S/390 PCI

Authors:
- Pierre Morel

Command line parameters and debugfs entries

Command line parameters

- nomio

  Do not use PCI Mapped I/O (MIO) instructions.

- norid

  Ignore the RID field and force use of one PCI domain per PCI function.

debugfs entries

The S/390 debug feature (s390dbf) generates views to hold various debug results in sysfs directories of the form:

- /sys/kernel/debug/s390dbf/pci_*/

For example:

- /sys/kernel/debug/s390dbf/pci_msg/sprintf Holds messages from the processing of PCI events, like machine check handling and setting of global functionality, like UID checking.

Change the level of logging to be more or less verbose by piping a number between 0 and 6 to /sys/kernel/debug/s390dbf/pci_*/level. For details, see the documentation on the S/390 debug feature at Documentation/s390/s390dbf.rst.

Sysfs entries

Entries specific to zPCI functions and entries that hold zPCI information.

- /sys/bus/pci/slots/XXXXXXXX

  The slot entries are set up using the function identifier (FID) of the PCI function. The format depicted as XXXXXXXX above is 8 hexadecimal digits with 0 padding and lower case hexadecimal digitis.

  - /sys/bus/pci/slots/XXXXXXXX/power

  A physical function that currently supports a virtual function cannot be powered off until all virtual functions are removed with: echo 0 > /sys/bus/pci/devices/XXXX:XX:XX.X/sriov_numvf

- /sys/bus/pci/devices/XXXX:XX:XX.X/
  - function_id A zPCI function identifier that uniquely identifies the function in the Z server.
  - function_handle Low-level identifier used for a configured PCI function. It might be useful for debuging.
  - pchid Model-dependent location of the I/O adapter.
  - pfgid PCI function group ID, functions that share identical functionality use a common identifier. A PCI group defines interrupts, IOMMU, IOTLB, and DMA specifics.
  - vfn The virtual function number, from 1 to N for virtual functions, 0 for physical functions.
  - pft The PCI function type
  - port The port corresponds to the physical port the function is attached to. It also gives an indication of the physical function a virtual function is attached to.
  - uid The user identifier (UID) may be defined as part of the machine configuration or the z/VM or KVM guest configuration. If the accompanying uid_is_unique attribute is 1 the platform guarantees that the UID is unique within that instance and no devices with the same UID can be attached during the lifetime of the system.
  - uid_is_unique Indicates whether the user identifier (UID) is guaranteed to be and remain unique within this Linux instance.
  - pfip/segmentX The segments determine the isolation of a function. They correspond to the physical path to the function. The more the segments are different, the more the functions are isolated.

Enumeration and hotplug

The PCI address consists of four parts: domain, bus, device and function, and is of this form: DDDD:BB:dd.f

- When not using multi-functions (norid is set, or the firmware does not support multi-functions):
  - There is only one function per domain.
  - The domain is set from the zPCI function's UID as defined during the LPAR creation.
- When using multi-functions (norid parameter is not set), zPCI functions are addressed differently:
  - There is still only one bus per domain.
  - There can be up to 256 functions per bus.
  - The domain part of the address of all functions for a multi-Function device is set from the zPCI function's UID as defined in the LPAR creation for the function zero.
  - New functions will only be ready for use after the function zero (the function with devfn 0) has been enumerated.