GSE | GUIDE SHARE EUROPE

UK Region

Getting the most out of the Latest Features in Linux and KVM on IBM Z and LinuxONE

Stefan Raspl
Principal Product Manager
Linux & Virtualization on IBM Z and LinuxONE

GSE UK Conference

2024 Charities

Every year, speakers and delegates at the GSE UK Conference donate to the charities we help. This year is no exception and we aim to collect donations for two worthy causes:
Air Ambulance UK and Muscular Dystrophy UK.

Please consider showing your appreciation by kindly donating a small sum to our charities this year!!

AIR AMBULANCES UK
SUPPORTING AIR AMBULANCE CHARITIES

MUSCULAR DYSTROPHY UK
OUR MUSCLES MATTER

Contents

- **Latest Linux on IBM Z & LinuxONE Features and Packages**
- **Latest KVM on IBM Z & LinuxONE News**

Latest on IBM Z & LinuxONE Features and Packages

Removal of 32-bit Support

LPAR z/VM KVM

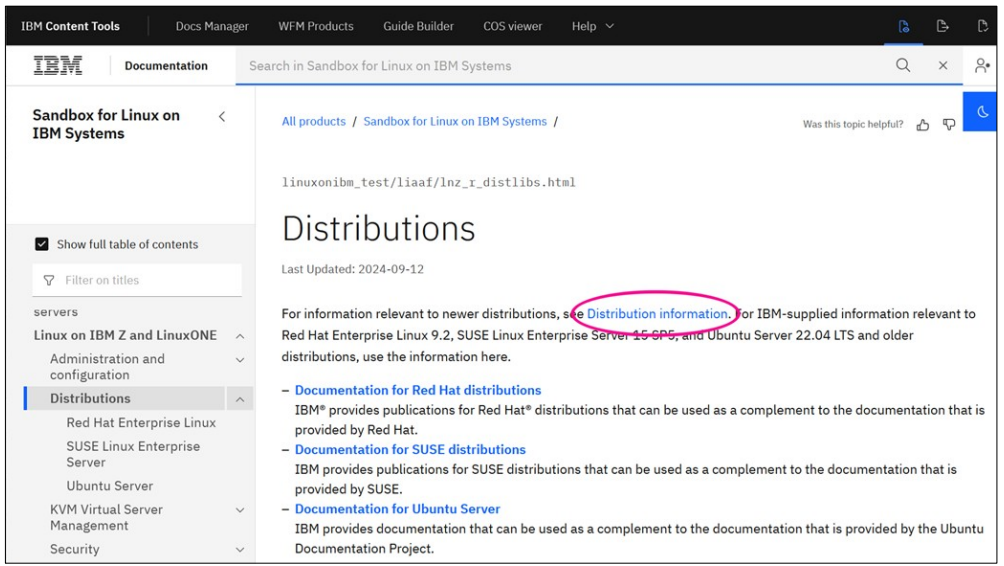
- **What it does**: Remove the ability to run ***any*** 32-bit application in forthcoming Linux distributions
- **Why you should care**: If you are still deploying 32-bit applications in ***any*** shape or form, it is about time to plan for a migration
- If kernel still supports 32-bit, userspace can run in 32-bit, provided all necessary libraries are available in 32-bit, too!
- Ways to provide 32-bit libraries:
  - Use in distro (see `compat` packages)
  - Link statically
  - Package in containers

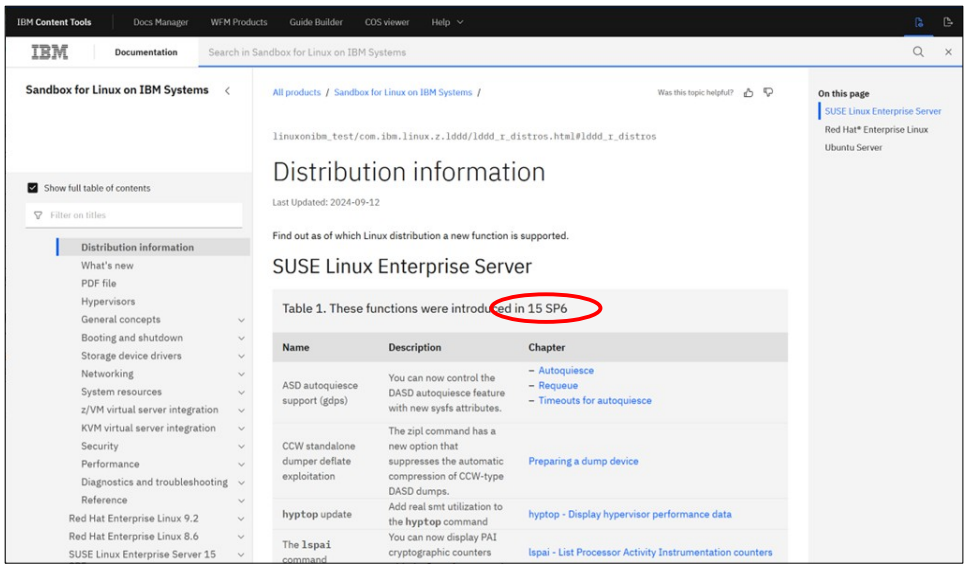
Distro	Properties	RHEL	SLES	Ubuntu
32-bit distro	Full 32-bit distro available	≤ RHEL 4	≤ SLES 9	-
64-bit distro	32-bit userspace	RHEL 5 RHEL 6 RHEL 7	-	-
	32-bit compat packages only	-	SLES 10 SLES 11 SLES 12	≤ Ubuntu 22.04
	32-bit compat kernel support, statically linked 32-bit & containers still work	RHEL 8 RHEL 9	SLES 15	-
	No 32-bit support at all	**RHEL next**	**SLES next**	**Ubuntu 24.04**

Fig. 1: *Distro support overview*

Linux on IBM Z – Content Design
Distribution information

On the Distributions page,
click Distribution information

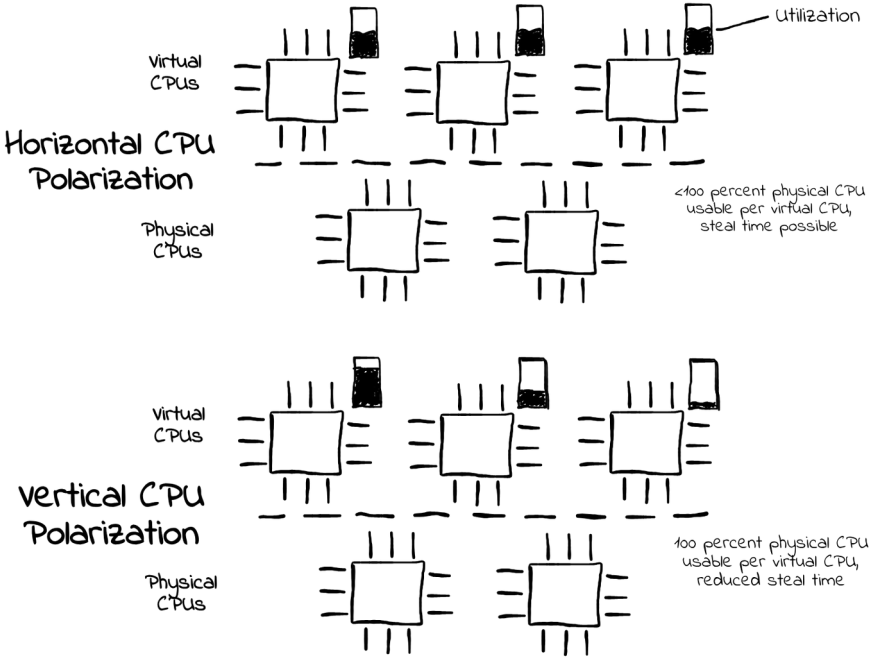




HiperDispatch Support aka Vertical CPU Polarization

- **What it does:** Prioritize process scheduling to CPUs with more consistent processing guarantees to avoid steal time
- **Why you should care:** Can yield substantial performance improvements for CPU-intensive workloads on highly utilized CECs
- Platform differentiates between *vertical high*, *medium* and *low* IFLs, with varying capacity grants
- Basically no steal time on *vertical high* IFLs
- Modifies the scheduler to prefer *vertical highs* and *mediums* for CPU-intensive workloads
- Workloads running large numbers of small tasks might perform better with horizontal CPU polarization
- **How to use:**
  - Enabled by default
  - Use sysctl `s390.hiperdispatch` to enable or disable:

```
sysctl -w s390.hiperdispatch=[0|1]
```



Tunables:

- `/sys/devices/system/cpu/hd_steal_threshold`
- `/sys/devices/system/cpu/hd_delay_factor`
  Steal time evaluation period. Reducing this value improves responsiveness to changes in workload behavior. Increasing it delays reaction to sudden changes in steal time.

Installation Assistant for Linux on IBM Z & LinuxONE

- **What it does**: Creating configuration files for starting Linux on IBM Z and LinuxONE installations
- **Why you should care**: Writing parameter files can be a challenge, with bugs triggering cycles with lengthy turnaround times
- Generates installer parameter files for the latest RHEL and SLES Linux distributions
- Supports OSA and PCI networking devices, IPv4/v6, and VLAN installations
- Provides easy to follow step by step instructions and context help
- Access at https://ibm.github.io/liz/

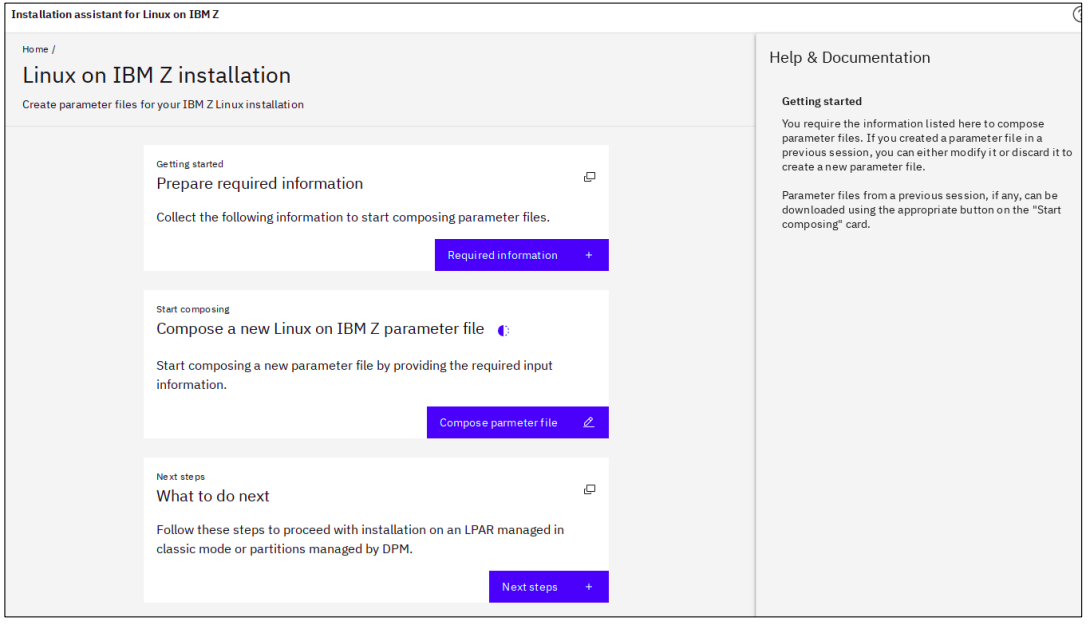
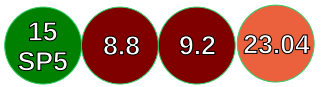


Fig. 1: *Landing page of the installation assistant*

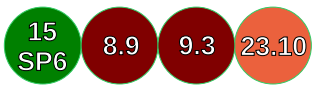
Secure Boot for ECKD DASD

LPAR z/VM

- **What it does:** Linux can boot from ECKD DASD in Secure Boot mode
- **Why you should care:** Secure Boot is a prerequisite for the NIAP certification, and deployment of Linux in environments with extra high security requirements
- **What you need:**
  - IBM z16 with GA1.5 firmware

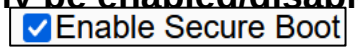


  - **For Basic boot support:**
    - s390-tools v2.25
  - **For Reboot and dump support:**



    - s390-tools v2.26
    - Linux kernel v6.2

- **How to use it:** With the new support, Linux DASDs contain 2 types of boot loader:
  - CCW IPL: Standard boot
  - LD-IPL (“List-Directed IPL”): Supports Secure Boot
- **Note: Secure Boot can only be enabled/disabled on the HMC Load panel**



- **zipl will always install both boot loader types:**

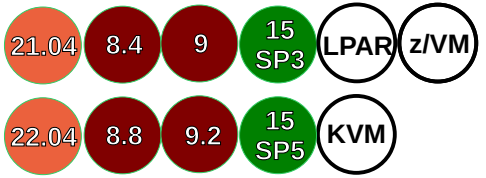
```
$ zipl
Using config file '/etc/zipl.conf'
...
Preparing boot device for CCW- and LD-IPL: dasda (1234).
Done.
```

- **For reboot, IPL-type must be chosen manually**
  - `chreipl eckd` for DASD LD-IPL with Secure Boot support
  - `chreipl ccw` for DASD CCW-IPL with standard boot

```
$ chreipl eckd 0.1.1002
Re-IPL type: eckd
Device:      0.1.1002
bootprog:    0
br_chr:      auto
Bootparm:    ""
Loadparm:    ""
clear:       0
```

```
$ chreipl ccw 0.1.1002
Re-IPL type: ccw
Device:      0.1.1002
Loadparm:    ""
clear:       0
```

Co-Location: SMC-Dv2



- **What it does**: Provides acceleration for TCP traffic
- **Why you should care**: v2 lifts limitations and greatly simplifies usage
- **Recap**
  - **Shared Memory Communications – Direct** provides intra-CEC acceleration for TCP traffic using *Internal Shared Memory* (ISM) devices
  - **Superior performance** (low latency, high throughput) at reduced CPU consumption
  - *However, SMC-Dv1 had limitations*:
    - Peers must be in **same IP subnet**
    - Devices need to be **paired using PNET IDs**
- **SMC-Dv2**
  - Peers can be in ***any* IP subnet**
  - No PNET IDs required
    ⇒ **Simplified configuration!**
  - Requires z15 or LinuxONE III
  - As with SMC-Dv1: Full **z/OS compatibility**
- **New performance paper available here**

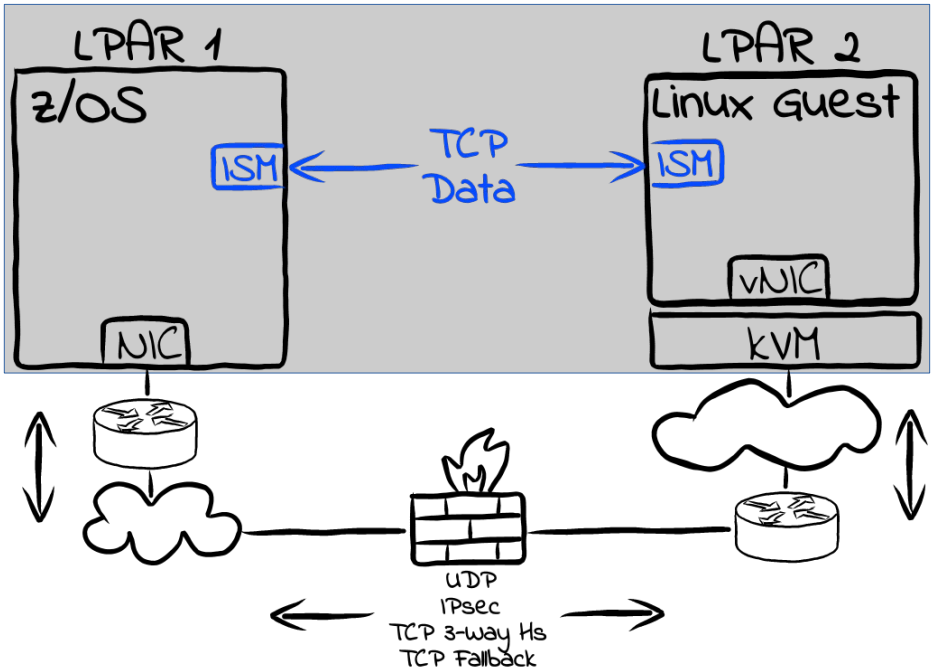


Fig.1: *Traffic flows with SMC-Dv2*

Eating our own dogfood: Leveraging SMC-Dv2 in IBM StorageScale

- **What we did**: Changed IBM Storage Scale to utilize SMC-Dv2 to benefit from superior performance and less CPU utilization
- **Why you should care**: Not only did this conversion yield strong results – it was also very easy to apply!

C++ code changes

- AF_INET sockets have been replaced with AF_SMC sockets for GPFS daemons: mmfsd and mmsdrserv (performance critical GPFS components)

```
socket(AF_INET, SOCK_STREAM, 0);
=>
socket(AF_SMC , SOCK_STREAM, 0);
```

- Other GPFS binaries, utilities, python/shell scripts are still using TCP

Python scripts changes

- New SMC-D Prerequisites Verification Tool: `tssmcdnodeverify`
- Enhanced `mmnetverify` tool: added SMC-D connections verification

Configuring IBM Storage Scale for SMC-Dv2 with Linux on Z nodes

1

Verify the Internal Shared Memory (ISM) device availability

```
# lspci | grep ISM
1014:00:00.0 Non-VGA unclassified device: IBM
Internal Shared Memory (ISM) virtual PCI device
```

2

Install `smc-tools` and `qclib` OS packages:

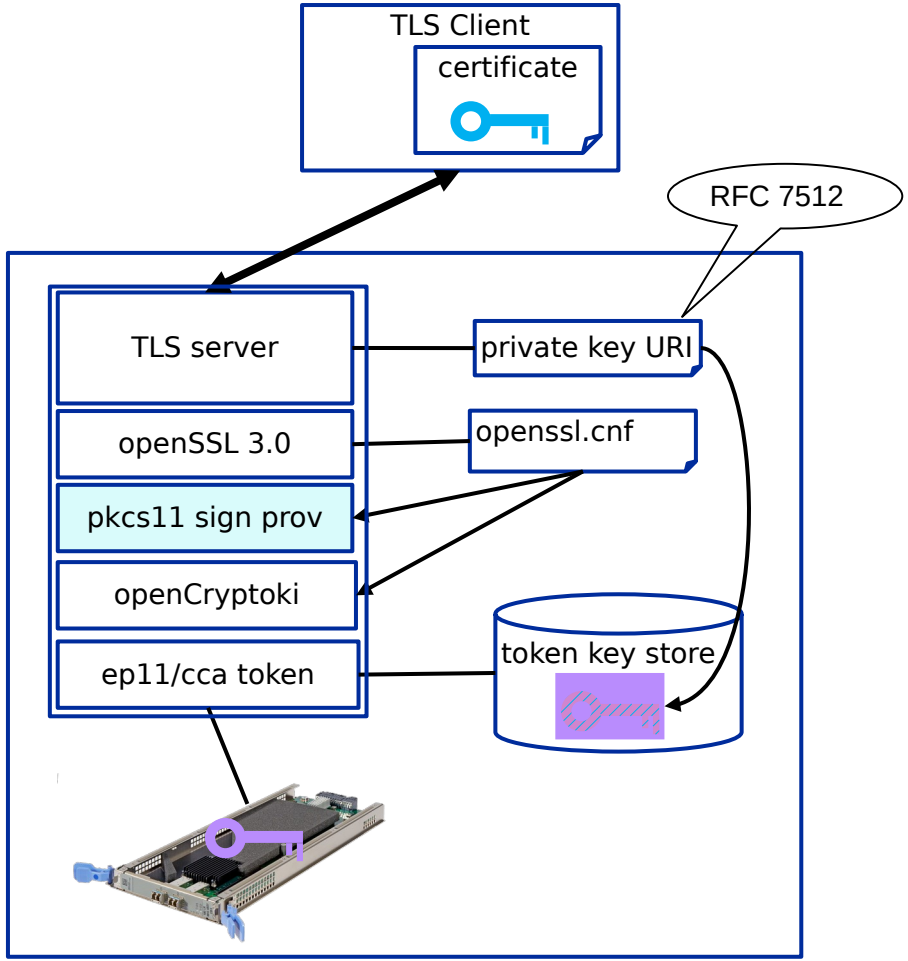
SLES: `zypper install smc-tools qclib`
RHEL: `dnf install smc-tools qclib`

3

Verify that SELINUX is set to "`permissive`" or "`disabled`"

SMC sockets are not included into the standard SELINUX policies and therefore SMC-D does not support SELINUX "enforcing" mode.

OpenSSL pkcs11 Signing Provider



15 SP6 | 23.10 | LPAR | z/VM | KVM

Problem

- a hacker who steals the private signing key of a TLS server can impersonate the TLS server

Solution

- protect the private signing key with an HSM

Note

- all other keys of a TLS connection are ephemeral and therefore less critical

Release 1.0 of openssl-pkcs11-sign-provider released on https://github.com/opencryptoki/openssl-pkcs11-sign-provider

p11sak from openCryptoki 3.21 supports key URIs

supports

- ECDSA
- RSA sign
- RSA decrypt

restriction

- process must not fork

Master Key Change Protocol for openCryptoki – so far

Disruptive Procedure!

CCA

Re-encipher token key repository of openCryptoki CCA token

1. *! Stop all processes using openCryptoki CCA token*
2. *Perform MK change on HSM*
3. *Use the `pkcscca` tool to re-encipher token keys*
4. *Restart processes using openCryptoki CCA token*

Note: CCA has 4 different MKs for different key types: (3)DES, AES, RSA, ECC

- ***Each MK can be changed independently***

Disruptive Procedure!

EP11

Re-encipher token key repository of openCryptoki ep11 token

1. *! Stop all processes using openCryptoki EP11 token*
2. *Commit new MK on HSM*
3. *Use the `pkcsep11_migrate` tool to re-encipher token keys*
4. *Activate new MK on HSM*
5. *Restart processes using openCryptoki EP11 token*

openCryptoki: Concurrent HSM Master Key Change

- **What it does**: Allow to concurrently change the master keys of an openCryptoki HSM token (e.g. CCA or EP11 token) concurrently to application linked to openCryptoki
- **Why you should care**: Avoid service outages for workloads that depend on HSM services (based on openCryptoki)
- Sample flow (applies to CCA & EP11):

TKE / HSM admin

- At TKE, load new HSM master key into adapter domains
- Inform OS admin that all new master keys are loaded
- At TKE, set the new HSM master keys. I.e., new HSM MKs become current HSM MK
- Inform OS admin that all master keys are now set

1

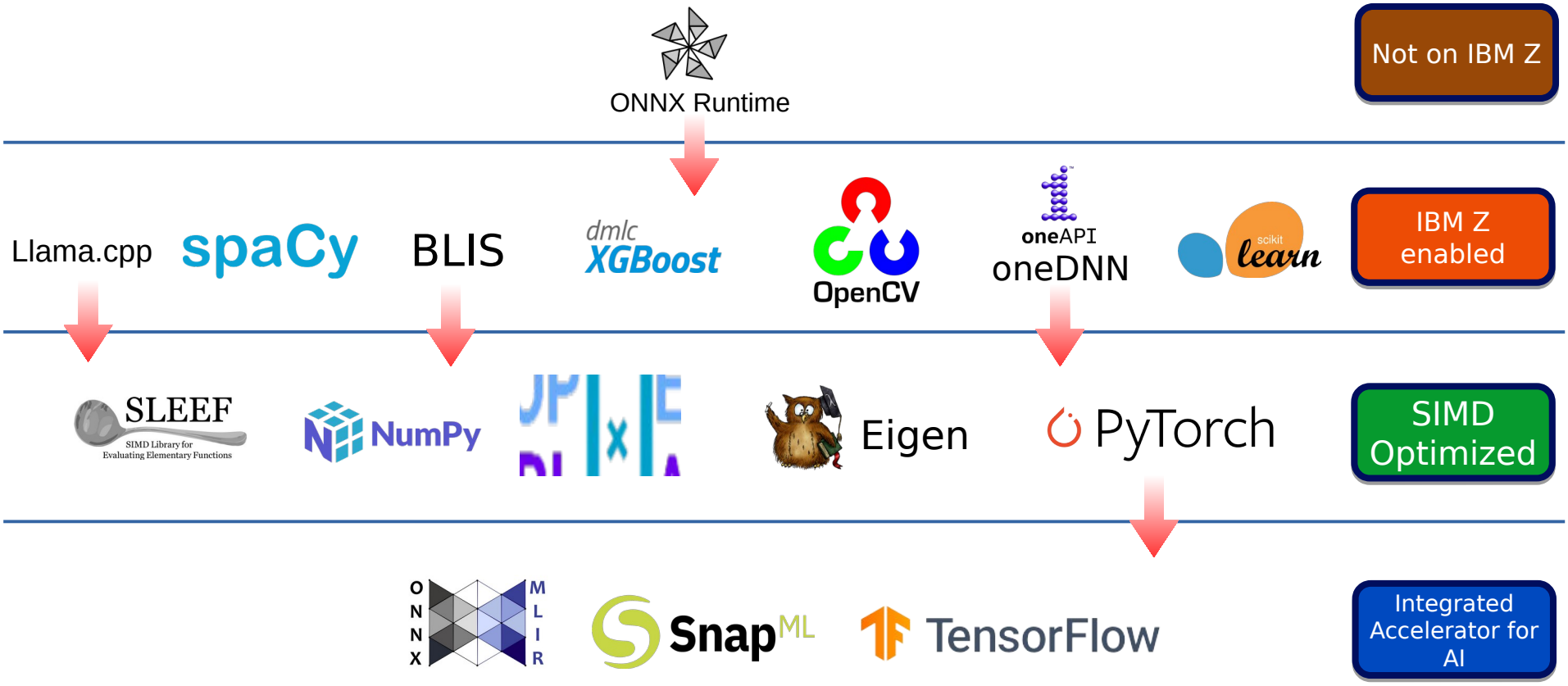
2

3

Linux / openCryptoki admin

- `pkcshsm_mk_change reencipher ...`
  - *initiates reenciphering process for all PKCS #11 keys objects used by all applications linked to openCryptoki with a set of specified APQNs and all token key objects for the specified APQNs*
- Inform HSM admin after reencipher process is complete
- `pkcshsm_mk_change finalize ...`
  - *activates re-enciphered keys*

AI Ecosystem – Hardware Exploitation



KVM for
Linux on IBM Z & LinuxONE

KVM Availability

KVM is available and supported in

- SLES12 SP2 and later
- RHEL 8 starting with RHEL 8.4 via Advanced Virtualization repository
- Ubuntu 16.04 and later

Community distributions with KVM support:

- Debian
- Fedora
- OpenSUSE

Documentation: KVM Virtual Server Management available here
Nov 2022 update now also covers

- Persistently configure VFIO mediated devices for both DASD and cryptographic resources
- Share parts of the KVM host file system with a virtual server (virtiofs)
- Dump automation improvements on the KVM host

Package versions

- Red Hat

	kernel	QEMU	Libvirt
RHEL 8.10	4.18	6.2	8.0
RHEL 9.4	5.14	8.2	10.0

- SUSE

	kernel	QEMU	Libvirt
SLES12 SP5	4.12	3.1	5.1
SLES15 SP4	6.4	8.2	10.0

- Ubuntu

	kernel	QEMU	Libvirt
16.04 ESM-only	4.4	2.5	1.3.1
18.04 LTS	4.15	2.11	4.0
20.04 LTS	5.4	4.2	6.0
22.04 LTS	5.15	6.2	8.0
24.04 LTS	6.8	8.2	10.0

KVM Hardware Support: Selecting the right CPU Model

```
<os>
    <type arch='s390x' machine='s390-ccw-virtio'>hvm</type>
    [...]
</os>
<cpu mode='host-model'/>
```

- E.g. z16 support provided by new model gen16a, enabling all z16 features per default
- Choose among the following CPU models:

Mode	Feature Set	Migration Safe	Syntax

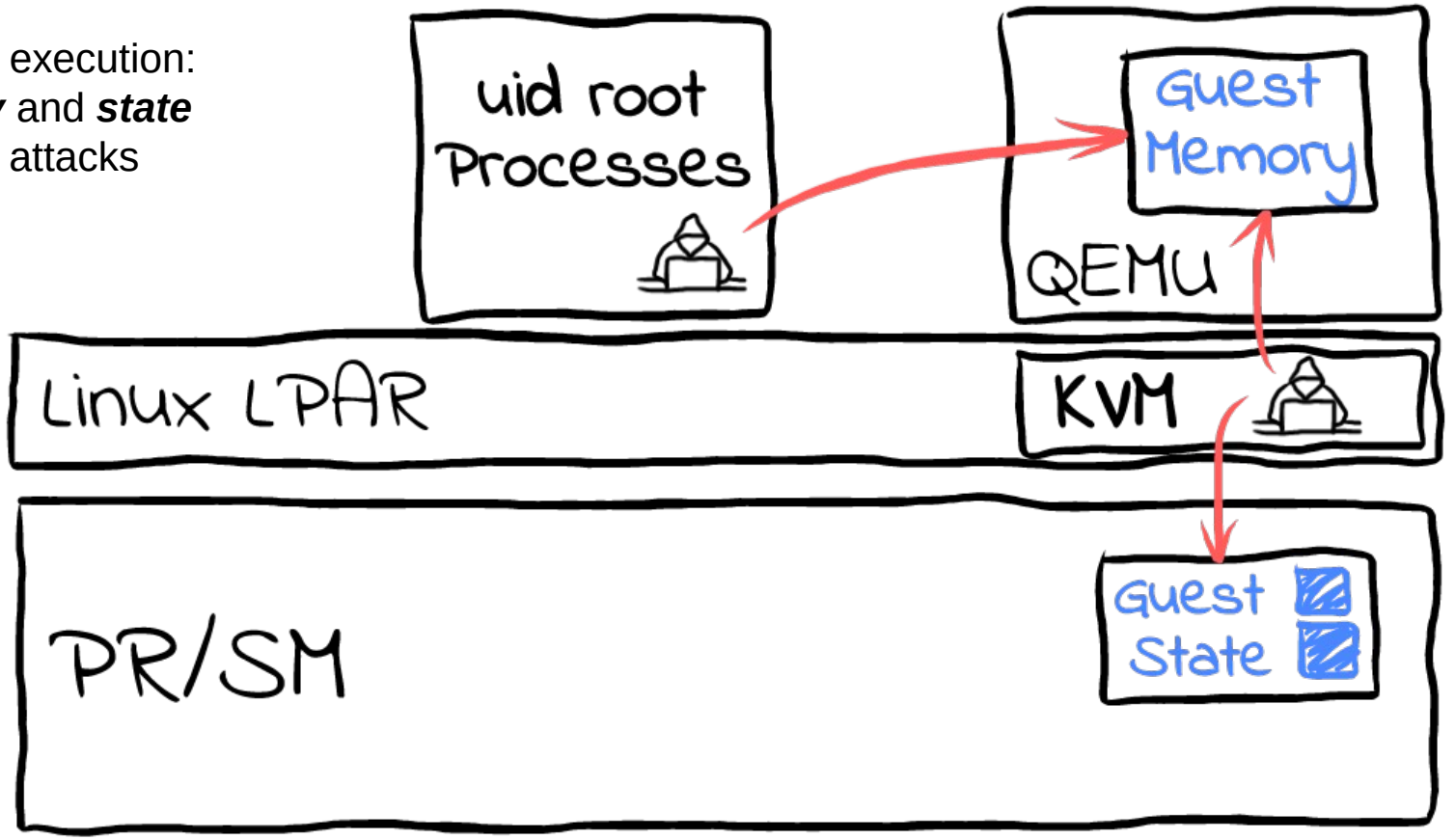
Pre-defined	Static	✔	`<cpu mode='custom'>`
			`  <model fallback='allow'>gen15a</model>`
			`</cpu>`

Host model (***recommended and default on new distros***)	Maximum (based on current host)	✔	`<cpu mode='host-model'/>`
Host passthrough	Maximum	✖	`<cpu mode='host-passthrough'/>`

- z16 CPU model readily available in RHEL 8.5, RHEL 9, SLES 15 SP4, and Ubuntu 21.10, and later.

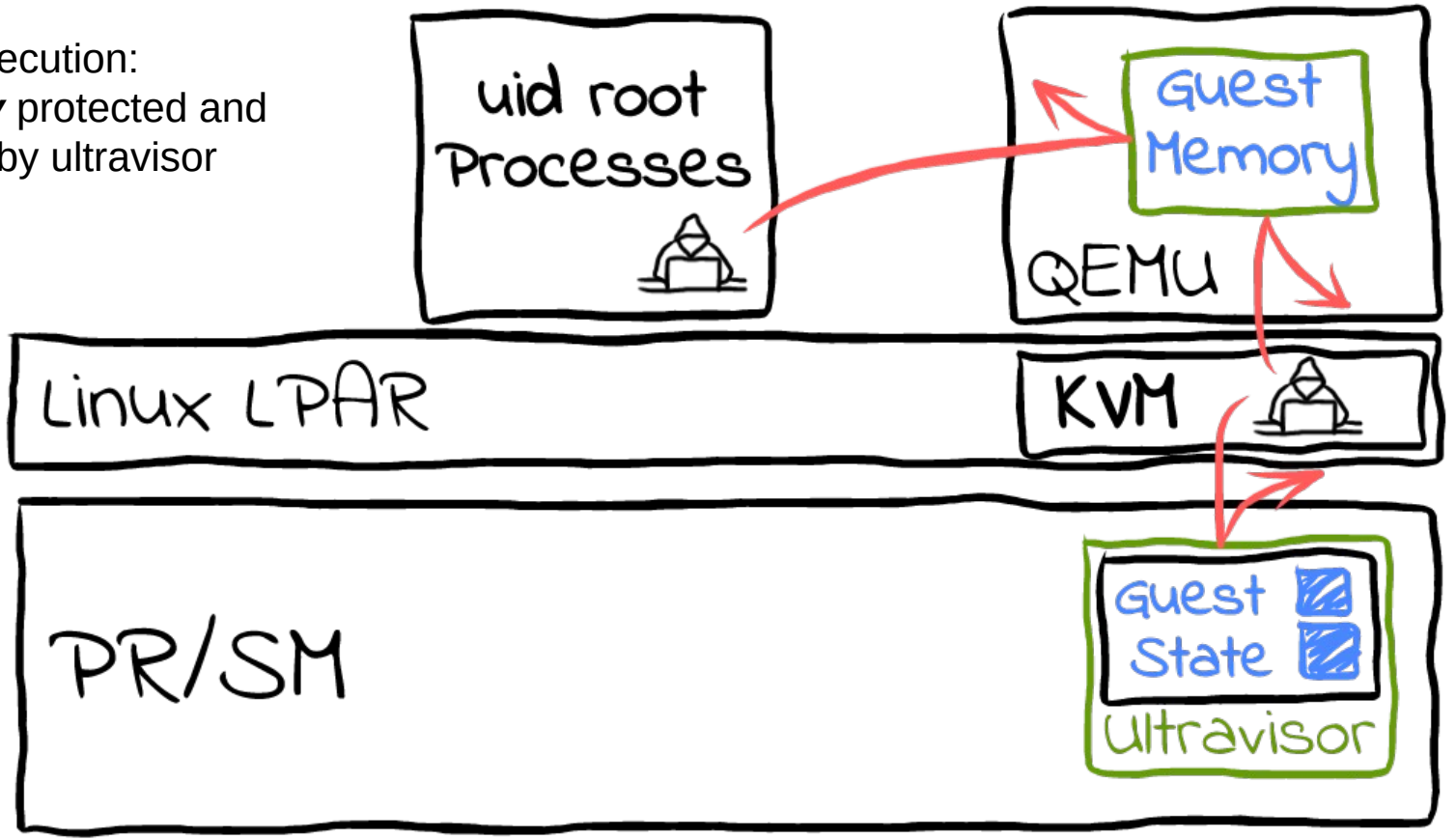
Secure Execution

Without secure execution:
Guest ***memory*** and ***state***
at risk of inside attacks



Secure Execution (*continued*)

With secure execution:
Guest ***memory*** protected and
state shielded by ultravisor

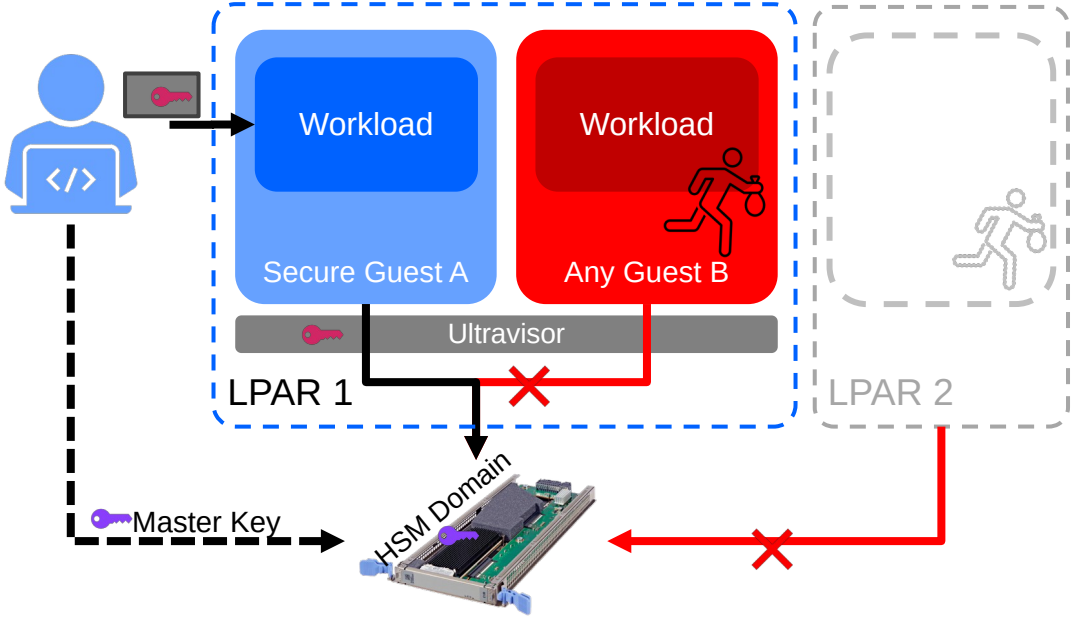


Secure Execution (*cont.*)

- **What it does**: Allows users to run their Linux workloads **with maximum** privacy by protecting system memory.
- **Why you should care**: ***Not even system administrators*** can access customer data
  ⇒ Protection against insider attacks
- Allows users to **run sensitive workloads** on and off premise with the same level of data protection
- Reduces the efforts of a cloud service provider to establish and document procedures for **compliance and certification**

- What is IBM Secure Execution for Linux?
  - Orderable feature of IBM z15 or LinuxONE III (feature code 115)
  - End-to-end memory protection realized in hardware
  - Trusted firmware controlling the separation and isolation of virtual machines
  - CA-certified public private keys to form a chain of trust
- What else is needed?
  - By the **machine owner**: a Linux operating system with KVM supporting IBM Secure Execution (RHEL 8.3, SLES 15 SP2, Ubuntu 20.04)
  - By the **workload owner**: a Linux operating system which supports running as KVM guest in an IBM Secure Execution virtual machine (RHEL 7.8, RHEL 8.2, SLES 12 SP5, SLES 15 SP2, Ubuntu 20.04)

Crypto Express support for Secure Execution



With an IBM z16 / LinuxONE 4 firmware upgrade (MCL P30725.009 in bundle S30 and later), secure guests are able to use Crypto Express domains:

- Adapter domains must be configured in passthrough mode (“dedicated”)
- For CEX8S adapters
  - EP11 or accelerator mode
    - EP11 adapters need firmware update
- Up to 12 adapter domains (virtual HSMs) per secure guest
- The TKE adapter domain administrator must be trusted
- No two domains of the same adapter might be configured to share secure key objects (For example, be configured with the same HSM master key)

A Confidential AI Power Couple

IBM LinuxONE and IBM Z Linux on Z Secure Execution

Hardware-based security for confidential computing

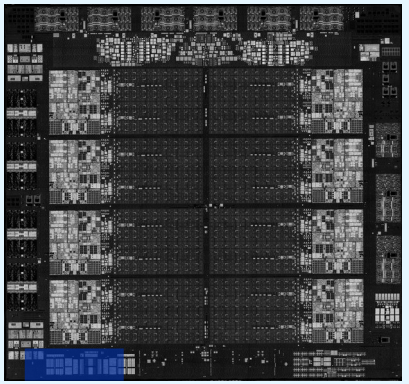
Technically enforced isolation of workloads
at massive scale with Secure Execution for Linux
on IBM Z and LinuxONE

Hyper Protect delivers data integrity and confidentiality taking advantage of IBM SEL to provide a turn-key and intuitive solution stack for workloads

Administrators can still perform their role but do not have data access through technical assurance

+

IBM z16 Integrated Accelerator for AI



- Very low and consistent inference latency and scalable capacity
- **Security – enterprise-grade memory virtualization and protection (data-in-use)**
- Fast direct storage access through new cache design
- Variety of AI models ranging from traditional Machine Learning (ML) to Deep Learning (DL)
- Extensibility with future firmware and hardware updates

Miscellaneous

Need something else for Linux & Virtualization?

Linux and KVM

(A) Use the Request for Enhancements (RFE) database:
- enter in your IBM ID
- select Brand "*Servers and System Software*"
- select Product "*Linux on System z*" (includes KVM)

(B) Reach to us at conferences:
SHARE, GSE, Tech. Univ., VM Workshop

How the Linux Distro Partners handle requirements

Red Hat defined RFE process for customers

SUSE requirements can be submitted to their sales reps as well as using the "feedback" button at the bottom of the SUSE Linux Enterprise Server for IBM Z and LinuxONE web site

Canonical is handling requirements for Ubuntu through Launchpad: Open a bug, put requirement in title and tag with s390x

More information about Linux & KVM

- Official web site https://www.ibm.com/it-infrastructure/z/os/linux
- Linux & KVM (see Backup) Key Documentation Links
- Secure Execution & Compression (see Backup) Videos & books
- Enterprise Key Management for Linux (see Backup)
  Videos & books

User forums

- Mailing lists at Maris college
- Linux on s390x forum at Open Mainframe Project

Need something else for Linux & Virtualization?

Linux and KVM

(A) Use the Request for Enhancements (RFE) database:
- enter in your IBM ID
- select Brand "*Servers and System Software*"
- select Product "*Linux on System z*" (includes KVM)

(B) Reach to us at conferences:
SHARE, GSE, Tech. Univ., VM Workshop

How the Linux Distro Partners handle requirements

Red Hat defined RFE process for customers

SUSE requirements can be submitted to their sales reps as well as using the "feedback" button at the bottom of the SUSE Linux Enterprise Server for IBM Z and LinuxONE web site

Canonical is handling requirements for Ubuntu through Launchpad: Open a bug, put requirement in title and tag with s390x

More information about Linux & KVM

- Official web site https://www.ibm.com/it-infrastructure/z/os/linux
- Linux & KVM (see Backup) Key Documentation Links
- Secure Execution & Compression (see Backup) Videos & books
- Enterprise Key Management for Linux (see Backup)
- Videos & books

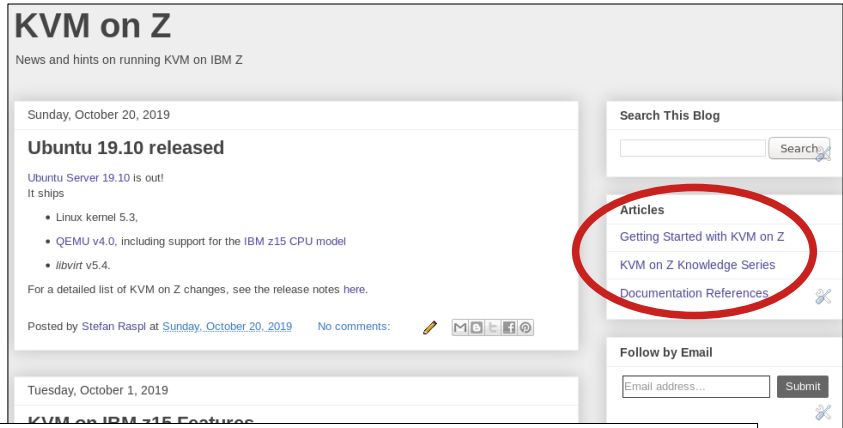
User forums

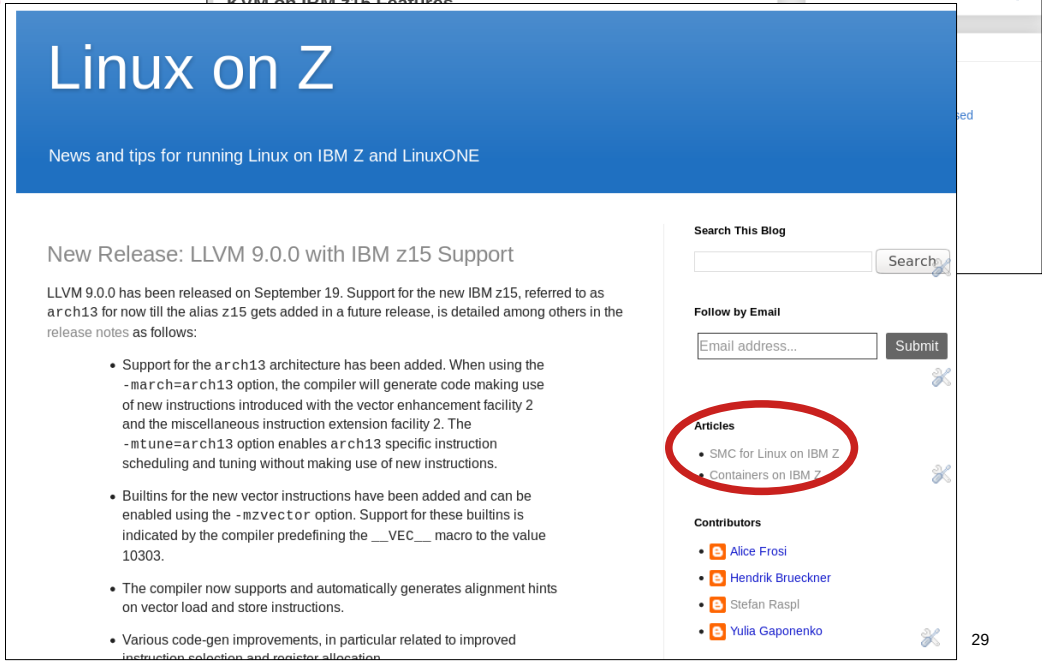
- Mailing lists at Maris college
- Linux on s390x forum at Open Mainframe Project

Staying Up-To-Date

Blogs

- Very latest news from the development team
  - KVM on Z: http://kvmonz.blogspot.com/
  - Linux on Z & containers: http://linux-on-z.blogspot.com/
- Focus primarily on upstream submissions, which will end up in Linux distributions later
- Also features in-depth articles on specific topics
- Provided by Linux & KVM on Z development teams





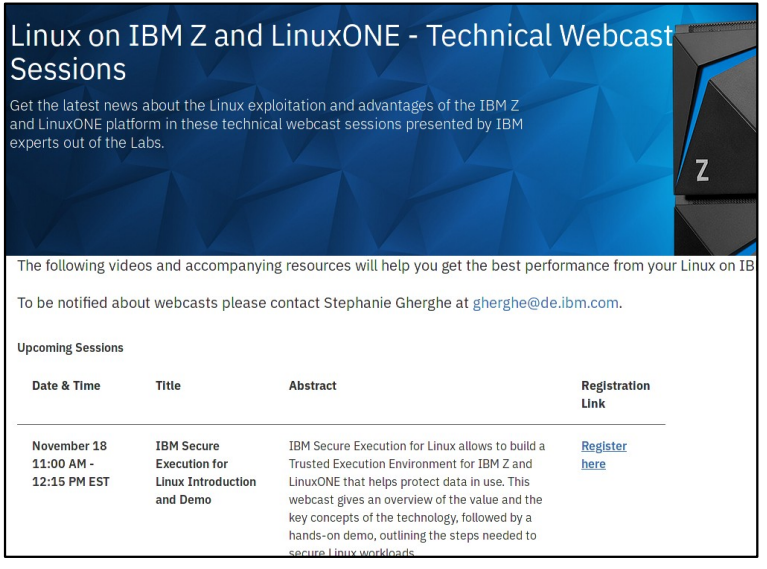
References

Documentation

- Linux on Z and LinuxONE on IBM Documentation
  https://www.ibm.com/docs/en/linux-on-systems?topic=linux-z-linuxone
- Videos explainers
  https://www.ibm.com/docs/en/linux-on-systems?topic=linuxone-video-explainers
- Solution assurance
  https://www.ibm.com/docs/en/linux-on-systems?topic=linuxone-solution-assurance
- z/VM Education Roadmap
  https://www.vm.ibm.com/education/

Webcasts

- In-depth sessions right from the Linux on Z development team
- Recordings available
  https://ibm.biz/Linux-on-IBMzSystems-LinuxONE-Webcasts



Please submit your session feedback!

- All done via the Whova App
- QR Code to the right to download the Whova App
- This session is NA

Tag Legend

- Supported distributions
  - x.y for SUSE SLES <X> Service Pack <Y>, e.g. 15 SP6 for SLES15 SP6
  - x.y for RHEL <x> Update <y>, e.g. 9.4 for RHEL9.4
  - x.y for Ubuntu x.y, e.g. 24.04 for Ubuntu 16.04 LTS
- Suppored environments
  - LPAR usable for systems running in LPAR mode
  - z/VM usable for guests running on z/VM
  - KVM usable for guests running on KVM