Power State Coordination Interface (PSCI)

KVM implements the PSCI (Power State Coordination Interface) specification in order to provide services such as CPU on/off, reset and power-off to the guest.

The PSCI specification is regularly updated to provide new features, and KVM implements these updates if they make sense from a virtualization point of view.

This means that a guest booted on two different versions of KVM can observe two different "firmware" revisions. This could cause issues if a given guest is tied to a particular PSCI revision (unlikely), or if a migration causes a different PSCI version to be exposed out of the blue to an unsuspecting guest.

In order to remedy this situation, KVM exposes a set of "firmware pseudo-registers" that can be manipulated using the GET/SET_ONE_REG interface. These registers can be saved/restored by userspace, and set to a convenient value if required.

The following register is defined:

- KVM_REG_ARM_PSCI_VERSION:
  - Only valid if the vcpu has the KVM_ARM_VCPU_PSCI_0_2 feature set (and thus has already been initialized)
  - Returns the current PSCI version on GET_ONE_REG (defaulting to the highest PSCI version implemented by KVM and compatible with v0.2)
  - Allows any PSCI version implemented by KVM and compatible with v0.2 to be set with SET_ONE_REG
  - Affects the whole VM (even if the register view is per-vcpu)
- KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_1:
  Holds the state of the firmware support to mitigate CVE-2017-5715, as offered by KVM to the guest via a HVC call. The workaround is described under SMCCC_ARCH_WORKAROUND_1 in [1].

  Accepted values are:

  KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_1_NOT_AVAIL:
  KVM does not offer firmware support for the workaround. The mitigation status for the guest is unknown.
  KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_1_AVAIL:
  The workaround HVC call is available to the guest and required for the mitigation.
  KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_1_NOT_REQUIRED:
  The workaround HVC call is available to the guest, but it is not needed on this VCPU.

- KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_2:
  Holds the state of the firmware support to mitigate CVE-2018-3639, as offered by KVM to the guest via a HVC call. The workaround is described under SMCCC_ARCH_WORKAROUND_2 in [1].

  Accepted values are:

  KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_2_NOT_AVAIL:
  A workaround is not available. KVM does not offer firmware support for the workaround.
  KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_2_UNKNOWN:
  The workaround state is unknown. KVM does not offer firmware support for the workaround.
  KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_2_AVAIL:
  The workaround is available, and can be disabled by a vCPU. If KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_2_ENABLED is set, it is active for this vCPU.
  KVM_REG_ARM_SMCCC_ARCH_WORKAROUND_2_NOT_REQUIRED:
  The workaround is always active on this vCPU or it is not needed.

[1] https://developer.arm.com/-/media/developer/pdf/ARM_DEN_0070A_Firmware_interfaces_for_mitigating_CVE-2017-5715.pdf