Subject: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by jmslkn on Mon, 30 Nov 2009 17:22:12 GMT
View Forum Message <> Reply to Message

I am trying to rebuild the kvm-kmod kernel module for the new openvz kernel but it seems some functions are missing:

rpmbuild --rebuild kvm-83-105.el5_4.9.src.rpm --define "kversion 2.6.18-164.2.1.el5.028stab066.7" --nodeps

[..]
make[1]: Entering directory `/usr/src/kernels/2.6.18-164.2.1.el5.028stab066.7-x86_64'
  LD   /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel- /ksm/built-in.o
  CC [M] /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel- /ksm/ksm_main.o
In file included from /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel-/x86/external-module-compat.h:14,
                 from <command line>:1:
/usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel- /include/asm/svm.h:219:1: warning: "MSR_VM_HSAVE_PA" redefined
In file included from include/asm/processor.h:16,
                 from /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel-/include/asm/kvm_para.h:89,
                 from /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel-/include/linux/kvm_para.h:64,
                 from /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel-/x86/../external-module-compat-comm.h:14,
                 from /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel-/x86/external-module-compat.h:9,
                 from <command line>:1:
include/asm/msr.h:448:1: warning: this is the location of the previous definition
In file included from /usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel-/ksm/ksm_main.c:45:
/usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel-/ksm/external-module-compat.h:39: error: static declaration of

include/linux/rmap.h:84: error: previous declaration of

make[2]: *** [/usr/src/redhat/BUILD/kvm-83-maint-snapshot-20090205/kernel -/ksm/ksm_main.o] Error 1
make[1]: *** [_module_/usr/src/redhat/BUILD/kvm-83-maint-snapshot-2009020 5/kernel-/ksm] Error 2
make[1]: Leaving directory `/usr/src/kernels/2.6.18-164.2.1.el5.028stab066.7-x86_64'
make: *** [all] Error 2

Any idea?

Thanks,

J.

Subject: Re: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by jormartr on Mon, 30 Nov 2009 17:58:37 GMT
View Forum Message <> Reply to Message

Is it even posible use kvm with the openvz patched 2.6.18 kernel?

I tried to do that, but I saw there is no kvm on this kernel.

Subject: Re: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by jmslkn on Tue, 01 Dec 2009 00:32:23 GMT
View Forum Message <> Reply to Message

I have created a small patch, it works to me, but it would be good to have some comments from the OpenVZ developers:

kvm-openvz.spec (apply on /usr/src/redhat/SPECS/kvm.spec)
--- kvm.spec 2009-10-05 18:28:37.000000000 +0200
+++ /usr/src/redhat/SPECS/kvm.spec 2009-12-01 00:23:25.000000000 +0100
@@ -796,6 +796,7 @@
 # end of kvm.git (kernel) patches
 ######################################3

+Patch600: kvm-openvz-page_lock_anon_vma.patch

 BuildRoot: %{_tmppath}/%{name}-%{version}-%{release}-root-%(%{__id_u} -n)
 BuildRequires: SDL-devel
@@ -1306,6 +1307,8 @@
 %patch582 -p1
 cd ../

+%patch600 -p0
+

 ## you can uncomment the code below to test the
 ## 'make sync' code

kvm-openvz-page_lock_anon_vma.patch (copy to /usr/src/redhat/SOURCES/):

--- kernel/ksm/external-module-compat.h.orig 2009-12-01 00:15:41.000000000 +0100

```
+++ kernel/ksm/external-module-compat.h 2009-12-01 00:15:45.000000000 +0100
@@ -35,7 +35,7 @@
 #define list_first_entry(ptr, type, member) \
        list_entry((ptr)->next, type, member)

-static struct anon_vma *page_lock_anon_vma(struct page *page)
+static struct anon_vma *page_lock_anon_vma2(struct page *page)
 {
  struct anon_vma *anon_vma;
  unsigned long anon_mapping;
@@ -241,7 +241,7 @@
  struct anon_vma *anon_vma;
  int ret = 0;

- anon_vma = page_lock_anon_vma(page);
+ anon_vma = page_lock_anon_vma2(page);
  if (!anon_vma)
   return ret;
```

Subject: Re: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by Ales on Fri, 04 Dec 2009 19:06:00 GMT
View Forum Message <> Reply to Message

We haven't started testing KVM on this kernel yet, but plan to do so next week or so. I'll report if we find anything unusual.

Some response from openVZ developers would be nice, though.

Perhaps we should file a bug in the openvz bugzilla for this KVM issue?

Subject: Re: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by dietmar on Tue, 08 Dec 2009 11:08:08 GMT
View Forum Message <> Reply to Message

I get a kernel panic very soon. Any ideas?

File Attachments
1) ksm-openvz-kernel-panic.jpg, downloaded 1750 times

Subject: Re: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by jmslkn on Wed, 09 Dec 2009 11:46:11 GMT
View Forum Message <> Reply to Message

I have the same panic here, when I start the ksm service. It seems there is a page that is

allocated on different path without "charging" it, and when the memory manager try to free that page (free_hot_cold_page), it crashes.

The relevant source lines are here:

/kernel/ub/ub_mem.c

```
/*
 * Pages accounting
 */

int ub_page_charge(struct page *page, int order, gfp_t mask)
{
	struct user_beancounter *ub;
	unsigned long flags;

	ub = NULL;
	if (!(mask & __GFP_UBC))
		goto out;

	ub = get_beancounter(get_exec_ub());
	if (ub == NULL)
		goto out;

	local_irq_save(flags);
	if (ub_kmemsize_charge(ub, CHARGE_ORDER(order),
				(mask & __GFP_SOFT_UBC ? UB_SOFT : UB_HARD)))
		goto err;

	inc_pages_charged(ub, order);
	local_irq_restore(flags);
out:
	BUG_ON(page_ub(page) != NULL);
	page_ub(page) = ub;
	return 0;

err:
	local_irq_restore(flags);
	BUG_ON(page_ub(page) != NULL);
	put_beancounter(ub);
	return -ENOMEM;
}

void ub_page_uncharge(struct page *page, int order)
{
	struct user_beancounter *ub;
	unsigned long flags;
```

```
	ub = page_ub(page);
	if (ub == NULL)
		return;

	BUG_ON(ub->ub_magic != UB_MAGIC);
	dec_pages_charged(ub, order);
	local_irq_save(flags);
	ub_kmemsize_uncharge(ub, CHARGE_ORDER(order));
	local_irq_restore(flags);
	put_beancounter(ub);
	page_ub(page) = NULL;
}
```

ksm/ksm_main.c

```
static int try_to_merge_two_pages_alloc(struct mm_struct *mm1,
					struct page *page1,
					struct mm_struct *mm2,
					struct page *page2,
					unsigned long addr1,
					unsigned long addr2)
{
	struct vm_area_struct *vma;
	pgprot_t prot;
	int ret = 1;
	struct page *kpage;


	kpage = alloc_page(GFP_HIGHUSER);
	if (!kpage)
		return ret;
	down_read(&mm1->mmap_sem);
	vma = find_vma(mm1, addr1);
	if (!vma) {
		put_page(kpage);
		up_read(&mm1->mmap_sem);
		return ret;
	}
	prot = vma->vm_page_prot;
	pgprot_val(prot) &= ~_PAGE_RW;

	copy_user_highpage(kpage, page1, addr1);
	ret = try_to_merge_one_page(mm1, vma, page1, kpage, prot);
	up_read(&mm1->mmap_sem);
[..]
```

However the alloc_page calls __alloc_page function, that is "charged". Any idea?

File Attachments
1) ksm.tar.gz, downloaded 494 times

Subject: Re: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by mainframe on Sat, 20 Feb 2010 22:51:30 GMT
View Forum Message <> Reply to Message

We have been rebuilding kvm module for openvz RHEL5 kernels and running these for 2 months - so far no kernel crashes, servers have been stable. We also build a bare-metal CentOS based ISO installer and are working on management solutions as well. ISO installer project can be found on http://opennode.sourceforge.net/

Subject: Re: KVM support on 2.6.18-164.2.1.el5.028stab066.7 (CentOS5.4)
Posted by jmslkn on Mon, 22 Feb 2010 18:44:35 GMT
View Forum Message <> Reply to Message

Thanks for the update. The same status here: stable server, no crashes.