KVM CPUID bits

Author: Glauber Costa <glommer@gmail.com>

A guest running on a kvm host, can check some of its features using cpuid. This is not always guaranteed to work, since userspace can mask-out some, or even all KVM-related cpuid features before launching a guest.

KVM cpuid functions are:

function: KVM_CPUID_SIGNATURE (0x40000000)

returns:

```
eax = 0x40000001
ebx = 0x4b4d564b
ecx = 0x564b4d56
edx = 0x4d
```

Note that this value in ebx, ecx and edx corresponds to the string "KVMKVMKVM". The value in eax corresponds to the maximum cpuid function present in this leaf, and will be updated if more functions are added in the future. Note also that old hosts set eax value to 0x0. This should be interpreted as if the value was 0x40000001. This function queries the presence of KVM cpuid leafs.

function: define KVM_CPUID_FEATURES (0x40000001)

returns:

```
ebx, ecx
eax = an OR'ed group of (1 << flag)
```

where `flag` is defined as below:

flag	value	meaning
KVM_FEATURE_CLOCKSOURCE	0	kvmclock available at msrs 0x11 and 0x12
KVM_FEATURE_NOP_IO_DELAY	1	not necessary to perform delays on PIO operations
KVM_FEATURE_MMU_OP	2	deprecated
KVM_FEATURE_CLOCKSOURCE2	3	kvmclock available at msrs 0x4b564d00 and 0x4b564d01
KVM_FEATURE_ASYNC_PF	4	async pf can be enabled by writing to msr 0x4b564d02
KVM_FEATURE_STEAL_TIME	5	steal time can be enabled by writing to msr 0x4b564d03
KVM_FEATURE_PV_EOI	6	paravirtualized end of interrupt handler can be enabled by writing to msr 0x4b564d04
KVM_FEATURE_PV_UNHALT	7	guest checks this feature bit before enabling paravirtualized spinlock support
KVM_FEATURE_PV_TLB_FLUSH	9	guest checks this feature bit before enabling paravirtualized tlb flush
KVM_FEATURE_ASYNC_PF_VMEXIT	10	paravirtualized async PF VM EXIT can be enabled by setting bit 2 when writing to msr 0x4b564d02
KVM_FEATURE_PV_SEND_IPI	11	guest checks this feature bit before enabling paravirtualized send IPIs
KVM_FEATURE_POLL_CONTROL	12	host-side polling on HLT can be disabled by writing to msr 0x4b564d05.
KVM_FEATURE_PV_SCHED_YIELD	13	guest checks this feature bit before using paravirtualized sched yield.
KVM_FEATURE_ASYNC_PF_INT	14	guest checks this feature bit before using the second async pf control msr 0x4b564d06 and async pf acknowledgment msr 0x4b564d07.
KVM_FEATURE_MSI_EXT_DEST_ID	15	guest checks this feature bit before using extended destination ID bits in MSI address bits 11-5.
KVM_FEATURE_HC_MAP_GPA_RANGE	16	guest checks this feature bit before using the map gpa range hypercall to notify the page state change
KVM_FEATURE_MIGRATION_CONTROL	17	guest checks this feature bit before using MSR_KVM_MIGRATION_CONTROL
KVM_FEATURE_CLOCKSOURCE_STABLE_BIT	24	host will warn if no guest-side per-cpu warps are expected in kvmclock

```
edx = an OR'ed group of (1 << flag)
```

Where `flag` here is defined as below:

flag	value	meaning

flag	value	meaning
KVM_HINTS_REALTIME	0	guest checks this feature bit to determine that vCPUs are never preempted for an unlimited time allowing optimizations