

# kdump - Recommendations for Linux® on IBM Z

—  
Solution Assurance

André Wild, David Stark

v1.4, March 2022

**IBM Z**



# Notices and disclaimers

© 2020 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.

## **U.S. Government Users Restricted Rights – use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.**

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed "as is" without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.

**Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.**

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

# Notices and disclaimers, *continued*

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.**

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com and IBM Z, IBM z15, and z/VM are trademarks of International Business Machines Corporation, 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](http://www.ibm.com/legal/copytrade.shtml)

# Agenda

## **kdump**

- **kernel command line**
- **page filter with makedumpfile**
- **hardware-accelerated compression**
- **configuration**
- **local storage setup**
- **remote storage setup**
- **debugging**

# kdump kernel command line

## crashkernel size

The crashkernel size should be properly set to avoid out of memory exception.

### Red Hat

supports `crashkernel=auto` which scales with changes like memory.

### SUSE

has a `kdumptool` which can calculate the estimate crashkernel size. Both variants are not bullet proof and need manual tweaking.

## crashkernel syntax

`crashkernel=<range1>:<size1>[,<range2>:<size2>,...][@offset]`  
range=start-[end] 'start' is inclusive and 'end' is exclusive.

Use this to scale the crashkernel size if `crashkernel=auto` does not suite your needs. The `@offset` value has to be greater or equal to your reserved memory size because the dump kernel will be relocated before beeing executed.

## Exclude unused devices

`cio_ignore` from `s390-utils` can be used to generate the kernel command line to exclude unused devices. This can help to reduce the memory consumption during the `kdump` boot process.

```
$ cio_ignore -k -u  
cio_ignore=all,!1234-123f,!1700
```

Example

```
KDUMP_COMMANDLINE_APPEND="cio_ignore=all,!1234-123f,!1700"
```

## LUKS encrypted devices

Using memory-hard functions might require adjustment of the crashkernel size. For example the default Aragon2 requires 1GB of extra memory to work properly.



`crashkernel=auto`



`/etc/sysconfig/kdump`



Low value of `#kdumptool calibrate` command



`/etc/sysconfig/kdump`

# kdump page filter with makedumpfile

dump level	zero page	non-private cache	private cache	user data	free page
0					
1	X				
2		X			
3	X	X			
4		X	X		
5	X	X	X		
6		X	X		
7	X	X	X		
8				X	
9	X			X	
10		X		X	
11	X	X		X	
12		X	X	X	
13	X	X	X	X	
14		X	X	X	
15	X	X	X	X	
16					X
17	X				X
18		X			X
19	X	X			X
20		X	X		X
21	X	X	X		X
22		X	X		X
23	X	X	X		X
24				X	X
25	X			X	X
26		X		X	X
27	X	X		X	X
28		X	X	X	X
29	X	X	X	X	X
30		X	X	X	X
31	X	X	X	X	X

source: man makedumpfile

## makedumpfile

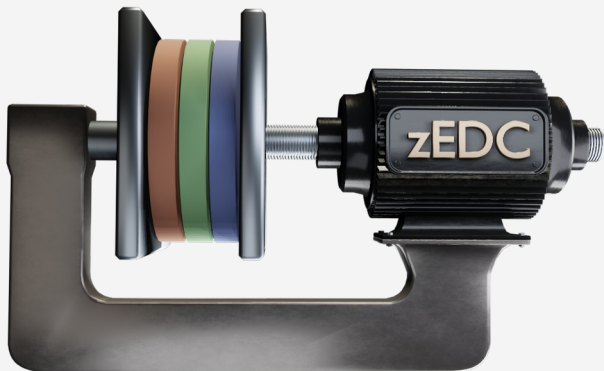
- default application used by kdump to copy the /proc/vmcore file
- supports dumplevel option to filter pages
  - should be used to reduce the size of your dump

## dump level

- consists of 5 bits
- each bit indicates a page type to discard
- add up it's value to calculate the appropriate dump level

Level value	Page filter
1	Exclude the pages filled with zero.
2	Exclude the non-private cache pages.
4	Exclude all cache pages.
8	Exclude the user process data pages.
16	Exclude the free pages.

# kdump hardware-accelerated compression



```
core_collector makedumpfile -c ....
```



```
KDUMP_DUMPFORMAT="compressed"
```

## + Pros

- Faster with compression turned on
- Less storage consumption
- Less network usage for remote dumps
- It's easier to share small dumps

## Requirements

- makedumpfile (or any other zlib/dftcc supported tool)
  - compression option -c for zlib compression
- zlib package

## Hardware supported compression

- zlib with dftcc support
- Introduced with IBM z15 and above

## Verify dftcc is supported

```
$ grep dflt /proc/cpuinfo
```

```
$ objdump -d $(ldd $(command -v makedumpfile) | awk '$1 ~ /lib\./ { print $3 }') | grep dftcc
```

## - Cons

- zlib/deflate is currently not the default in RHEL
- zlib/deflate will always use hardware-accelerated compression
  - This is because makedumpfile uses compression level 1 which uses dftcc by default

# kdump configuration

## How to configure kdump

### Steps to perform changes

1. Make changes to the configuration
2. restart kdump service

```
sudo systemctl restart kdump
```

3. Verify initrd/initramfs for kdump has been recreated

```
ls -lt /boot
```

Now you should see a new initrd/initramfs file for kdump

## How to verify your configuration

### Trigger kernel panic

Enable all sysrq

```
# echo 1 > /proc/sys/kernel/sysrq
```

Trigger kernel panic

```
# echo c > /proc/sysrq-trigger
```

### Verify Dump exists

The default location is /var/crash. This could have been changed by your kdump configuration.



/etc/kdump.conf  
/etc/sysconfig/kdump

Remark: /etc/sysconfig/kdump does not support all options from SUSE



[Documentation Link](#)



/etc/sysconfig/kdump



[Documentation Link](#)



# kdump local storage setup

## Pros

- No network protocol overhead
- Filtering extra pages reduces the overall size
- Compression can help to reduce the size
  - With z15 and dfltcc enabled zlib it's a lot faster

## Cons

- Storage can only be used for one system
- Dumps can pollute the system
  - SUSE is able to reduce such impacts (KDUMP\_KEEP\_OLD\_DUMPS)

## Use case

- Default setup for any type of system
- No shared storage is available for dumps

## Hints

### Red Hat

- Uses lzo compression per default which is not as fast as zlib dfltcc
- No built-in feature to set the amount of kept dumps
  - This could be achieved with the kdump pre scripts see man kdump.conf

### SUSE

- Use the internal dump rollover feature



```
default reboot
path /var/crash
core_collector makedumpfile --message-level 7 -c -F -d 31
```



```
KDUMP_IMMEDIATE_REBOOT="yes"
KDUMP_SAVEDIR="/var/crash"
KDUMP_KEEP_OLD_DUMPS="2"
KDUMP_VERBOSE="3"
KDUMP_FREE_DISK_SIZE="64"
KDUMP_DUMPLEVEL="31"
KDUMP_DUMPFORMAT="compressed"
KDUMP_CONTINUE_ON_ERROR="true"
KDUMP_COPY_KERNEL="yes"
```

# kdump remote storage setup

## Pros

- One setup for all participating systems
- Overall less storage consumption/reservation
- Decentralized dump rollover and storage provisioning
- Improved performance and reduced dump size using compression

## Cons

- Network dumps could impact network performance
- The network protocol overhead reduces performance slightly
- Network dumps should be secured in order to protect sensitive information

## Use case

- Scales better with large environments
- Reduces storage costs because storage is not permanently reserved
- Reduces service downtime - dumps are already located on another system

## Hints

### Red Hat

- With z15 and above zlib/dfltc is faster and creates smaller dumps in comparison to lzo which is currently the default

### SUSE

- You should reduce/increase the default number of local dumps that should be kept to suit your needs

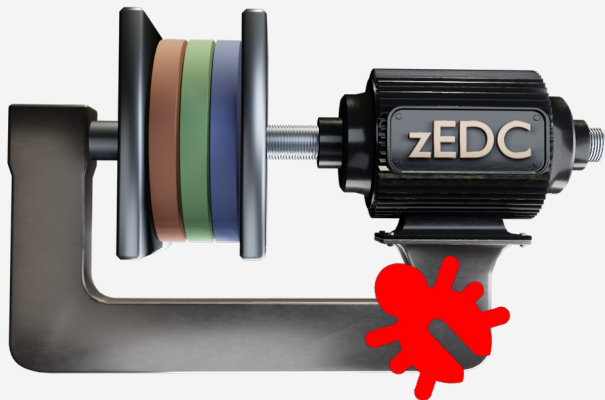


```
default reboot
path /kdump/hostname
ssh user@server
sshkey /root/.ssh/id_rsa.kdump
core_collector makedumpfile --message-level 7 -c -F -d 31
```



```
KDUMP_IMMEDIATE_REBOOT="yes"
KDUMP_SAVEDIR="ssh://user@hostname/kdump/hostname"
KDUMP_KEEP_OLD_DUMPS="2"
KDUMP_VERBOSE="3"
KDUMP_DUMPLEVEL="31"
KDUMP_DUMPFORMAT="compressed"
KDUMP_CONTINUE_ON_ERROR="true"
KDUMP_COPY_KERNEL="yes"
KDUMP_NETCONFIG="auto"
KDUMP_NET_TIMEOUT="30"
KDUMP_SSH_HOST_KEY="AAAAE2V..."
KDUMP_SSH_IDENTITY="/root/.ssh/id_rsa.kdump"
```

# kdump debugging configuration



```
default shell
core_collector ... -D --message-level 31
failure_action shell
```



```
KDUMP_IMMEDIATE_REBOOT="no"
KDUMP_VERBOSE="8"
KDUMP_CONTINUE_ON_ERROR="false"
MAKEDUMPFIL_OPTIONS="-D --message-level 31"
```



## Notes

Troubleshooting kdump can be done with the following options and tools.

- **Increase verbosity**
  - Red Hat: `core_collector makedumpfile --message-level X ...`
  - SUSE: `KDUMP_VERBOSE | MAKEDUMPFIL_OPTIONS`
- **Open shell on error or after dump**
  - Red Hat: `default shell | failure_action shell`
  - SUSE: `KDUMP_IMMEDIATE_REBOOT`
- **Disable continue on error**
  - Red Hat: `failure_action shell`
  - SUSE: `KDUMP_CONTINUE_ON_ERROR`
- **Use pre/post kdump actions to copy logs or open a shell**
  - Red Hat: `kdump_pre | kdump_post`
  - SUSE: `KDUMP_PRESCRIPT | KDUMP_POSTSCRIPT`

### View logs

```
$ systemctl status kdump
$ journalctl -fu kdump
```

# Resources

- Linux on IBM Z and IBM LinuxONE
  - Official homepage: <http://www.ibm.com/systems/z/os/linux>
  - Using the dump tools: <https://www.ibm.com/docs/en/linux-on-systems?topic=troubleshooting-using-dump-tools>
  - IBM z15 Hardware Compression  
<https://community.ibm.com/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=95d7c2ae-190a-5de4-b0f5-a210200eb821&forceDialog=0>
- Red Hat
  - Documentation:  
[https://access.redhat.com/documentation/en-us/red\\_hat\\_enterprise\\_linux/8/html/system\\_design\\_guide/installing-and-configuring-kdump\\_system-design-guide](https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/8/html/system_design_guide/installing-and-configuring-kdump_system-design-guide)
- SUSE
  - Documentation: <https://documentation.suse.com/sles/15-SP3/html/SLES-all/cha-tuning-kexec.html>
- Ubuntu
  - Documentation: <https://ubuntu.com/server/docs/kernel-crash-dump>
- Linux Kernel
  - kdump: <https://www.kernel.org/doc/Documentation/kdump/kdump.txt>
  - kexec: <https://man7.org/linux/man-pages/man8/kexec.8.html>
  - crash: <https://man7.org/linux/man-pages/man8/crash.8.html>
  - makedumpfile: <https://linux.die.net/man/8/makedumpfile>
- Logos & Icons
  - <https://www.redhat.com/de/about/brand/standards/logo>
  - <https://brand.suse.com>
  - <https://www.iconsdb.com/red-icons/bug-3-icon.html>