Linux on Z and LinuxONE

Configuring Crypto Express Adapters for KVM Guests
August 2019

SC34-7717-00

Note

Before using this information and the product it supports, read the information in "Notices" on page 13.

This edition applies to the Linux on Z Development stream, libvirt version, and QEMU release as available at the time of writing, and to all subsequent releases and modifications until otherwise indicated in new editions.

Contents

About this document....................v
Other publications for Linux on Z and LinuxONE....................v

Chapter 1. AP queues on IBM Z.................... 1

Chapter 2. Prerequisites and assumptions.................... 3

Chapter 3. Setup on the KVM host.................... 5
Free AP queues for use by KVM guests.................... 5
Create a mediated device with AP queues....................6
Assign a mediated device to a KVM guest....................8

Accessibility....................11

Notices....................13
Trademarks.................... 13

About this document

This document describes the steps you must perform on the KVM host and in a virtual server configuration to make AP queues of cryptographic adapters available to a KVM guest.

Other publications for Linux on Z and LinuxONE

You can find publications for Linux on Z and LinuxONE on IBM® Knowledge Center.

These publications are available on IBM Knowledge Center at www.ibm.com/support/knowledgecenter/linuxonibm/liaaf/lnz_r_lib.html

- *Device Drivers, Features, and Commands*
- *Using the Dump Tools*
- *Running Docker Containers on IBM Z*, SC34-2781
- *KVM Virtual Server Quick Start*, SC34-2753
- *KVM Virtual Server Management*, SC34-2752
- *How to use FC-attached SCSI devices with Linux on z Systems®*, SC33-8413
- *libica Programmer's Reference*, SC34-2602
- *Exploiting Enterprise PKCS #11 using openCryptoki*, SC34-2713
- *Secure Key Solution with the Common Cryptographic Architecture Application Programmer's Guide*, SC33-8294
- *Pervasive Encryption for Data Volumes*, SC34-2782
- *How to set an AES master key*, SC34-7712
- *Troubleshooting*, SC34-2612
- *Kernel Messages*, SC34-2599
- *How to Improve Performance with PAV*, SC33-8414
- *How to Set up a Terminal Server Environment on z/VM*, SC34-2596

Chapter 1. AP queues on IBM Z

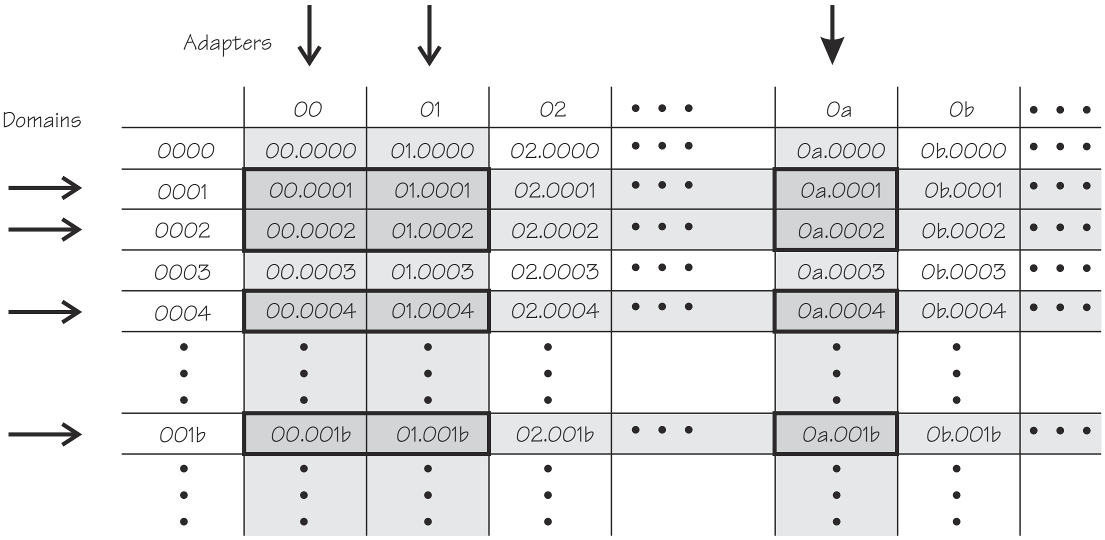
On IBM Z®, you assign cryptographic adapter resources in the form of AP queues.

IBM Z cryptographic adapters are partitioned into multiple cryptographic domains, each with its own state, including its own master key. A specific domain on a specific adapter is called an *AP queue*. In effect, an AP queue is a virtual cryptographic adapter.

AP queues are usually identified by expressions of the form *<adapter_ID>.<domain_ID>*, where *<adapter_ID>* is a two-digit ID for the cryptographic adapter in hexadecimal notation and *<domain_ID>* is a four-digit ID for the domain in hexadecimal notation. For example, `0a.001b` denotes domain 27 on adapter 10.

The number of available AP queues in a particular mainframe environment depends on the number of installed cryptographic adapters and on the number of domains that are supported by each adapter. To provide for redundancy and workload balancing, typical environments include multiple adapters.

Generally, AP queues are assigned as a matrix of intersecting adapters and domains. You use the Support Element (SE) to assign adapters and domains to an LPAR. For example, configuring adapters `00`, `01`, and `0a` and usage domains `0001`, `0002`, `0004`, and `001b` assigns 12 AP queues to the LPAR, as illustrated in Figure 1 on page 1.



	Adapters	↓	↓			↓		
Domains		00	01	02	• • •	0a	0b	• • •
	0000	00.0000	01.0000	02.0000	• • •	0a.0000	0b.0000	• • •
→	0001	00.0001	01.0001	02.0001	• • •	0a.0001	0b.0001	• • •
→	0002	00.0002	01.0002	02.0002	• • •	0a.0002	0b.0002	• • •
	0003	00.0003	01.0003	02.0003	• • •	0a.0003	0b.0003	• • •
→	0004	00.0004	01.0004	02.0004	• • •	0a.0004	0b.0004	• • •
	⋮	⋮	⋮	⋮		⋮	⋮	
→	001b	00.001b	01.001b	02.001b	• • •	0a.001b	0b.001b	• • •
	⋮	⋮	⋮	⋮		⋮	⋮	

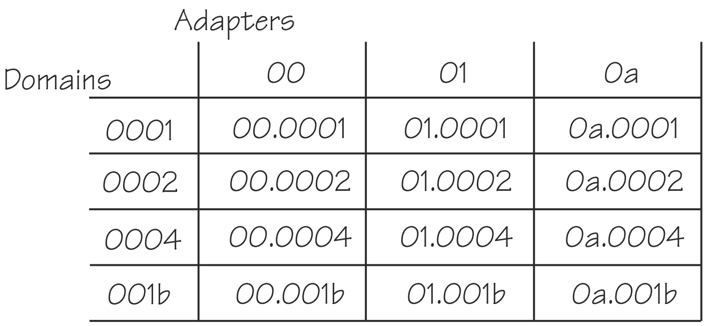
Figure 1. Matrix of adapters and domains with AP queues as intersections

Use the **`lszcrypt`** command to list the available AP queues.

```
# lszcrypt
CARD.DOMAIN TYPE    MODE         STATUS  REQUESTS
--------------------------------------------------
00          CEX6A   Accelerator  online         0
00.0001     CEX6A   Accelerator  online         0
00.0002     CEX6A   Accelerator  online         0
00.0004     CEX6A   Accelerator  online         0
00.001b     CEX6A   Accelerator  online         0
01          CEX6C   CCA-Coproc   online        31
01.0001     CEX6C   CCA-Coproc   online        10
01.0002     CEX6C   CCA-Coproc   online         7
01.0004     CEX6C   CCA-Coproc   online         9
01.001b     CEX6C   CCA-Coproc   online         5
0a          CEX6P   EP11-Coproc  online         0
0a.0001     CEX6P   EP11-Coproc  online         0
0a.0002     CEX6P   EP11-Coproc  online         0
0a.0004     CEX6P   EP11-Coproc  online         0
0a.001b     CEX6P   EP11-Coproc  online         0
```

The adapter mode is an LPAR setting that is configured on the SE, along with the assignment of adapters and domains to the LPAR. Adapters can be configured as CCA or EP11 coprocessors, or as accelerators. The mode of an adapter applies to all domains of the adapter.

The configuration scenario in Chapter 3, “Setup on the KVM host,” on page 5 is based on the LPAR configuration of Figure 1 on page 1. For the scenario, only the 12 AP queues that are available to the LPAR require consideration.



Domains \ Adapters	00	01	0a
0001	00.0001	01.0001	0a.0001
0002	00.0002	01.0002	0a.0002
0004	00.0004	01.0004	0a.0004
001b	00.001b	01.001b	0a.001b

Figure 2. Matrix of AP queues configured for an LPAR

Chapter 2. Prerequisites and assumptions

Your IBM Z hardware and KVM host must support AP queues for KVM guests.

Hardware requirements

- IBM Z hardware with one or more cryptographic adapters.

KVM host requirements

- The KVM host must be a Linux instance in LPAR mode, in an LPAR that has access to one or more AP queues.
- The Linux instance must support the vfio-ap device driver. This device driver was integrated into the upstream kernel with Linux 4.20.
- s390-tools upstream version 2.7 or later.
- QEMU upstream version 3.1 or later.
- libvirt upstream version 4.9 or later.

Distributions include the required modules and packages as of the following versions:

- Red Hat Enterprise Linux 8.0
- SUSE Linux Enterprise Server 15 SP1
- Ubuntu Server 18.04 LTS

KVM guest requirements

The KVM guest can be any distribution that is supported on IBM Z hardware.

Chapter 3. Setup on the KVM host

You must set up both the KVM host and the virtual server configuration of the KVM guest.

Procedure

1. “Free AP queues for use by KVM guests” on page 5
2. “Create a mediated device with AP queues” on page 6
3. “Assign a mediated device to a KVM guest” on page 8

Free AP queues for use by KVM guests

By default, all AP queues that are available to a KVM host are controlled by the `zcrypt` device driver on the host, and so unavailable to guests.

About this task

In a common setup, the KVM host acts as a broker of AP queues for its guests without using AP queues itself. You use two bit masks in sysfs, `/sys/bus/ap/apmask` and `/sys/bus/ap/aqmask`, to manage host control of AP queues. This procedure describes a fast path for freeing all queues for use by guests.

Procedure

1. Load the `vfio-ap` device driver.

   ```
   # modprobe vfio_ap
   ```

2. Free all adapters by specifying the following command:

   ```
   # echo 0x0 > /sys/bus/ap/apmask
   ```

3. Free all domains by specifying the following command:

   ```
   # echo 0x0 > /sys/bus/ap/aqmask
   ```

4. Optional: Issue **`lszcrypt -V`** to confirm your settings.
   With the verbose option, the **`lszcrypt`** command shows the AP queues that were freed from `zcrypt` control as being controlled by the `vfio-ap` device driver. The output of **`lszcrypt`** without the verbose option omits AP queues that are not controlled by `zcrypt`. The adapters themselves always remain under control of the applicable `zcrypt` submodule, `cex4card` in the example.

   Example:

```
# lszcrypt -V
CARD.DOMAIN TYPE    MODE         STATUS  REQUESTS  PENDING HWTYPE QDEPTH FUNCTIONS  DRIVER
--------------------------------------------------------------------------------------------
00          CEX6A   Accelerator  online         0        0     12     08 -MC-A-NF-  cex4card
00.0001     CEX6A   Accelerator  online         0        0     12     08 -MC-A-NF-  vfio_ap
00.0002     CEX6A   Accelerator  online         0        0     12     08 -MC-A-NF-  vfio_ap
00.0004     CEX6A   Accelerator  online         0        0     12     08 -MC-A-NF-  vfio_ap
00.001b     CEX6A   Accelerator  online         0        0     12     08 -MC-A-NF-  vfio_ap
01          CEX6C   CCA-Coproc   online        31        0     12     08 S--D--NF-  cex4card
01.0001     CEX6C   CCA-Coproc   online        10        0     12     08 S--D--NF-  vfio_ap
01.0002     CEX6C   CCA-Coproc   online         7        0     12     08 S--D--NF-  vfio_ap
01.0004     CEX6C   CCA-Coproc   online         9        0     12     08 S--D--NF-  vfio_ap
01.001b     CEX6C   CCA-Coproc   online         5        0     12     08 S--D--NF-  vfio_ap
0a          CEX6P   EP11-Coproc  online         0        0     12     08 -----XNF-  cex4card
0a.0001     CEX6P   EP11-Coproc  online         0        0     12     08 -----XNF-  vfio_ap
0a.0002     CEX6P   EP11-Coproc  online         0        0     12     08 -----XNF-  vfio_ap
0a.0004     CEX6P   EP11-Coproc  online         0        0     12     08 -----XNF-  vfio_ap
0a.001b     CEX6P   EP11-Coproc  online         0        0     12     08 -----XNF-  vfio_ap
```

If the `vfio-ap` device driver has not been loaded, the DRIVER column in the verbose output shows `-no-driver-` instead of `vfio_ap`.

Important: The mask specifications in sysfs do not persist across reboots.

What to do next

Optional: If the Linux instance of your KVM host needs to use AP queues, assign these AP queues to the host before you proceed to assign any queues to guests. Follows these steps to assign an AP queue to the KVM host:

1. Write the numerical value of the AP queue's adapter ID, prefixed with a plus sign (+), to `/sys/bus/ap/apmask`. Repeat this step to assign AP queues from multiple adapters to the KVM host.
2. Write the numerical value of the AP queue's domain ID, prefixed with a plus sign (+), to `/sys/bus/ap/aqmask`. Repeat this step to assign AP queues from multiple domains to the KVM host.

For example, to assign AP queue `00.001b` to the KVM host, issue:

```
# echo +0x0 > /sys/bus/ap/apmask
# echo +0x1b > /sys/bus/ap/aqmask
```

For more details about setting the adapter and domain masks, see *Device Drivers, Features, and Commands*.

AP queues that are not assigned to the host can now be configured for KVM guests.

Create a mediated device with AP queues

KVM guests access AP queues through an AP Virtual Function I/O (VFIO) mediated device. The configuration of the mediated device defines the AP configuration of the KVM guest to which it is assigned.

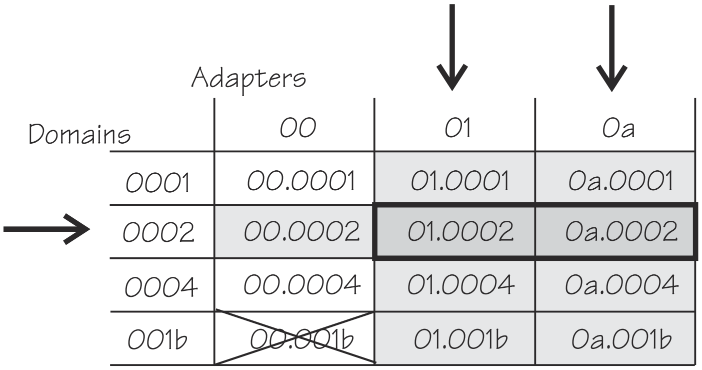
About this task

In the steps that follow, a mediated device is first created, then adapters and domains are configured for the device. The adapter and domain specifications define a matrix of AP queues. After the mediated device is included in a KVM virtual server configuration, these AP queues become available to the guest that runs in the virtual server.

The examples in the steps that follow assume that 11 AP queues are available. These AP queues correspond to intersections of 3 adapters with IDs `00`, `01`, and `0a` and 4 domains with IDs `0000`, `0001`, `0002`, and `001b`, omitting `00.001b`.

According to the example in “Free AP queues for use by KVM guests” on page 5, 00.001b has already been assigned to the KVM host and is no longer available to guests.

In the example, AP queues 01.0002 and 0a.0002 are to be assigned to a mediated device. The following figure illustrates how this assignment results from a specification of two adapters and a domain.



	Adapters		
Domains	00	01	0a
0001	00.0001	01.0001	0a.0001
0002	00.0002	01.0002	0a.0002
0004	00.0004	01.0004	0a.0004
001b	~~00.001b~~	01.001b	0a.001b

Figure 3. Matrix of AP queues available to KVM guests, omitting a queue that is used by the host

Procedure

1. Generate a UUID as an identifier for the mediated device.

   Example:

   ```
   # uuidgen
   4b0518fd-9237-493f-93c8-c5597f8006a3
   ```

2. Create the device by writing the UUID to /sys/devices/vfio_ap/matrix/mdev_supported_types/vfio_ap-passthrough/create

   Example:

   ```
   # echo 4b0518fd-9237-493f-93c8-c5597f8006a3 \
   > /sys/devices/vfio_ap/matrix/mdev_supported_types/vfio_ap-passthrough/create
   ```

   This command creates a mediated device that is represented by a sysfs directory /sys/devices/vfio_ap/matrix/*<device_id>*, where *<device_id>* is the UUID that was generated in the previous step.

3. Assign an adapter to the mediated device by writing the adapter ID, with a 0x prefix, to the device's assign_adapter sysfs attribute.

   Repeat this step to assign multiple adapters.

   Example: To assign adapters 01 and 0a:

   ```
   # echo 0x01 > /sys/devices/vfio_ap/matrix/4b0518fd-9237-493f-93c8-c5597f8006a3/assign_adapter
   # echo 0x0a > /sys/devices/vfio_ap/matrix/4b0518fd-9237-493f-93c8-c5597f8006a3/assign_adapter
   ```

4. Assign a domain to the mediated device by writing the domain ID, with a 0x prefix, to /sys/devices/vfio_ap/matrix/*<device_id>*/assign_domain

   Repeat this step to assign multiple domains.

   Example: To assign domain 0002:

   ```
   # echo 0x0002 > /sys/devices/vfio_ap/matrix/4b0518fd-9237-493f-93c8-c5597f8006a3/assign_domain
   ```

5. For each domain that you assigned in step “4” on page 7, assign a control domain, so you can manage your domains from the guest that uses the mediated device.

   Other than for z/VM® guests, usage domains on KVM guests are not automatically also control domains.

Example: To assign domain `0002` as a control domain:

```
# echo 0x0002 > /sys/devices/vfio_ap/matrix/4b0518fd-9237-493f-93c8-c5597f8006a3/assign_control_domain
```

6. Read the `matrix` attribute of the mediated device to confirm that the assignment of adapters and domains resulted in the intended AP queue assignment.

   Example:

   ```
   # cat /sys/devices/vfio_ap/matrix/4b0518fd-9237-493f-93c8-c5597f8006a3/matrix
   01.0002
   0a.0002
   ```

What to do next

You can repeat this procedure to create multiple mediated devices, but you must not assign a specific AP queue to multiple mediated devices. You can use the attributes of the mediated device to investigate and control the device's properties.

```
ls -1 /sys/devices/vfio_ap/matrix/<device_id>
assign_adapter
assign_control_domain
assign_domain
control_domains
driver
iommu_group
matrix
mdev_type
power
remove
subsystem
uevent
unassign_adapter
unassign_control_domain
unassign_domain
```

In particular, you can write to the `assign_*` and `unassign_*` attributes to modify the mediated device, and you can use the `remove` attribute to remove the mediated device.

Important: The mask specifications in sysfs do not persist across reboots.

Assign a mediated device to a KVM guest

Add the mediated device to the domain configuration-XML of the guest.

Procedure

1. On the KVM host, open the configuration-XML of the guest in edit mode.

   Example:

   ```
   # virsh edit guest1
   ```

2. Add an entry for the mediated device to the <devices> element within the guest XML.

   Use the following template, replacing *<device_id>* with the UUID that identifies the mediated device in sysfs.

   ```
        <hostdev mode='subsystem' type='mdev' managed='no' model='vfio-ap'>
          <source>
            <address uuid='<device_id>'/>
          </source>
        </hostdev>
   ```

   Example:

```
<domain type='kvm'>
   <name>guest1</name>
   ...
   <devices>
    ...
    <hostdev mode='subsystem' type='mdev' managed='no' model='vfio-ap'>
       <source>
         <address uuid='99e714ec-8eee-40fd-a26e-80ff3b1a2564'/>
       </source>
     </hostdev>
     ...
   </devices>
   ...
</domain>
```

3. Save and close the XML document.

What to do next

After the virtual server is started, you can use the **lszcrypt** command on the guest to display the AP queues. The queues are controlled by the zcrypt device driver of the guest and can be used as usual.

```
# lszcrypt -V
CARD.DOMAIN TYPE    MODE         STATUS  REQUESTS  PENDING HWTYPE QDEPTH FUNCTIONS  DRIVER
---------------------------------------------------------------------------------------------
01          CEX6C   CCA-Coproc   online        31        0     12     08 S--D--NF-  cex4card
01.0002     CEX6C   CCA-Coproc   online        10        0     12     08 S--D--NF-  cex4queue
0a          CEX6P   EP11-Coproc  online         0        0     12     08 -----XNF-  cex4card
0a.0002     CEX6P   EP11-Coproc  online         0        0     12     08 -----XNF-  cex4queue
```

Accessibility

Accessibility features help users who have a disability, such as restricted mobility or limited vision, to use information technology products successfully.

Documentation accessibility

The Linux on Z and LinuxONE publications are in Adobe Portable Document Format (PDF) and should be compliant with accessibility standards. If you experience difficulties when you use the PDF file and want to request a Web-based format for this publication send an email to eservdoc@de.ibm.com or write to:

IBM Deutschland Research & Development GmbH
Information Development
Department 3282
Schoenaicher Strasse 220
71032 Boeblingen
Germany

In the request, be sure to include the publication number and title.

When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

IBM and accessibility

See the IBM Human Ability and Accessibility Center for more information about the commitment that IBM has to accessibility at

```
www.ibm.com/able
```

Notices

This information was developed for products and services offered in the U.S.A. IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

The licensed program described in this information and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement, or any equivalent agreement between us.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at
www.ibm.com/legal/copytrade.shtml

SC34-7717-00