浅析之驱动篇

本文属本人原创,欢迎转载,转载请注明出处。由于个人的见识和能力有限,不可能面面俱到,也可能存在谬误,敬请网友指出,本人的邮箱是 yzq.seen@gmail.com,博客是http://zhiqiang0071.cublog.cn。

Linux 设备模型,仅仅看理论介绍,比如 LDD3 的第十四章,会感觉太抽象不易理解,而通过阅读内核代码就更具体更易理解,所以结合理论介绍和内核代码阅读能够更快速的理解掌握 linux 设备模型。这一序列的文章的目的就是在于此,看这些文章之前最好能够仔细阅读 LDD3 的第十四章。大部分 device 和 driver 都被包含在一个特定 bus 中,platform_device 和 platform_driver 就是如此,包含在 platform_bus_type 中。这里就以对 platform_bus_type 的调用为主线,浅析 platform_driver 的注册过程,从而理解 linux 设备模型。platform_bus_type 用于关联 SOC 的 platform device 和 platform driver,比如在内核 linux-2.6.29 中所有 S3C2410 中的 platform device 都保存在 devs.c 中。这里就以 S3C2410 RTC 的驱动程序 rtc-s3c.c 为例来分析 platform_driver_register()例程的调用过程。在文章的最后贴有一张针对本例的 device model 图片,可在阅读本文章的时候作为参照。阅读这篇文章之前,最好先阅读文章《Linux 设备模型 浅析之设备篇》。

```
一、S3C2410 RTC 的 platform driver 定义在 drivers/rtc/rtc-s3c.c 中,代码如下:
static struct platform driver s3c2410 rtc driver = {
      .probe
                  = s3c rtc probe,
                  = devexit p(s3c rtc remove),
      .remove
      .suspend
                  = s3c rtc suspend,
                  = s3c rtc resume,
      .resume
      .driver
            .name = "s3c2410-rtc",
            .owner = THIS MODULE,
      },
};
由于 name = "s3c2410-rtc",后面会分析到,它将成为一个目录的名字,
即/sys/bus/platform/s3c2410-rtc。s3c rtc probe()的调用后面也会讲到,其实在之前《Linux 设备
模型浅析之设备篇》中已经分析过了。使用 platform driver register()注册 s3c2410 rtc driver,
下面就分析它。
    二、platform driver register()例程定义在 drivers/base/ platform.c 中,代码如下:
int platform driver register(struct platform driver *drv)
      drv->driver.bus = &platform bus type; // 设置为 platform bus type
      if (drv->probe)
            drv->driver.probe = platform drv probe; // 转存到 device driver 中
      if (drv->remove)
            dry->driver.remove = platform dry remove;
      if (drv->shutdown)
            drv->driver.shutdown = platform drv shutdown;
      if (drv->suspend)
            dry->driver.suspend = platform dry suspend;
      if (drv->resume)
```

dry->driver.resume = platform dry resume; return driver register(&drv->driver); // 该例程将 device driver 注册到 bus,后面将分析 代码中,

1. 设置成 platform bus type 这个很重要,因为 driver 和 device 是通过 bus 联系在一起的,具 体在本例中是通过 platform bus type 中注册的回调例程和属性来是实现的,在《Linux 设备模型 浅析之设备篇》中讲到的 driver 与 device 的匹配就是通过 platform bus type 注册的回到例程 mach()来完成的。关于 platform bus type 的分析请参考《Linux 设备模型浅析之设备篇》。下面 就分析 driver register()。

```
三、driver register()例程定义在 drivers/base/ driver.c 中,代码如下:
   int driver register(struct device driver *drv)
      int ret;
      struct device driver *other;
      // 做些判断
      if ((drv->bus->probe && drv->probe) ||
        (drv->bus->remove && drv->remove) ||
        (drv->bus->shutdown && drv->shutdown))
            printk(KERN WARNING "Driver '%s' needs updating - please use "
                  "bus type methods\n", drv->name);
      /*通过驱动的名字查找 driver,如果找到了,说明已经注册过,返回错误代码,后面会
      分析*/
      other = driver find(drv->name, drv->bus);
      if (other) {
            put driver(other);
            printk(KERN ERR "Error: Driver '%s' is already registered, "
                  "aborting...\n", drv->name);
            return -EEXIST;
      }
      /* 将 driver 加入到 bus 的 kset, 并生成些文件夹和链接文件, 后面会分析 */
      ret = bus add driver(drv);
      if (ret)
            return ret;
      /* 添加 attribute group, 本例中没有设置 drv->groups */
      ret = driver add groups(drv, drv->groups);
      if (ret)
            bus remove driver(drv);
      return ret;
代码中,
   1. 正如该例程的英文注释所言,大部分工作在 bus add driver()例程中完成。
```

bus_add_driver()例程的作用类似于《Linux 设备模型浅析之设备篇》分析过的 bus add device(), 只不过后者将 device 添得到 bus 中。下面分析 driver find()例程。

```
2. driver find()例程通过驱动所属 bus 的 driver 容器 drivers kset 来查找,该例程定义在
drivers/base/ driver.c 中,代码如下:
struct device driver *driver find(const char *name, struct bus type *bus)
      struct kobject *k = kset find obj(bus->p->drivers kset, name); // 在 drivers kset 容器中查找
      struct driver private *priv;
      if (k) {
            priv = to driver(k);
            return priv->driver; // 返回找到的 driver
      return NULL;
代码中,
   2.1. 通过 kset find obj(bus->p->drivers kset, name)查找该 driver 的 kobj, 其代码如下,
struct kobject *kset find obj(struct kset *kset, const char *name)
      struct kobject *k;
      struct kobject *ret = NULL;
      spin lock(&kset->list lock);
      list for each entry(k, &kset->list, entry) { // 遍历 kset->list 列表获取 kobj
            if (kobject_name(k) &&!strcmp(kobject_name(k), name)) { // 比较 name 字符
                   ret = kobject get(k); // 如果找到就增加引用并返回
                   break;
      spin unlock(&kset->list lock);
      return ret;
显然,所有同类型的 driver 都注册到了一个 bus->p->drivers kset->list 中,所以可通过其查找已
经注册的 driver。下面分析 bus add driver()例程。
    3. bus add driver()例程将 driver 注册到 bus 中,在本例中是 platform bus type,该例程定
义在 drivers/base/ bus.c 中,代码如下:
int bus add driver(struct device driver *drv)
{
      struct bus type *bus;
      struct driver private *priv;
      int error = 0;
      bus = bus get(drv->bus);
                              // 增加对 bus 的引用
      if (!bus)
            return -EINVAL;
      pr debug("bus: '%s': add driver %s\n", bus->name, drv->name);
```

```
priv = kzalloc(sizeof(*priv), GFP KERNEL); // 这个结构体中存放着 kobj 相关的数据
if (!priv) {
      error = -ENOMEM;
      goto out put bus;
klist init(&priv->klist devices, NULL, NULL);
priv->driver = drv; // 反向指向包含其的 drv, 以便后续使用
drv -> p = priv:
                 // 将 priv 保存到 device driver
/* 指向 bus 的 drivers kset 容器,该容器的作用与 device kset 容器相同,前者是包含所
有注册到该 bus 的 driver,后者是包含所有注册到该 bus 的 device。*/
priv->kobj.kset = bus->p->drivers kset;
/* 初始化 priv->kobj,并将其添加到 bus->p->drivers kset 中,在本例中生
成/sys/bus/platform/drivers/s3c2410-rtc 目录,后面会分析 drivers kset 的初始化
及/sys/bus/platform/drivers/目录的生成 */
error = kobject init and add(&priv->kobj, &driver ktype, NULL,
                    "%s", drv->name);
if (error)
      goto out unregister;
if (drv->bus->p->drivers autoprobe) { // 在 bus register()例程中已经设置为1了
     error = driver attach(drv); // 所以会寻找匹配的 device, 后面分析
     if (error)
            goto out unregister;
}
// 将 driver 链接到 klist drivers,方便后续快速查找 driver
klist add tail(&priv->knode bus, &bus->p->klist drivers):
// 将 driver 添加到 module 模块,后面会分析
module add driver(drv->owner, drv);
/* 生成/sys//sys/bus/platform/drivers/s3c2410-rtc/uevent 属性文件,其作用与 device 中的
uevent 类似,可参照《Linux 设备模型浅析之设备篇》*/
error = driver create file(dry, &driver attr uevent);
if (error) {
     printk(KERN ERR "%s: uevent attr (%s) failed\n",
             func , drv->name);
/* 添加 bus 的公有属性文件到/sys//sys/bus/platform/drivers/s3c2410-rtc/目录,所有的 driver
都添加 */
error = driver add attrs(bus, drv);
if (error) {
     /* How the hell do we get out of this pickle? Give up */
     printk(KERN ERR "%s: driver add attrs(%s) failed\n",
            func , drv->name);
}
```

```
/* 如果配置了"CONFIG HOTPLUG",则生成"bind"和"unbind"属性文件,可用于手动匹
     配和移除 device 与 driver 之间的关联 */
     error = add bind files(drv);
     if (error) {
           /* Ditto */
           printk(KERN ERR "%s: add bind files(%s) failed\n",
                 func , drv->name);
     }
     // 通过 uevent 设置几个环境变量并通知用户空间,以便调用程序来完成相关设置
     kobject uevent(&priv->kobj, KOBJ ADD);
     return error:
out unregister:
     kobject put(&priv->kobj);
out_put_bus:
     bus put(bus);
     return error;
代码中,
   3.1. 老一点的 2.6 的内核版本如 2.6.10 直接将 kobj 嵌入在 struct device driver 中,新版的内
核用 struct driver private 来存放 kobj 相关的数据,将 struct driver private 嵌入在 struct
device driver 中。struct driver private 定义如下:
struct driver private {
     struct kobject kobj; // kobj
                             // 用于链接关联到该 driver 的 device
     struct klist klist devices;
     struct klist node knode bus; // 用于链接到 bus->p->klist drivers
     struct module kobject *mkobj;
                                 // 模块的 kobi, 后面会讲到 module
     struct device driver *driver; // 反向指向包含其的 driver
}
    3.2. 使用 bus register()例程注册 bus(本例是 platform bus type)的时候,除了生成该 bus
的 kset 容器 subsys, 还生成 devices kset 和 drivers kset 容器,都包含在 struct bus type private 里
。当然也先后生成他们的目录/sys/bus/platform、/sys/bus/platform/devices
、/sys/bus/platform/drivers。先看看 struct bus type private 的定义:
struct bus type private {
                       // 代表该 bus, 里面的 kobj 是该 bus 的主 kobj, 也就是最顶层
     struct kset subsys:
     struct kset *drivers kset;
                             // 包含该 bus 所有的 driver
     struct kset *devices kset;
                            // 包含该 bus 所有的 device
     struct klist klist devices;
                             // 其作用与 devices kset->list 作用相同
                             // 其作用与 drivers kset->list 作用相同
     struct klist klist drivers;
     struct blocking notifier head bus notifier; // 通知 bus 相关的模块
     unsigned int drivers autoprobe:1; // 设置是否在 driver 注册的时候 probe device
     struct bus type *bus; // 回指包含自己的 bus
其实该结构体的英文注释写的很清楚,可到内核的 drivers/base/ base.h 中查看。
    3.3. bus_register()例程定义在 drivers/base/bus.c 中 ,部分代码如下:
int bus register(struct bus type *bus)
```

```
{
      int retval;
      struct bus type private *priv;
      priv = kzalloc(sizeof(struct bus type private), GFP KERNEL);
      if (!priv)
            return -ENOMEM;
                        // 反向指向 bus
      priv->bus = bus;
      bus->p = priv;
      // 将 bus 的名字设置为 kobj 的名字,本例中是"platform"
      retval = kobject_set_name(&priv->subsys.kobj, "%s", bus->name);
      if (retval)
            goto out;
      priv->subsys.kobj.kset = bus kset; // 指向其父kset, bus kset 在 buses init()例程中添加
      priv->subsys.kobj.ktype = &bus ktype;
                                         // 设置读取属性文件的默认方法
      priv->drivers autoprobe = 1;
                                   // 设置为注册时 probe device
      /* 注册 priv->subsys 容器,初试化完后,调用了 kobject add internal()注册其 kobj 到
      bus kset 父容器里,并生成目录/sys/bus/platfrom */
      retval = kset register(&priv->subsys):
      /* 生成 devices kset 容器,命名为"devices",由于其父 kobj 是 priv->subsys.kobj,所以生
      成的目录是/sys/bus/platform/devices */
      priv->devices kset = kset create and add("devices", NULL,
                                     &priv->subsys.kobj);
      /* 生成 drivers kset 容器,命名为"drivers",由于其父 kobj 是 priv->subsys.kobj,所以生
      成的目录是/sys/bus/platform/ drivers */
      priv->drivers kset = kset create and add("drivers", NULL,
                                     &priv->subsys.kobj);
下面该分析 driver attach()例程了。
    3.4. driver attach()例程从 bus 中查找匹配的 device, 该例程定义在 drivers/base/ dd.c 中,代
码如下:
int driver attach(struct device driver *drv)
      /* 遍历 bus 的 klist devices 列表,对每个 device 使用回调函数 driver attach()来鉴别是否
      和 driver 匹配 */
      return bus for each dev(dry->bus, NULL, dry, driver attach);
代码中,
    3.4. 1. bus for each dev()定义在在 drivers/base/ bus.c 中,代码如下:
int bus for each dev(struct bus type *bus, struct device *start,
```

```
void *data, int (*fn)(struct device *, void *))
{
      struct klist iter i;
      struct device *dev;
      int error = 0;
      if (!bus)
             return -EINVAL;
      // 设置 i
      klist iter init node(&bus->p->klist devices, &i,
                      (start? &start->knode bus: NULL));
      // 使用 i 遍历
      while ((dev = next device(\&i)) \&\& !error)
             error = fn(dev, data); // 使用回调例程处理
      klist iter exit(&i);
      return error;
接着分析回调例程 driver attach()。
    3.4. 2. 回调例程 driver attach()定义在 drivers/base/ dd.c 中,代码如下:
static int driver attach(struct device *dev, void *data)
      struct device driver *drv = data;
       * Lock device and try to bind to it. We drop the error
       * here and always return 0, because we need to keep trying
       * to bind to devices and some drivers will return an error
       * simply if it didn't support the device.
       * driver probe device() will spit a warning if there
       * is an error.
       */
      /* 调用 bus 的 match (),在这里是 platform bus type 的 mach(),即 platform match()例
      程,其在《Linux设备模型浅析之设备篇》中分析过*/
      if (drv->bus->match && !drv->bus->match(dev, drv))
             return 0;
      if (dev->parent)
                          /* Needed for USB */
             down(&dev->parent->sem);
      down(&dev->sem);
      if (!dev->driver) // 显然本例中 s3c device rtc 在注册时没有找到 driver, 所以这里会执行
             driver probe device(drv, dev);
                                              // 这里开始 probe
      up(&dev->sem):
      if (dev->parent)
             up(&dev->parent->sem);
      return 0;
```

```
3.4.2.1. driver probe device()在之前的《Linux 设备模型浅析之设备篇》文章已经分析,所
以这里直接拷贝过来。
driver probe device()也是定义在 drivers/base/dd.c 中,代码如下:
int driver probe device(struct device driver *drv, struct device *dev)
{
     int ret = 0:
     if (!device is registered(dev)) // 判断 dev 是否已经注册
           return -ENODEV;
     /* 调用 bus 的 match (),在这里是 platform bus type 的 mach(),即 platform match()例
     程,其在《Linux设备模型浅析之设备篇》中分析过*/
     if (drv->bus->match && !drv->bus->match(dev, drv))
           goto done;
     pr debug("bus: '%s': %s: matched device %s with driver %s\n",
            drv->bus->name, func , dev name(dev), drv->name);
     // 这里真正开始调用用户在 device driver 中注册的 probe()例程
     ret = really probe(dev, drv);
done:
     return ret;
下面分析 really probe()例程,顾名思义,其真正开始 probe 了。之前的《Linux 设备模型浅析之
设备篇》文章已经分析,所以这里直接拷贝过来。
    3.4.2.2. really probe()定义在drivers/base/dd.c中,代码如下:
static int really probe(struct device *dev, struct device driver *drv)
     int ret = 0:
     atomic inc(&probe count);
     pr debug("bus: '%s': %s: probing driver %s with device %s\n",
           drv->bus->name, __func__, drv->name, dev name(dev));
     WARN ON(!list empty(&dev->devres head));
     dev->driver = drv: // 将匹配的 driver 指针关联到 dev, 以便后续使用
     /*如果设备和驱动已经关联了,则在 dev 目录下,即 s3c2410-rtc 目录下生成名
     为"driver"的链接文件,指向其关联的驱动 dev->driver 的 sys 目录,并且在 dev->
     driver 的 sys 目录下生成链接文件, 名字和 dev 的名字一样, 即"3c2410-wdt", 指
     向/sys/devices/platform/s3c2410-rtc 目录
     */
     if (driver sysfs add(dev)) {
           printk(KERN ERR "%s: driver sysfs add(%s) failed\n",
                   func , dev name(dev));
           goto probe failed;
```

```
}
      // 如果设置了 dev->bus->probe,则调用,在 platform bus type 没有设置
      if (dev->bus->probe) {
            ret = dev->bus->probe(dev);
            if (ret)
                   goto probe failed;
      /* 所以,调用驱动注册在 device driver 里的 probe,这个很常用,用于获得硬件资源,初
      始化硬件等,在本例中就是调用注册到 driver 的 s3c rtc probe ()例程。
      } else if (drv->probe) {
            ret = drv - probe(dev);
            if (ret)
                   goto probe failed;
      }
      // 将 device 添加到 driver 列表中,并通知 bus 上的设备,表明 BOUND DRIVER。
      driver bound(dev);
      ret = 1;
      pr debug("bus: '%s': %s: bound device %s to driver %s\n",
             drv->bus->name, __func , dev name(dev), drv->name);
      goto done;
probe failed:
      devres release all(dev);
      driver sysfs remove(dev);
      dev->driver = NULL;
      if (ret != -ENODEV && ret != -ENXIO) {
            /* driver matched but the probe failed */
            printk(KERN WARNING
                "%s: probe of %s failed with error %d\n",
                drv->name, dev name(dev), ret);
       * Ignore errors returned by ->probe so that the next driver can try
       * its luck.
       */
      ret = 0;
done:
      atomic dec(&probe count);
      wake up(&probe waitqueue);
      return ret;
代码中,
```

3.4.2.2.1.在 s3c_rtc_probe ()中获取了硬件资源和注册了 rtc device, 所以会产生相关的文件夹和文件,并调用 rtc_device_register("s3c", &pdev->dev, &s3c_rtcops, THIS_MODULE)注册 rtc 类设备,在《Linux 设备模型浅析之设备篇》中做了具体分析,请参照。

```
3.5. 接着执行 module add driver()例程,其定义在 drivers/base/module.c 中,代码如下:
void module add driver(struct module *mod, struct device driver *drv)
      char *driver name;
      int no warn;
      struct module kobject *mk = NULL;
      if (!drv)
            return;
      if (mod)
                  // 本例中设置为 THIS MODULE, 所以执行
            mk = \& mod > mkobi;
      else if (drv->mod name) { // 如果设置了模块的名字,则到 module kset 容器列表中查找
            struct kobject *mkobj;
            /* Lookup built-in module entry in /sys/modules */
            mkobj = kset find obj(module kset, drv->mod name);
                                                            // 根据模块名查找
            if (mkobj) {
                  mk = container of(mkobj, struct module kobject, kobj);
                  /* remember our module structure */
                  drv - p - mkobj = mk;
                  /* kset find obj took a reference */
                  kobject put(mkobj);
      }
      if (!mk)
            return;
      /* Don't check return codes; these calls are idempotent */
      /* 本例中,假设 rtc-s3c.c 驱动编译成模块,手工在 shell 中使用 insmod 命令加载。所以,
      会在/sys/bus/platform/drivers/s3c2410-rtc/目录下生成名为"module"的链接文件,指
      向/sys/modules/rtc-s3c 目录,至于/sys/modules/rtc-s3c 目录是如何产生的,稍后将做分析
      no warn = sysfs create link(&drv->p->kobj, &mk->kobj, "module");
      // 本例中, 生成的 driver name 是"platform:s3c2410-rtc", 你看了该例程的实现就会明白
      driver name = make driver name(drv);
      if (driver name) {
            // 生成/sys/modules/rtc-s3c/drivers 目录
            module create drivers dir(mk);
            /* 本例中,在/sys/modules/rtc-s3c/drivers 目录下生成名为"platform:s3c2410-rtc"的
            链接文件,指向/sys/bus/platform/drivers/s3c2410-rtc/目录*/
            no warn = sysfs create link(mk->drivers dir, &drv->p->kobj,
                                driver name);
            kfree(driver name);
      }
}
```

代码中,

- 3.5.1. 看了上面的分析,一定会产生一个疑问,/sys/modules/rtc-s3c 目录是如何产生的呢? 下面就说这个问题。首先说说/sys/modules 目录是如何产生的。在 kernel/params.c 中有个初始化 例程 param sysfs init()在系统初始化的时候会调用,在该例程中调用了 module kset = kset create and add("module", &module uevent ops, NULL), 显然生成了一个 kset 容器, 产生 了/svs/module 目录,该 kset 容器被赋给了全局指针 module kset,所以我们所有的驱动模块的 mkobi 都挂在在它的名下。
- 3.5.2. 再说说/sys/modules/rtc-s3c 目录是如何产生的。rtc-s3c.c 驱动程序被编译成模块rtcs3c.ko。insmod 加载时会产生系统调用,调用到的内核入口程序是定义在 kernel/module.c 中的 init module()例程(其实是 sys init module())。该例程会调用在同一个文件中的 load_module() 例程,该例程会生成一个 struct module 并根据记载的模块进行初始化并将其加入到一个链表 中,以便以后进行引用。在 init module()例程中会调用到 mod sysfs init()例程,其代码如下: int mod sysfs init(struct module *mod)

```
int err;
struct kobject *kobj;
if (!module sysfs initialized) {
     printk(KERN ERR "%s: module sysfs not initialized\n",
         mod->name);
     err = -EINVAL;
     goto out;
}
// 先在 module kset 容器的列表中查找,看是否该 mod 已经加载
kobj = kset find obj(module kset, mod->name);
if (kobj) {
     // 加载过了则打印错误信息并返回
     printk(KERN ERR "%s: module is already loaded\n", mod->name);
     kobject put(kobj);
     err = -EINVAL;
     goto out;
}
mod->mkobj.mod = mod;
memset(&mod->mkobj.kobj, 0, sizeof(mod->mkobj.kobj));
// 将 kobj.kset 指向 module kset, 也就是包含在它的名下
mod->mkobj.kobj.kset = module kset;
/* 因为传入的 parent 参数为 NULL,所以会使用 module kset.kobj 作为 mkobj.kobj 的
partent kobj。mod->name 就是模块名,根据生成的模块来获得,本例中模块为rtc-
s3c.ko, 显然 mod->name = "rtc-s3c", 所以会产生/sys/module/rtc-s3c 目录。*/
err = kobject init and add(&mod->mkobj.kobj, &module ktype, NULL,
                   "%s", mod->name);
if (err)
     kobject put(&mod->mkobj.kobj);
```

/* delay uevent until full sysfs population */

out:

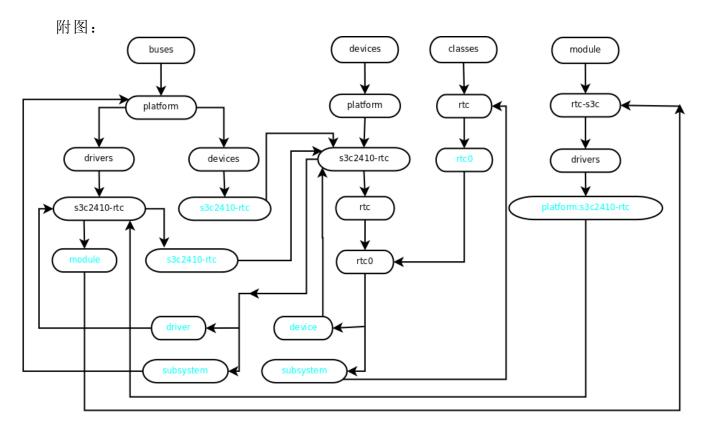
return err;

}

3.5.3. 顺便说说本例中 platform_driver_register(&s3c2410_rtc_driver)是被如何调用的,该例程被 模块初始化例程 s3c_rtc_init(void)调用。module_init(s3c_rtc_init)使得其会被赋给__this_module.init,而该 init 函数指针会在 init_module()例程中被调用。这样 platform driver register(&s3c2410 rtc driver)就被调用了,呵呵。

至此,platform_driver_register()例程完成调用。由于 platform_device s3c_device_rtc 是在系统初始化的时候注册的,所以 driver 能够找到匹配的 device,并产生一系列的文件夹和文件,可看附图。

作个小结,从上面的分析可以看出,sys 文件系统中 devices、bus、class 和 dev 目录里的内容之间的关联是通过调用 device_register()、driver_register()和 init_module()例程来完成的。很显然,linux 设备模型就这样建立起来了。



注:

- 1. 其中黑色字体的椭圆形表示是个文件夹;
- 2. 其中青色字体的椭圆形表示是个链接文件;
- 3. 用箭头表示文件夹之间的隶属关系和链接文件与文件夹之间的链接关系。