

ARM Virtual Interrupt Translation Service (ITS)

Device types supported:

KVM_DEV_TYPE_ARM_VGIC_ITS ARM Interrupt Translation Service Controller

The ITS allows MSI(-X) interrupts to be injected into guests. This extension is optional. Creating a virtual ITS controller also requires a host GICv3 (see arm-vgic-v3.txt), but does not depend on having physical ITS controllers.

There can be multiple ITS controllers per guest, each of them has to have a separate, non-overlapping MMIO region.

Groups

KVM_DEV_ARM_VGIC_GRP_ADDR

Attributes:

KVM_VGIC_ITS_ADDR_TYPE (rw, 64-bit)

Base address in the guest physical address space of the GICv3 ITS control register frame. This address needs to be 64K aligned and the region covers 128K.

Errors:

-E2BIG	Address outside of addressable IPA range
-EINVAL	Incorrectly aligned address
-EEXIST	Address already configured
-EFAULT	Invalid user pointer for attr->addr.
-ENODEV	Incorrect attribute or the ITS is not supported.

KVM_DEV_ARM_VGIC_GRP_CTRL

Attributes:

KVM_DEV_ARM_VGIC_CTRL_INIT

request the initialization of the ITS, no additional parameter in kvm_device_attr.addr.

KVM_DEV_ARM_ITS_CTRL_RESET

reset the ITS, no additional parameter in kvm_device_attr.addr. See "ITS Reset State" section.

KVM_DEV_ARM_ITS_SAVE_TABLES

save the ITS table data into guest RAM, at the location provisioned by the guest in corresponding registers/table entries.

The layout of the tables in guest memory defines an ABI. The entries are laid out in little endian format as described in the last paragraph.

KVM_DEV_ARM_ITS_RESTORE_TABLES

restore the ITS tables from guest RAM to ITS internal structures.

The GICV3 must be restored before the ITS and all ITS registers but the GITS_CTLR must be restored before restoring the ITS tables.

The GITS_IIDR read-only register must also be restored before calling KVM_DEV_ARM_ITS_RESTORE_TABLES as the IIDR revision field encodes the ABI revision.

The expected ordering when restoring the GICv3/ITS is described in section "ITS Restore Sequence".

Errors:

-ENXIO	ITS not properly configured as required prior to setting this attribute
-ENOMEM	Memory shortage when allocating ITS internal data
-EINVAL	Inconsistent restored data
-EFAULT	Invalid guest ram access
-EBUSY	One or more VCPUS are running
-EACCES	The virtual ITS is backed by a physical GICv4 ITS, and the state is not available without GICv4.1

KVM_DEV_ARM_VGIC_GRP_ITS_REGS

Attributes:

The attr field of `kvm_device_attr` encodes the offset of the ITS register, relative to the ITS control frame base address (`ITS_base`).

`kvm_device_attr.addr` points to a `__u64` value whatever the width of the addressed register (32/64 bits). 64 bit registers can only be accessed with full length.

Writes to read-only registers are ignored by the kernel except for:

- `GITS_CREADR`. It must be restored otherwise commands in the queue will be re-executed after restoring `CWRITER`. `GITS_CREADR` must be restored before restoring the `GITS_CTLR` which is likely to enable the ITS. Also it must be restored after `GITS_CBASER` since a write to `GITS_CBASER` resets `GITS_CREADR`.
- `GITS_IIDR`. The Revision field encodes the table layout ABI revision. In the future we might implement direct injection of virtual LPIs. This will require an upgrade of the table layout and an evolution of the ABI. `GITS_IIDR` must be restored before calling `KVM_DEV_ARM_ITS_RESTORE_TABLES`.

For other registers, getting or setting a register has the same effect as reading/writing the register on real hardware.

Errors:

-ENXIO	Offset does not correspond to any supported register
-EFAULT	Invalid user pointer for attr->addr
-EINVAL	Offset is not 64-bit aligned
-EBUSY	one or more VCPUS are running

ITS Restore Sequence:

The following ordering must be followed when restoring the GIC and the ITS:

- a. restore all guest memory and create vcpus
- b. restore all redistributors
- c. provide the ITS base address (`KVM_DEV_ARM_VGIC_GRP_ADDR`)
- d. restore the ITS in the following order:
 1. Restore `GITS_CBASER`
 2. Restore all other `GITS_` registers, except `GITS_CTLR`!
 3. Load the ITS table data (`KVM_DEV_ARM_ITS_RESTORE_TABLES`)
 4. Restore `GITS_CTLR`

Then vcpus can be started.

ITS Table ABI REV0:

Revision 0 of the ABI only supports the features of a virtual GICv3, and does not support a virtual GICv4 with support for direct injection of virtual interrupts for nested hypervisors.

The device table and ITT are indexed by the DeviceID and EventID, respectively. The collection table is not indexed by CollectionID, and the entries in the collection are listed in no particular order. All entries are 8 bytes.

Device Table Entry (DTE):

```
bits:      | 63| 62 ... 49 | 48 ... 5 | 4 ... 0 |
values:    | V |  next      | ITT_addr | Size    |
```

where:

- V indicates whether the entry is valid. If not, other fields are not meaningful.
- next: equals to 0 if this entry is the last one; otherwise it corresponds to the DeviceID offset to the next DTE, capped by $2^{14} - 1$.
- ITT_addr matches bits [51:8] of the ITT address (256 Byte aligned).
- Size specifies the supported number of bits for the EventID, minus one

Collection Table Entry (CTE):

```
bits:      | 63| 62 .. 52 | 51 ... 16 | 15 ... 0 |
values:    | V |  RES0    | RDBase   | ICID     |
```

where:

- V indicates whether the entry is valid. If not, other fields are not meaningful.
- RES0: reserved field with Should-Be-Zero-or-Preserved behavior.

- RDBase is the PE number (GICR_TYPER.Processor_Number semantic),
- ICID is the collection ID

Interrupt Translation Entry (ITE):

bits:	63 ... 48 47 ... 16 15 ... 0
values:	next pINTID ICID

where:

- next: equals to 0 if this entry is the last one; otherwise it corresponds to the EventID offset to the next ITE capped by $2^{16} - 1$.
- pINTID is the physical LPI ID; if zero, it means the entry is not valid and other fields are not meaningful.
- ICID is the collection ID

ITS Reset State:

RESET returns the ITS to the same state that it was when first created and initialized. When the RESET command returns, the following things are guaranteed:

- The ITS is not enabled and quiescent `GITS_CTLR.Enabled = 0` .Quiescent=1
- There is no internally cached state
- No collection or device table are used `GITS_BASER<n>.Valid = 0`
- `GITS_CBASER = 0`, `GITS_CREADR = 0`, `GITS_CWRITER = 0`
- The ABI version is unchanged and remains the one set when the ITS device was first created.