QEMU-KVM的初始化与客户系统的执行

原创 Lux_Veritas 最后发布于2013-07-19 17:11:34 阅读数 7474 ☆ 收藏

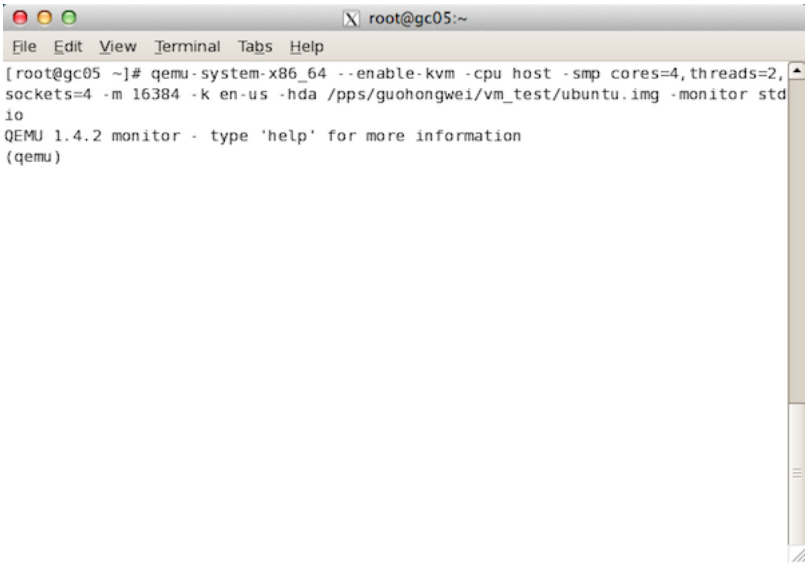
虚拟机运行概览

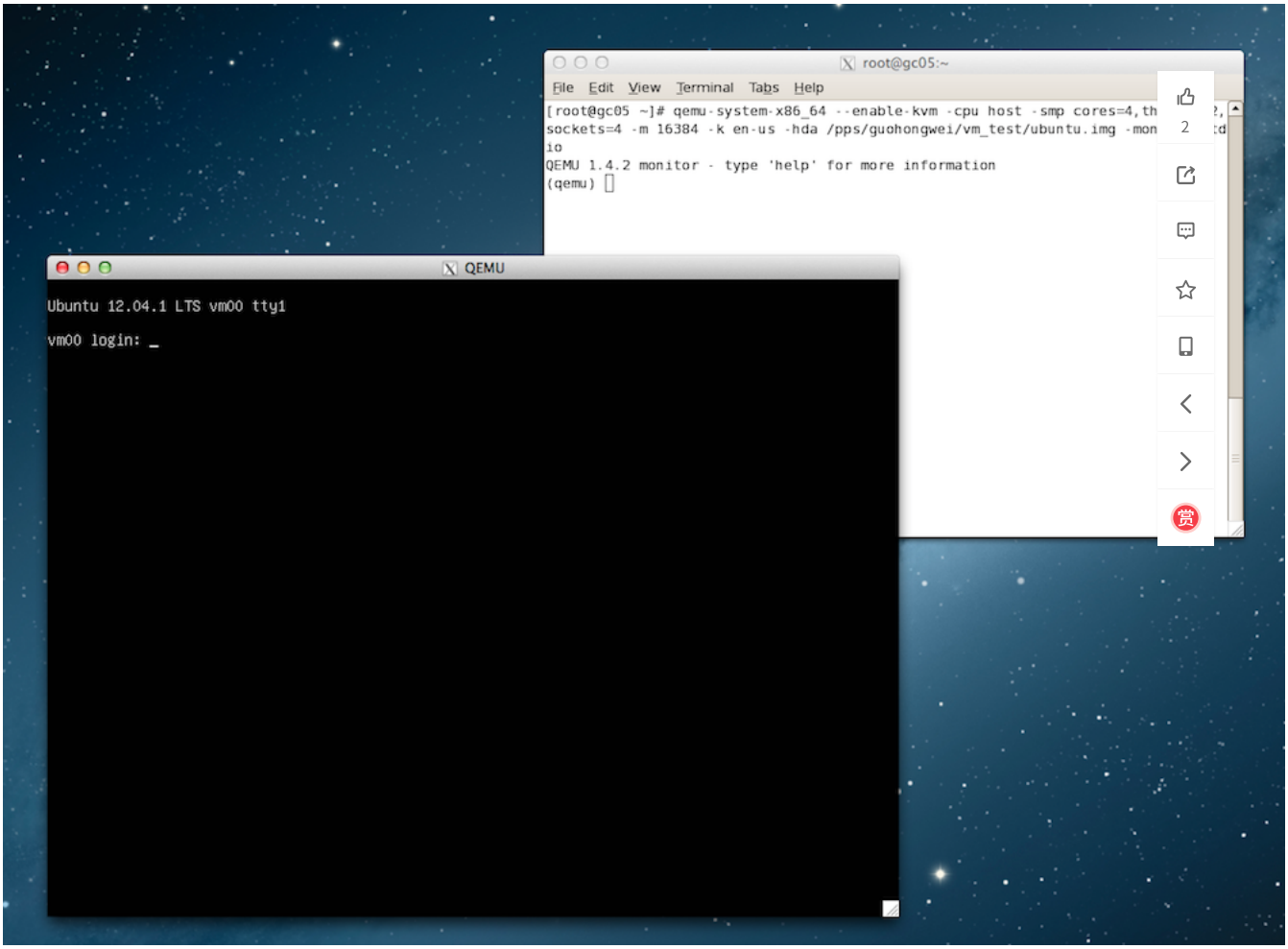
首先直观的了解一下利用QEMU运行客户虚拟机的流程。

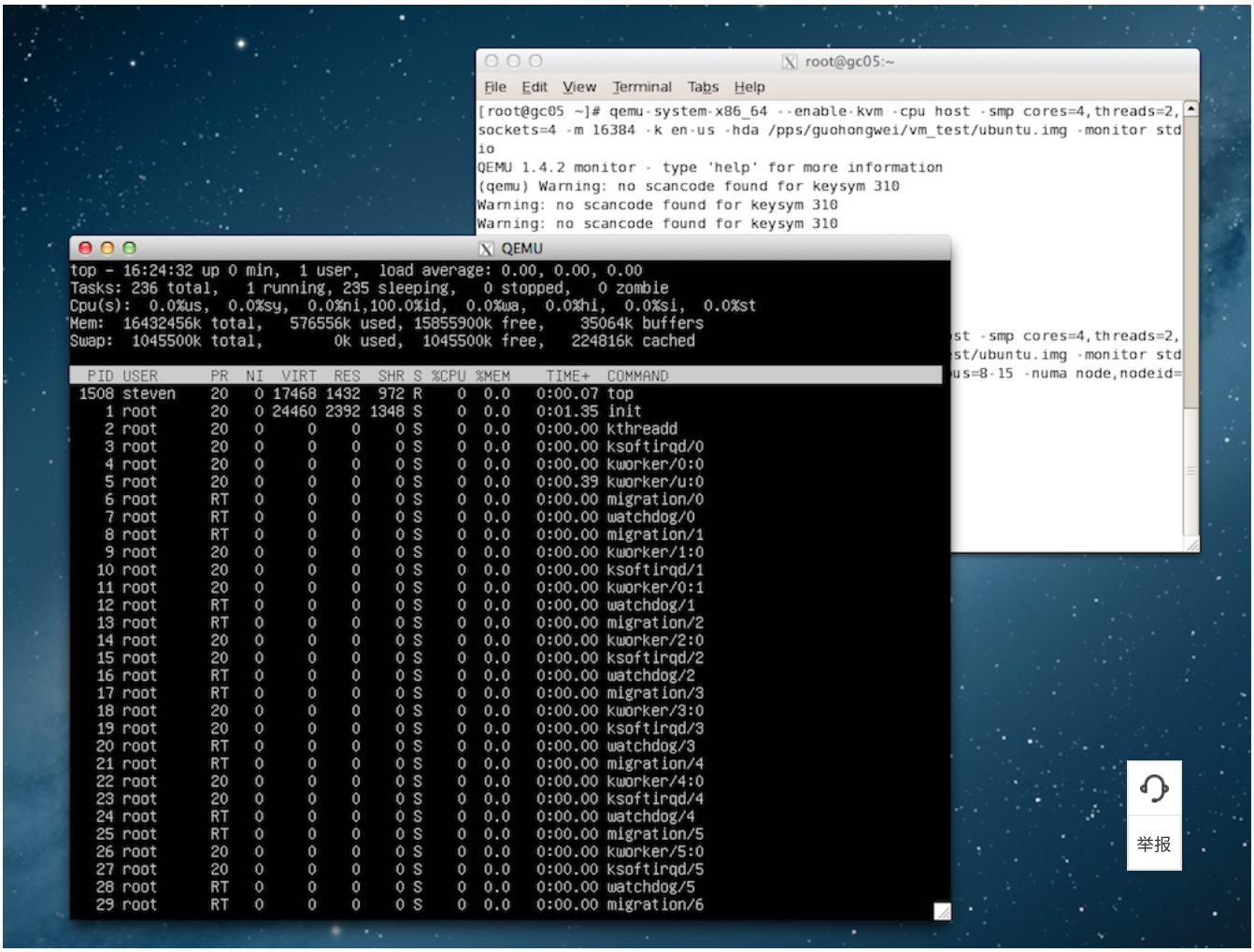
在命令行中运行QEMU的系统模式的可执行文件，参数声明虚拟CPU的个数，内存大小，指定已经安装好的硬盘镜像，启动QEMU虚拟机主窗口。启 格式举例：

```
qemu-system-x86_64 --enable-kvm -cpu host \
-smp cores=4,threads=2,sockets=4 \
-m 16384 -k en-us -hda /pps/guohongwei/vm_test/ubuntu.img -monitor stdio
```

以下为笔者的Mac-mini上弹出QEMU虚拟机界面的图示：




root@gc05:~
File Edit View Terminal Tabs Help
[root@gc05 ~]# qemu-system-x86_64 --enable-kvm -cpu host -smp cores=4,th
sockets=4 -m 16384 -k en-us -hda /pps/guohongwei/vm_test/ubuntu.img -mon
io
QEMU 1.4.2 monitor - type 'help' for more information
(qemu)
QEMU
Ubuntu 12.04.1 LTS vm00 tty1
vm00 login: _
2

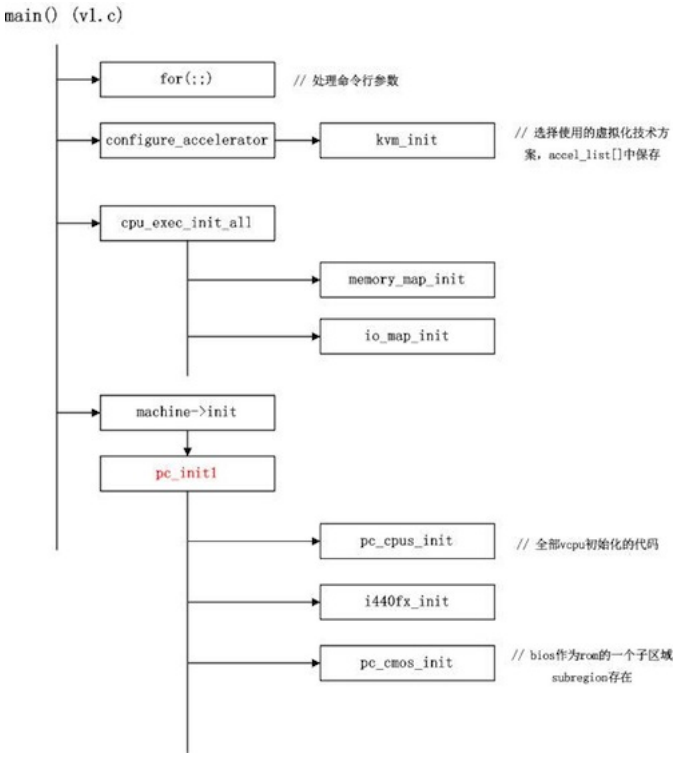

root@gc05:~
File Edit View Terminal Tabs Help
[root@gc05 ~]# qemu-system-x86_64 --enable-kvm -cpu host -smp cores=4,threads=2,
sockets=4 -m 16384 -k en-us -hda /pps/guohongwei/vm_test/ubuntu.img -monitor std
io
QEMU 1.4.2 monitor - type 'help' for more information
(qemu) Warning: no scancode found for keysym 310
Warning: no scancode found for keysym 310
Warning: no scancode found for keysym 310
QEMU
top - 16:24:32 up 0 min, 1 user, load average: 0.00, 0.00, 0.00
Tasks: 236 total, 1 running, 235 sleeping, 0 stopped, 0 zombie
Cpu(s): 0.0%us, 0.0%sy, 0.0%ni,100.0%id, 0.0%wa, 0.0%hi, 0.0%si, 0.0%st
Mem: 16432456k total, 576556k used, 15855900k free, 35064k buffers
Swap: 1045500k total, 0k used, 1045500k free, 224816k cached
PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
1508 steven 20 0 17468 1432 972 R 0 0.0 0:00.07 top
1 root 20 0 24460 2392 1348 S 0 0.0 0:01.35 init
2 root 20 0 0 0 0 S 0 0.0 0:00.00 kthreadd
3 root 20 0 0 0 0 S 0 0.0 0:00.00 ksoftirqd/0
4 root 20 0 0 0 0 S 0 0.0 0:00.00 kworker/0:0
5 root 20 0 0 0 0 S 0 0.0 0:00.39 kworker/u:0
6 root RT 0 0 0 0 S 0 0.0 0:00.00 migration/0
7 root RT 0 0 0 0 S 0 0.0 0:00.00 watchdog/0
8 root RT 0 0 0 0 S 0 0.0 0:00.00 migration/1
9 root 20 0 0 0 0 S 0 0.0 0:00.00 kworker/1:0
10 root 20 0 0 0 0 S 0 0.0 0:00.00 ksoftirqd/1
11 root 20 0 0 0 0 S 0 0.0 0:00.00 kworker/0:1
12 root RT 0 0 0 0 S 0 0.0 0:00.00 watchdog/1
13 root RT 0 0 0 0 S 0 0.0 0:00.00 migration/2
14 root 20 0 0 0 0 S 0 0.0 0:00.00 kworker/2:0
15 root 20 0 0 0 0 S 0 0.0 0:00.00 ksoftirqd/2
16 root RT 0 0 0 0 S 0 0.0 0:00.00 watchdog/2
17 root RT 0 0 0 0 S 0 0.0 0:00.00 migration/3
18 root 20 0 0 0 0 S 0 0.0 0:00.00 kworker/3:0
19 root 20 0 0 0 0 S 0 0.0 0:00.00 ksoftirqd/3
20 root RT 0 0 0 0 S 0 0.0 0:00.00 watchdog/3
21 root RT 0 0 0 0 S 0 0.0 0:00.00 migration/4
22 root 20 0 0 0 0 S 0 0.0 0:00.00 kworker/4:0
23 root 20 0 0 0 0 S 0 0.0 0:00.00 ksoftirqd/4
24 root RT 0 0 0 0 S 0 0.0 0:00.00 watchdog/4
25 root RT 0 0 0 0 S 0 0.0 0:00.00 migration/5
26 root 20 0 0 0 0 S 0 0.0 0:00.00 kworker/5:0
27 root 20 0 0 0 0 S 0 0.0 0:00.00 ksoftirqd/5
28 root RT 0 0 0 0 S 0 0.0 0:00.00 watchdog/5
29 root RT 0 0 0 0 S 0 0.0 0:00.00 migration/6
举报

QEMU的核心初始化流程

客户系统运行之前，QEMU作为全系统模拟软件，需要为客户系统模拟出CPU、主存以及I/O设备，使客户系统就像运行在真实硬件之上，而不用对 统代码做修改

如概览部分所示，由用户为客户系统指定需要的虚拟CPU资源（包括CPU核心数，SOCKET数目，每核心的超线程数，是否开启NUMA等等），虚拟 源，具体参数 ${QEMU}/qemu-options.hx。创建QEMU主线程，执行QEMU系统的初始化，在初始化的过程中针对每一个虚拟CPU，单独创建一个posix线程。每 虚拟CPU线程 上执行时，该VCPU对应的一套完整的寄存器集合被加载到物理CPU上，通过VM-LAUNCH或VM-RESUME指令切换到非根模式执行。直到该线程时间 或者其它中 VCPU退出到根模式，进行异常处理。

如下图所示，当用户运行QEMU的System Mode的可执行文件时，QEMU从${QEMU}/vl.c的main函数执行主线程。以下着重分析，客户系统启动之 EMU所做的初



1.处理命令行参数：

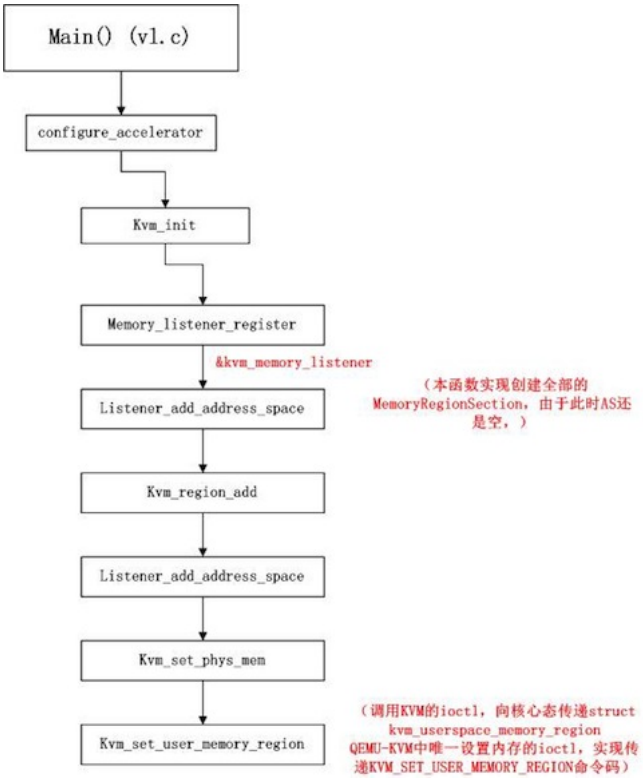
进入vl.c的main函数，首先有一个很长的for(;;)循环，用于分析处理通过命令行传进来的参数，进行相应系统的初始化设置。比如创建多少VCPU，是否开启NUMA，分 等等。

2.选择虚拟化方案：

configure_accelerator()函数，选择使用哪种虚拟化解决方案。

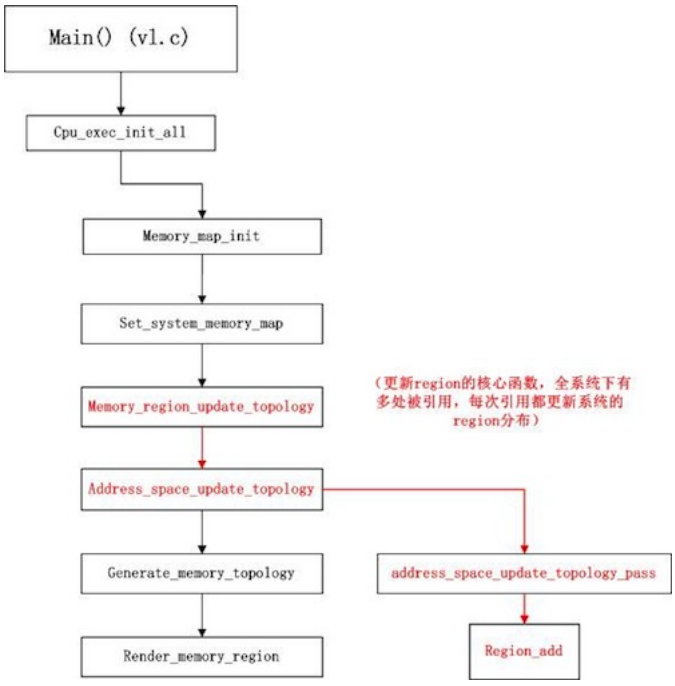
```
accel_list[] = {
  { "tcg", "tcg", tcg_available, tcg_init, &tcg_allowed },
  { "xen", "Xen", xen_available, xen_init, &xen_allowed },
  { "kvm", "KVM", kvm_available, kvm_init, &kvm_allowed },
  { "qtest", "QTest", qtest_available, qtest_init, &qtest_allowed },
};
```

accel_list[]数组声明了QEMU使用的系统模拟方案。“tcg”模式是不使用任何硬件虚拟化辅助方式，采用基于二进制指令翻译的方式，将目标平台的指令代码通过一个 为本机可以执行的指令。“xen”、“kvm”分别为两种主流的开源虚拟化解决方案。本文主要针对kvm这种硬件辅助的虚拟化解决方案。



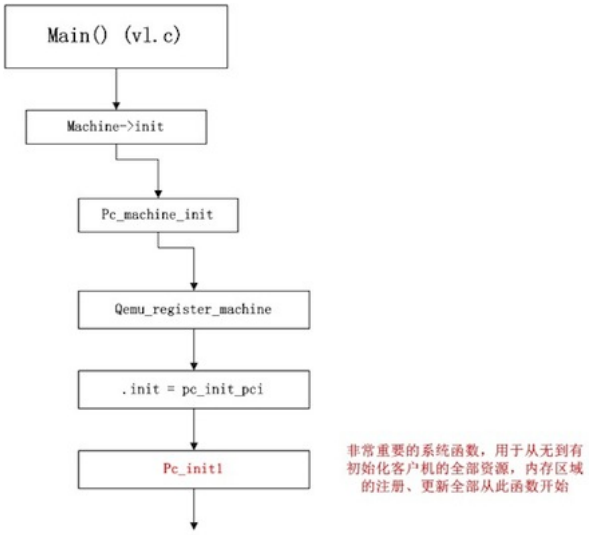
3.初始化内存布局：

新版本的QEMU(1.4)中，cpu_exec_init_all()函数只负责注册主存与IO内存两个顶层的memory_region，并且注册memory_listener。



4.虚拟客户机硬件初始化：

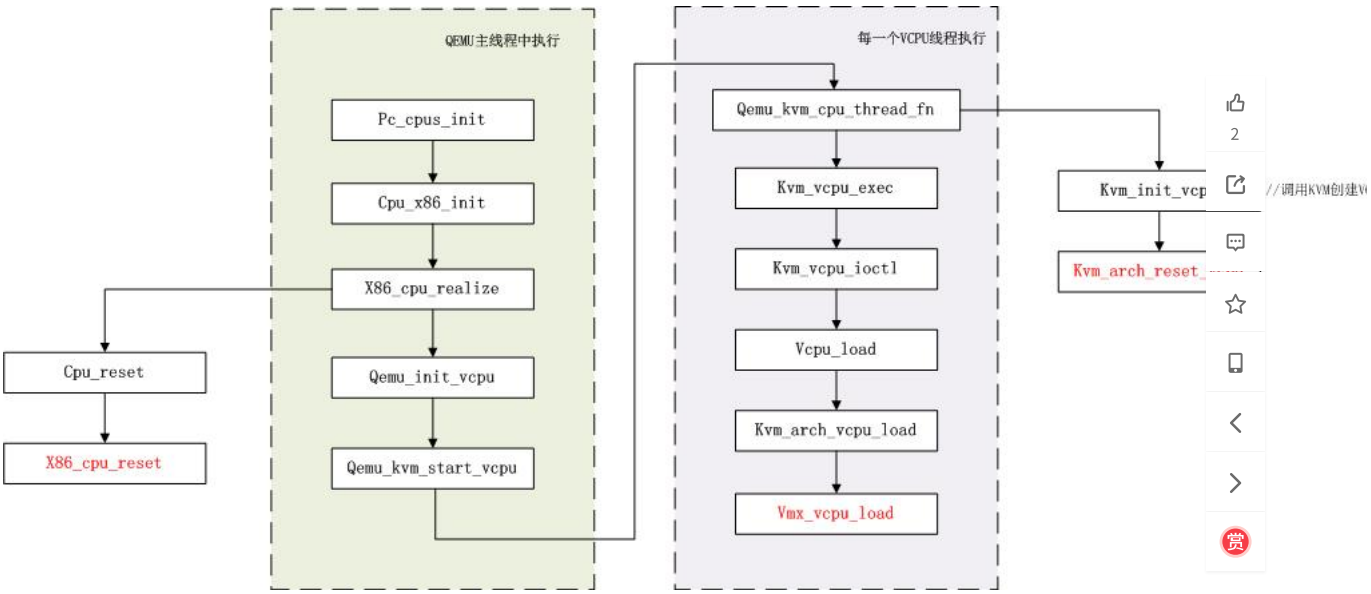
在完成了QEMU自身的初始化工作后，便开始了客户系统核心的初始化工作，主要是QEMU根据命令行参数，为客户系统创建虚拟的CPU、内存、I/ 核心过程是>init(&args)，对于x86目标平台，实际调用的是pc_init1函数。下面着重分析该函数。



4.1 VCPU初始化pc_cpus_init()

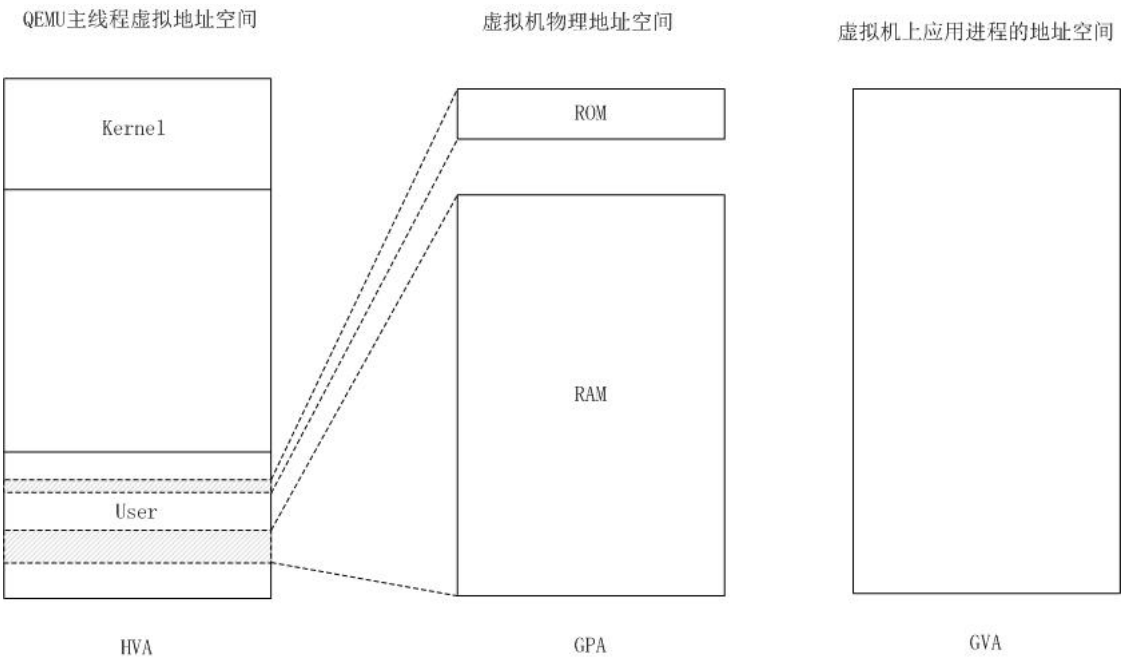
```
void pc_cpus_init(const char *cpu_model)
{
    int i;


    /* init CPUs */
    if (cpu_model == NULL) {
#ifdef TARGET_X86_64
        cpu_model = "qemu64";
#else
        cpu_model = "qemu32";
#endif
    }


    for (i = 0; i < smp_cpus; i++) {
        if (!cpu_x86_init(cpu_model)) {
            fprintf(stderr, "Unable to find x86 CPU definition\n");
            exit(1);
        }
    }
}
```

pc_init1的函数调用关系如下图所示，对于每一个即将创建的VCPU（个数由命令行传入smp_cpus），执行cpu_x86_init，逐层调用后，由qemu_kvm_start_vcpu创
新的VCPU线程将执行qemu_kvm_cpu_thread_fn函数，逐层调用后经过kvm_vcpu_ioctl系统调用切换到核心态，由KVM执行VCPU的创建工作，包括创建VMCS等非
的核心数据结构。



4.2 pc_memory_init初始化主存空间

利用mmap系统调用，在QEMU主线程的虚拟地址空间中申明一段连续的大小的空间用于客户机物理内存映射。在QEMU的内存管理结构中逐步添加subregion。此处添
memory_region，高于4g的memory_region，BIOS的memory_region。并调用bochs_bios_init，初始化BIOS。



4.3 i440fx_init初始化桥片及总线结构

4.4 pc_cmos_init初始化CMOS及时钟

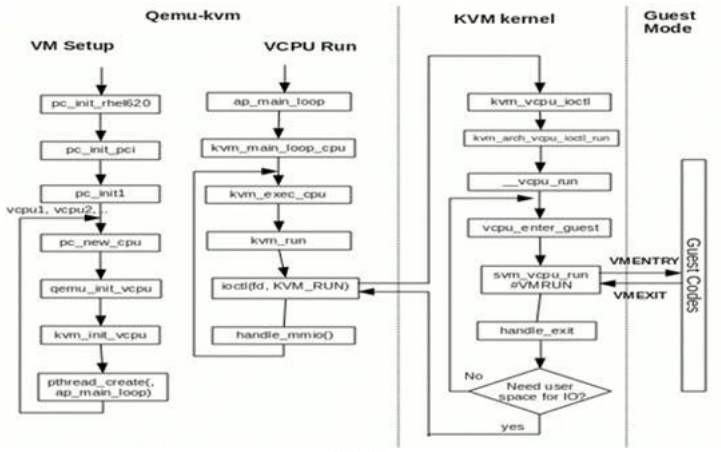
客户系统的执行流程

在创建了全部VCPU后，这些VCPU线程并没有被立即调度执行，直至vl.c的main函数执行完全部初始化工作后，调用resume_all_vcpus()，将pcpu->stop和 pcpu->st
VCPU线程再次被调度到物理CPU上执行时，VCPU正式开始工作。

由于pc_memory_init阶段已将BIOS初始化，创建了“bios.bin”文件到内存的映射，已处于非根模式下的VCPU到客户物理地址0xFFFF0（即“bio 文件被映射
偏移）处，获取第一条指令，开始执行客户系统代码。

此时多个VCPU线程由宿主系统轮流调度执行，QEMU主线程处于循环中，用于接收来自客户系统传回的I/O模拟请求。每一个VCPU线程，只要被调度的物理CPU上，便
客户系统代码，产生需要退出的异常时（如EPT缺页，处理I/O指令等），保存异常原因到VMCS，切换到根模式下，由KVM捕获该异常，查询VMCS异常号执行相应处理

非根模式，如此循环往复执行下去。



Lux_Veritas
发布了28 篇原创文章 · 获赞 23 · 访问量 19万+

kvm_cpu_exec 阅读数 605

http://blog.csdn.net/dashulu/article/details/17090293接着KVM虚拟机IO处理过程中Guest Vm IO处理过程(http://… 博文 来自： tycoon的专栏

qemu-kvm 之 CPU 绑定 阅读数 1831

cpu绑定场景CPU的绑定设置，是指将进程绑定到特定的一个或多个CPU上去执行，而不允许调度到其他的CPU上。… 博文

cento7安装kvm并通过kvm命令行安装centos7 阅读数 4236

这里写自定义目录标题一.KVM简介二.KVM虚拟化平台构建功能快捷键合理的创建标题，有助于目录的生成如何改变… 博文 来自： 飞龙的博客

qemu-kvm 命令行方式启动虚拟机 阅读数 4750

对于KVM虚拟机，一般启动我们会用virshcreate×××.xml方式启动，其实底层还是调用了qemu-kvm命令行去执行… 博文 来自： lonely_geek的博客

Qemu-KVM虚拟机初始化及创建过程源码简要分析（二） 阅读数 3539

前面我们讲了KVM内核层创建及初始化虚拟机的一些工作过程，现在讲一下Qemu层的流程以及与KVM内核层的配合… 博文 来自： Mr.Buffoon

Qemu-KVM虚拟机初始化及创建过程源码简要分析（一） 阅 51

我们知道，Qemu-KVM实际上包括Qemu和KVM两部分，那么在创建以及初始化虚拟机时，实际上也是在这两部分进… 博文 来自： Mr.Buffoon

QEMU使用简介 阅读数 1万+

QEMU使用简介。 博文 来自： jiangwei0512的博客

qemu-kvm分析 阅读数 449

虚拟机运行概览首先直观的了解一下利用QEMU运行客户虚拟机的流程。在命令行中运行QEMU的系统模式的可执行… 博文 来自： sdulibh的专栏

QEMU/seaBIOS启动流程分析 阅读数 524

1 QEMU函数执行流程 machine_init(pc_machine_init) –> pc_machine_init(void) –&… 博文 来自： 北方南方

虚拟化与QEMU-KVM系统分析系列_运维_Lux_Veritas的专栏-CSDN博客

Qemu-KVM虚拟机初始化及创建过程源码简要分析(一)_运维..._CSDN博客

Qemu&Kvm虚拟计算机系统启动流程图 阅读数 27

详细地跟踪分析了Qemu&Kvm虚拟计算机系统构建实例化与启动流程

...EPT页表的建立过程_运维_Lux_Veritas的专栏-CSDN博客

KVM初始化过程_网络_li_jiejun的专栏-CSDN博客

kvm组建安装与虚拟机初始化 阅读数 318

kvm组建安装与虚拟机初始化一 检查系统基础配置1查看该服务器是否支持虚拟化grep -E -o 'vmx|svm' /proc/cpuinf… 博文 来自： 技术之家

Qemu-KVM虚拟机初始化及创建过程源码简要分析(一)_Mr.B..._CSDN博客

KVM初始化过程 - dashulu的专栏 - CSDN博客

创建kvm虚拟机模板，及以模板新建虚拟机 阅读数 2555

环境：利用centos-7 虚拟机 。 先创建一个虚拟机模板 1.配置yum 导入gpg key2.禁用 selinux vim /etc/selinu… 博文 来自： wubude的博客

Centos7环境下制作kvm模板 阅读数 782

取消网卡自动改名cat > /etc/sysconfig/grub <<EOFGRUB_TIMEOUT=5GRUB_DEFAULT=s... 博文 来自： Ropon运维

KVM的初始化过程 - weixin_30411819的博客 - CSDN博客

Qemu创建KVM虚拟机内存初始化流程_xelatex KVM-CSDN博客

KVM虚拟化Windows 模版制作步骤 阅读数 97

1. 安装安装虚拟机设置： CPU内存大小（MB）硬盘方式、大小（G）网卡方式Windows 2003 系列2个512IDE、5G… 博文 来自： weixin_34087307…

QEMU安装和启动客户机 - bin_linux96的专栏

qemu 4.0.0安装 阅读数 994

安装qemucd qemu1) ./configure --target-list=arm-softmmu,arm-linux-user #只安装arm平台2) ./configure #安装… 博文 来自： 深空深蓝的博客

KVM虚拟机IO处理过程(二) ----QEMU/KVM I/O 处理过程 阅读数 384

原址：http://blog.csdn.net/dashulu/article/details/17090293 接着KVM虚拟机IO处理过程中Guest Vm IO处理过… 博文 来自： 慢慢游

qemu-kvm部分流程/源代码分析 阅读数 1125

接触虚拟化只有几个月，阅读qemu-kvm代码过程中，作了一点总结，画成流程图，如下(后续还会画qemu-kvm中… 博文 来自： ithinkwalk的专栏

QEMU代码分析（1）－module_init()构造函数 阅读数 2066

最近在看QEMU2.0源代码，决定把看的东西j 博文 来自： 无为而无不…

1.1Qemu 用户态架构 阅…517

本节首先分析Qemu的初始化的顶层流程;从而引出Qemu各大功能模块的描述；最后分析Qemu与内核态KVM的通讯… 博文 来自： wanthelpin…客

qemu-kvm设备初始化 阅…769

1. vm_config_groups是一个数组，数组每一个成员是一个链表表头，这些链表包含了qemu-kvm的各种启动参数，… 博文 来自： yuxinghai…

qemu初始化代码分析 …-04

qemu初始化代码分析qemu初始化代码分析qemu初始化代码分析qemu初始化代码分析qemu初始化代码分析qemu初始… …载

kvm性能优化方案---cpu/内存/磁盘/网络 阅读数 2万+

kvm性能优化kvm性能优化，主要集中在cpu、内存、磁盘、网络，4个方面，当然对于这里面的优化，也是要分场… 博文 来自： 简单生活

KVM详解，太详细太深入了，经典 阅读数 2万+

KVM介绍（1）：简介及安装http://www.cnblogs.com/sammyliu/p/4543110.html学习KVM的系列文章：（1）介… 博文 来自： sdulibh的专栏

QEMU 3.0.0 新特性一览 阅读数 891

QEMU 在 2018年8月15发布了版本3.0.0， 正式从 2.12 进入了3.0 时代。而且到今年位为止，QEMU 已经有15个年头… 博文 来自： 专注虚拟化的Linuxer

QEMU,KVM及QEMU-KVM介绍 阅读数 142

What's QEMUQEMU是一个主机上的VMM（virtual machine monitor）,通过动态二进制转换来模拟CPU，并提供一… 博文 来自： weixin_33755554…

QEMU中的IOCTL 阅读数 706

5. QEMU中的IOCTL在QEMU-KVM中，用户空间的QEMU是通过IOCTL与内核空间的KVM模块进行通讯的。1. 创建KV… 博文 来自： tycoon的专栏

QEMU之初始化——ARM vexpress-a9（一） 阅读数 2908

在上一篇的介绍的main()函数中，其实QEMU并没有很多真正具体化的实质性的初始化，在main()函数中做的最多的… 博文 来自： CurtisGuo的专栏

10.3QEMU基于VNC的桌面虚拟化原理 阅读数 3623

本节分析qemu基于vnc的桌面虚拟化的工作原理 博文 来自： wanthelping的博客

编译qemu和libvirt使支持SDL 阅读数 2046

登录centos官网，分别下载版本源码包：qemu-kvm-1.5.3-60.el7.src.rpmlibvirt-1.1.1-29.el7.src.rpm要安装rpmbu… 博文 来自： 酒醉东坡的专栏

在中国程序员是青春饭吗? 阅读数 17万+

今年，我也32了 ，为了不给大家误导，咨询了猎头、圈内好友，以及年过35岁的几位老程序员……舍了老脸去揭人… 博文 来自： 启舰

Synchronized关键字深析（小白慎入，深入jvm源码，两万字长文） 阅读数 1万+

从jvm层面解析synchronized，看完绝对可以超越绝大数人 博文 来自： Java新生代

大学四年，我决定把Java学习过的书籍都分享一遍 阅读数 1万+

给岁月以文明，而不是给文明以岁月，技术人读书我觉得很有必要，那这份书单的大部分书我觉得对您有用。... 博文 来自： 敖丙

程序员请照顾好自己，周末病魔差点一套带走我。 阅读数 8万+

程序员在一个周末的时间，得了重病，差点当场去世，还好及时挽救回来了。... 博文 来自： 敖丙

卸载 x 雷某度！GitHub 标星 1.5w+，从此我只用这款全能高速下载工具！ 阅读数 17万+

作者 | Rocky0429来源 | Python空间大家好，我是 Rocky0429，一个喜欢在网上收集各种资源的蒟蒻…网上资源眼花… 博文 来自： Rocky0429

为什么猝死的都是程序员，基本上不见产品经理猝死呢? 阅…万+

相信大家时不时听到程序员猝死的消息，但是基本上听不到产品经理猝死的消息，这是为什么呢？我们先百度搜一下… 博文 来自： 曹银飞的专…

毕业5年，我问遍了身边的大佬，总结了他们的学习方法 阅读数 15万+

我问了身边10个大佬，总结了他们的学习方法，原来成功都是有迹可循的。 博文 来自： 敖丙

推荐10个堪称神器的学习网站 阅读数 25万+

每天都会收到很多读者的私信，问我：“二哥，有什么推荐的学习网站吗？最近很浮躁，手头的一些网站都看烦了，… 博文 来自：沉默王二

阿里程序员写了一个新手都写不出的低级bug，被骂惨了。 阅读数 [illegible]万+

这种新手都不会范的错，居然被一个工作好几年的小伙子写出来，差点被当场开除了。... 博文 来自：敖丙

大学四年自学走来，这些私藏的实用工具/学习网站我贡献出来了 阅读数 [illegible]万+

大学四年，看课本是不可能一直看课本的了，对于学习，特别是自学，善于搜索网上的一些资源来辅助，还是非常有… 博文 来自：帅地

Java C语言 Python C++ C# Visual Basic .NET JavaScript PHP SQL Go语言 R语言 Assembly language Swift Ru
MATLAB PL/SQL Perl Visual Basic Objective-C Delphi/Object Pascal Unity3D

Lux_Veritas

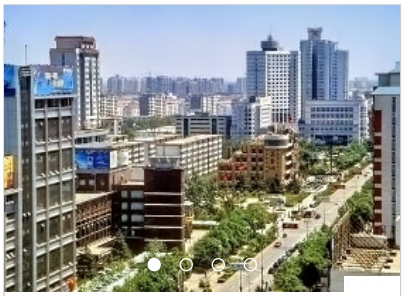
TA的个人主页 >

原创	粉丝	获赞	评论	访问
28	31	23	15	19万+

等级: 博客 4 周排名: 17万+

积分: 1551 总排名: 5万+

关注 私信



固安楼盘

最新文章

dnsmasq作DHCP服务器配置

Bandwidth内存带宽测试工具

VIM常用命令介绍

Gtk-WARNING **: cannot open display问题的解决

linux查看系统空闲内存的方法

分类专栏

Most Important 2篇

杂谈 3篇

tricky things 3篇

内核学习 8篇

内功修炼 5篇

展开

举报

归档

展开

最新评论