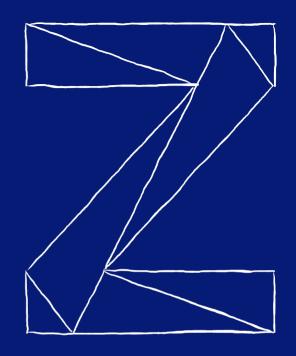
Linux on IBM Z and LinuxONE: What's New

_

Stefan Raspl
IBM Research & Development Germany



VM Workshop, June 2021



Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

AIX* BladeCenter*	DB2* DFSMS	HiperSockets* HyperSwap	MQSeries* NetView*	PowerHA* PR/SM	RMF Smarter Planet*	System z* System z10*	zEnterprise* z10	z/VM* z/VSE*
CICS*	EASY Tier	IMS	OMEGAMON*	PureSystems	Storwize*	Tívoli*	z10 EC	
Cognos*	FICON*	InfiniBand*	Parallel Sysplex*	Rational*	System Storage*	WebSphere*	z/OS*	
DataPower*	GDPS*	Lotus*	POWER7*	RACF*	System x*	XIV*		

^{*} Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Intel, Intel logo, Intel Inside, Intel Inside, Intel Corporation or its subsidiaries in the United States and other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the OpenStack website.

TEALEAF is a registered trademark of Tealeaf, an IBM Company.

Windows Server and the Windows logo are trademarks of the Microsoft group of countries.

Worklight is a trademark or registered trademark of Worklight, an IBM Company.

UNIX is a registered trademark of The Open Group in the United States and other countries.

* Other product and service names might be trademarks of IBM or other companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

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

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This information provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g, zIIPs, zAAPs, and IFLs) ("SEs"). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT"). No other workload processing is authorized for execution on an SE. IBM offers SE at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Agenda

- Linux on Z Distributions
- IBM z15 and LinuxONE III
- Latest Linux on Z Features & Packages



Linux on Z Distributions

Linux Distributions & Hardware Certification

	z15	z14 (all models)	z13	z13s	zEnterprise –	zEnterprise – z196, z114	
	LinuxONE III	Emperor II Rockhopper II	Emperor	Rockhopper	zEC12, zBC12		
RHEL 8							
RHEL 7							
RHEL 6							
RHEL 5					•		
SLES 15	•	•		•	•		
SLES 12		•					
SLES 11				•	•		
Ubuntu 20.04	•	•	•	•			
Ubuntu 18.04	•	•			•		
Ubuntu 16.04		•					

Last update 22/03/2021

See www.ibm.com/systems/z/os/linux/resources/testedplatforms.html for latest updates and details, including *certified Linux distributions by machine*.

Linux on IBM Z Distributions: SUSE

SUSE Linux Enterprise Server 15

- 07/2018 SLES 15 GA: Kernel 4.12, GCC 7.1 / 7.3
- 07/2020 SLES 15 SP2: Kernel 5.3. GCC 7.5 / 9.3
 - EOS 31 July 2028; LTSS: 31 July 2031

But don't forget:
Keep it current,
apply maintweb
updates!!!

SUSE Linux Enterprise Server 12

- 10/2014 SLES 12 GA: Kernel 3.12, GCC 4.8
- 12/2019 SLES 12 SP5: Kernel 4.12, GCC 4.8
 - EOS 31 Oct. 2024; LTSS: 31 Oct 2027

SUSE Linux Enterprise Server 11

- 03/2009 SLES 11 GA: Kernel 2.6.27, GCC 4.3.3
- 07/2015 SLES 11 SP4: Kernel 3.0, GCC 4.3.4
 - EOS 31 Mar. 2019; LTSS: 31 March 2022

Lifecycle Details

https://www.suse.com/en-en/lifecycle/

Linux on IBM Z Distributions: Red Hat

Red Hat Enterprise Linux 8

- 05/2019 RHEL 8 GA: Kernel 4.18, GCC 8.2.1
- 05/2021 RHEL 8.4
- EOS: May 2029; ELS: tbd

But don't forget:
Keep it current,
apply z-Stream
updates!!!

Red Hat Enterprise Linux 7

- 06/2014 RHEL 7 GA: Kernel 3.10, GCC 4.8
- 09/2020 RHEL 7.9
- EOS 30 June 2024; ELS: tbd

Red Hat Enterprise Linux 6

- 11/2010 RHEL 6 GA: Kernel 2.6.32, GCC 4.4.0
- 06/2018 RHEL 6.10
- EOS 30 Nov. 2020; ELS: 30 June 2024

Lifecycle Details

https://access.redhat.com/support/policy/update s/errata

Linux on IBM Z Distributions: Canonical

Ubuntu 20.04 (Focal Fossa)

- 04/2020 GA: Kernel 5.4, GCC 9.3.0, LTS-Release
- 08/2020 Ubuntu 20.04.1:
- EOS: April 2025; ESM: Apr 2030

Ubuntu 21.04 (Hirsute Hippo)

- 04/2021 GA: Kernel 5.11, GCC 10.3
- EOS: January 2022

But don't forget:
Keep it current,
apply SRU
updates!!!

Ubuntu 18.04 (Bionic Beaver)

- 04/2018 GA: Kernel 4.15, GCC 7.2.0, LTS-Release
- 08/2019 Ubuntu 18.04.3 Kernel 4.15/4.18 GCC 7.2.0
- EOS: April 2023; ESM: April 2028

Ubuntu 16.04 (Xenial Xerus)

- 04/2016 GA: Kernel 4.4, GCC 5.3.0+, LTS-Release
- 02/2019 Ubuntu 16.04.06 LTS
- EOS: April 2021; ESM: April 2024

Lifecycle

- Regular releases every 6 months, supported for 9 months
- LTS releases every 2 years, supported for 5 years
- LTS enablement stack provides newer kernels
- http://www.ubuntu.com/info/release-end-of-life

IBM z15 & LinuxONE III

IBM z15 Machine Type: 8561 Model T01



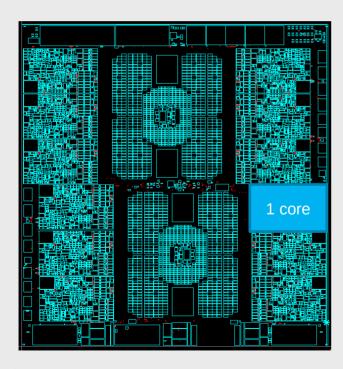
IBM z15
IBM LinuxONE III

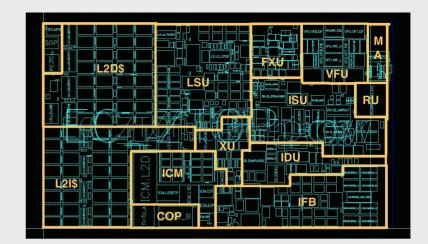
IBM LinuxONE III Machine Type: 8561 Model LT1



IBM z15 Overview

- 19" industry standard form factor
- New on-chip compression acceleration





- 14nm SOI technology
- 17 layers of metal
- 5.2 GHz
- 12 cores per CP-chip
- Increased cache sizes:
 - 2x L2 on-chip
 - 1.4x L4
- 14% single thread performance improvement
- 40TB max memory
- 190 usable cores

as IBM z14

NEW

IBM z15 Machine Type: 8562 Model T02



IBM z15 Single Frame Offerings

Processor

- Up to 6 cores for CP and up to 65 IFLs
- 14% Single Thread Performance Improvement
- 14% maximum system capacity growth over z14 ZR1
- 19" frame with 8U or 16U Reserved Space
- New on-chip functions Integrated Accelerator for zEDC – for more efficient storage of data

Memory

Up to 2X more RAIM Memory - 16 TB Max Per System

To the Data

 Faster SSL/TLS handshake performance on T02 with Crypto Express7S compared to z14 ZR1 with Crypto Express6S

IBM LinuxONE III Machine Type: 8562 Model LT2



IBM z15 Toleration Support:

Linux Distros and z/VM (Minimum Versions)

Linux Distributions

- Red Hat Enterprise Linux 8.0 (z stream if needed)
- Red Hat Enterprise Linux 7.7.z
- Red Hat Enterprise Linux 6.10.z
- SUSE Linux Enterprise Server 15 SP1 (maintweb if needed)
- SUSE Linux Enterprise Server 12 SP4 maintweb
- Canonical Ubuntu 18.04 LTS
- Canonical Ubuntu16.04 LTS

z/VM Hypervisor

- z/VM7.1

Install z15 with the Linux environment you use today!

Latest Linux on Z Features & Packages

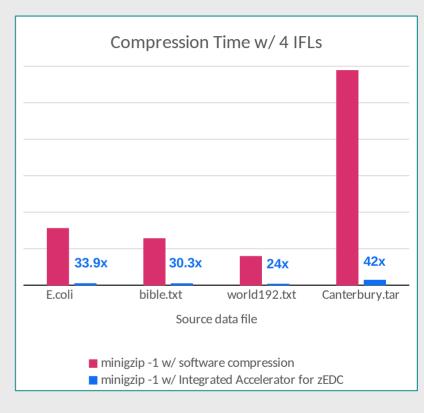
IBM z15 Support: New Vector Instructions

- Reported with new feature flags in /proc/cpuinfo
 - vxp
 - vxe2
- Examples for use of new vector instructions:
 - Vector alignment hints
 - Vector Byte and element swaps
 - Vector substring search in strstr() and memmem()
- Exploited (among others) in
 - GCC 9.1
 - glibc 2.30
 - LLVM 9.0.0

IBM z15 Support: Deflate

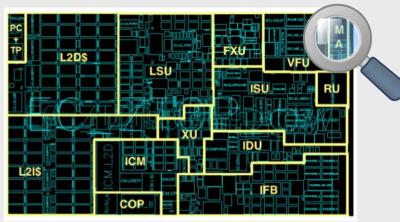


- Data compress and uncompress through new instruction
- Compression equivalent to gzip -1
 - −1 is fastest, −9 slowest, default is −6
- Can be exploited e.g. by zlib, gzip, Java et al
- Compress data with zlib on IBM z15 with 4 processors up to 42x faster as compared to software compression
- Enabled for the default compression level starting with RHEL 8.4, SLES 15 SP3, and Ubuntu 20.10
- Linux enablement:
 - Java: Use Java 8 SR6 FP16 on any Linux distribution
 - Reported with new feature flag in /proc/cpuinfo: dflt
 - Use env variable DFLTCC_LEVEL_MASK to enable for arbitrary compression levels
 - See here for further details on usage



IBM z15 Support: CPACF

- New Message Security Assist MSA9 for *Elliptic Curve Cryptography* (ECC)
- Supports
 - message authentication
 - generation of elliptic curve keys
 - scalar multiplication
- Used with SSL/TLS protocol
 - securing client-server network connection
 - handshake establishes the secure connection
- TLS v1.2 and v1.3 support ECDH (key exchange) and ECDSA (signature)



z15 Processor Unit

Supported curves:

- ECDSA (sign/verify) P256, P384, P521 Ed 25519, Ed448
- ECDH (key exchange) P256, P384, P521,
 X25519, X448

Performance

- Up to **20x key exchange** operations
- Up to **38x sign** operations
- Up to **10x verify** operations

IBM z15 Support: Secure Boot for SCSI IPL



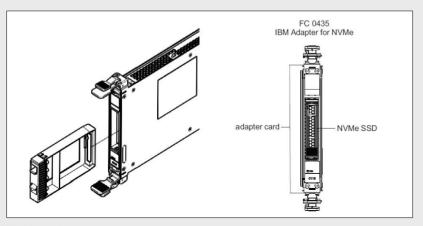
- Ensure that only code is loaded during IPL that is
 - signed by a trusted distribution vendor (currently: Red Hat, SUSE or Canonical)
 - unmodified
- Kernel image and zipl boot record must be signed
- zipl tool creates signature entries for SCSI IPL
- New switch on HMC enables secure boot
- Firmware checks signatures and stops IPL on mismatch

- /sys/firmware/ipl/has_secure indicates support
- /sys/firmware/ipl/secure indicates IPL
 using secure boot
- zipl option secure="auto/0/1"
 - 0 disable secure boot
 - 1 enforce secure boot
 - auto enable secure boot if system supports it and image/stage3 signed
- Support available in
 - Linux kernel 5.3

IBM LinuxONE support for NVMe drives

IBM Adapter for NVMe

- Carrier card for industry standard U.2 NVMe drives
 - Common capacities up to 16 TB per drive
 - 1 drive per carrier, up to 16 cards per CEC
- Available for IBM LinuxONE starting with Emperor II and Rockhopper II



NVMe drive characteristics

- Low-cost, low-latency, high-throughput storage
- PCI direct-attached (no SAN)
 - No cabling, switches, etc. required
- No virtualization or shared access: can use one drive only in one LPAR/VM
- Linux on Z support for NVMe 18.04

18.04 7.8 8.0 SP.

- Uses standard Linux NVMe driver
- Always apply latest service levels!

Working with NVMe drives in Linux

Listing available devices

- Use lspci to show PCI device information
- Use lsblk to list NVMe block devices:

```
$ lsblk
                         SIZE RO TYPE MOUNTPOINT
NAME
             MAJ:MTN RM
nvme0n1
             259:0
                              0 disk
⊢nvme0n1p1 259:1
                        300M
                              0 part
-nvme0n1p2 259:2
                              0 part
-nvme0n1p3 259:3
                              0 part
└nvme0n1p4 259:4
                              0 part
```

Storing data

- Use NVMe for swap, root, boot* and data file systems (*=requires IPL support)
- Use of software RAID recommended

Management tools

- nvme-cli: query and manage NVMe device functions
- zpcictl: recovery and service actions

IPL support



- Available on LinuxONE III with latest firmware
 - LPAR only
- NVMe IPL support added with
 - Linux kernel v5.8
 - s390-tools v2.14
- Distribution installer support available

zdsfs



- What it does: Provides read-only access to z/OS data sets on DASD
- Why you should care:
 - Transfer speed closer to FICON limits
 - No CPU cycles spent in z/OS during data transfer
- Latest Improvements
 - Robustness
 - Use storage mechanisms to prevent concurrent access
 - Get an updated dataset list without need to unmount/mount DASDs
 - Coordinated read access
 - Establish authorization checking and audit capabilities
 - DASD no longer required offline in z/OS during Linux access
 - Transparent dataset conversion (upcoming)
 - Read EBCDIC encoded datasets as ASCII without any add'l codepage conversion required

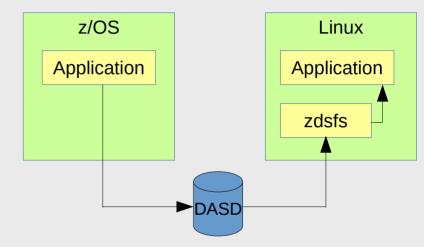


Fig.1: Architecture overview

\$ zdsfs /dev/disk/by-path/ccw-0.0.edc0 /mnt/
Using z/OSMF REST services on https://192.168.1.1/zosmf/
\$ cat /mnt/HABERL.ZDSFSDV.LARGE.TEST2
[...]
\$ fusermount -u /mnt

Fig.1: Usage example

SMC-Dv2

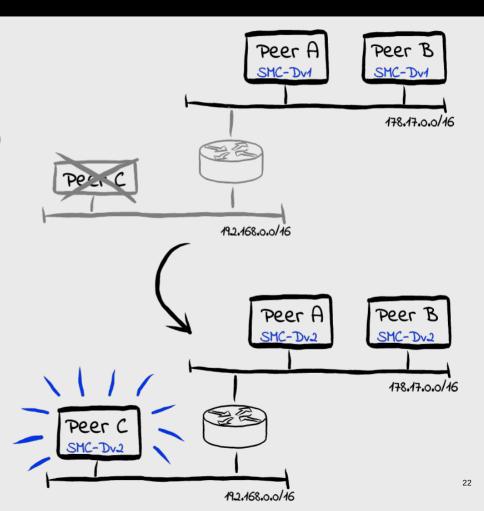


Recap:

- SMC-Dv1 provides intra-CEC acceleration for TCP traffic using Internal Shared Memory (ISM) devices
- Superior performance (low latency, high throughput)
 at reduced CPU consumption
- However:
 - 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
- Support available in Linux kernel 5.10

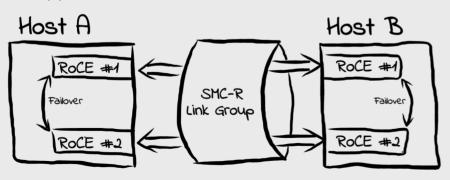


SMC-R / smc-tools

SMC-R Link Group Support



- Transparently moves connections between links in link group upon link failure - think channel bonding for SMC-R
- Compatible to z/OS
- Reference Architecture:
 - 2x OSA for IP connectivity
 - 2x RoCE for RDMA
- Support available in Linux kernel 5.8



smc-tools v1.5

- Utilities in support of SMC-R and SMC-D
- Latest additions:
 - New tools smcd/smcr, e.g

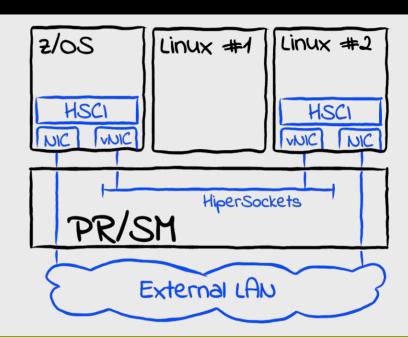
```
smcd info
Kernel Capabilities
SMC Version:
                    2.0
SMC Hostname:
                    tux
                   v1 v2
SMC-D Features:
SMC-R Features:
                    \nabla 1
Hardware Capabilities
SEID:
                    IBM-SYSZ-ISMSEID0000...
ISM:
                    v1 v2
RoCE:
                    \nabla 1
```

- New tool smc_chk to verify setup/peer capabilities

HiperSockets Converged Interface



- Forms a single logical LAN segment across
 - external LAN, and
 - HiperSockets
- I.e. one IP to rule them all!
- The best of both worlds: Prefer HiperSockets over external LAN, taking advantage of the low latencies of HiperSockets in a simple setup
- Can register a single MAC only per Linux instance for now –
 i.e. not usable with Open vSwitch
- Requirements/compatibility:
 - z15 or LinuxONE III
 - HiperSockets VCHID with External Bridge option set
 - All participants accessing the same HiperSockets channel need to setup HSCI, too
 - In z/VM, use VSWITCH Bridgeport
 - Compatible with z/OS, providing layer 2 connectivity to z/OS over HiperSockets
- Support available in Linux kernel 5.10 and s390-tools v2.16



\$ hsci add enc8410 encb040

Verifying net dev encb040 and HiperSockets dev enc8410 Adding hsci8410 with a HiperSockets dev enc8410 and an external dev encb040 Set enc8410 MAC fe:c2:f4:35:00:12 on encb040 and hsci8410 Successfully added HSCI interface hsci8410

\$ ip addr add 192.1.2.3/16 dev hsci8410

\$ hsci sh	ow		
HSCI	PNET_ID	HiperSockets	External
hsci8410	NET1	enc8410	encb040

Other Packages

- **s390-tools** v2.16 (02/2021)
 - Userspace tools for use with the Linux kernel and its device drivers on IBM Z
 - Homepage: https://github.com/ibm-s390-linux/s390-tools
 - v2.16 supports Linux kernel 5.11
 - New tool: hsci, manages HiperSockets Converged Interfaces
 - See CHANGELOG for further details
- *Open Cryptoki* v3.15 (10/2020)
 - Implements Cryptoki API as defined by the PKCS#11 specification
 - Homepage: https://github.com/opencryptoki/opencryptoki/
 - Latest Changes:
 - Conform to PKCS#11 3.0 Baseline Provider profile
 - p11sak tool: add remove-key command version
 - · CCA: Support key wrapping
 - SOFT: Support ECC

- *qclib* v2.3 (04/2021)
 - C library providing information on system, capacity, and virtualization layers
 - Latest changes (as of 2.2.1):
 - Support for zCX environment

• Added new commands zname and zhypinfo

	Madea new commands zhame and zhiyethio								
\$	\$ zname								
IB	IBM z15								
\$	<pre>\$ zhypinfo</pre>								
#	Layer_Type	Categ	Name	IFLS	CPs	Tot			
3	z/VM-guest	GUEST	R3545038	3	0	3			
2	z/VM-hypervisor	H0ST	B0ER3545	2	1	3			
1	LPAR	GUEST	R35LP45	2	1	3			
0	CEC	H0ST	P35	60	3	63			

- *smc-tools* v1.5 (01/2021)
 - Utilities in support of SMC-R and SMC-D
 - Latest changes:
 - New tools smcd/smcr for link group and device information
 - New tool smc_chk to verify setup/peer capabilities 25

Compilers

- Development focus on GCC and LLVM
- IBM z15 support available in GCC 9.1 and LLVM 9.0.0 or later as follows:
 - -march=arch13 enables GCC and LLVM z15 instruction set exploitation
 - New bit operations
 - 2 way conditional register moves (select)
 - Vector byte and element swaps
 - -mtune=arch13 for z15 specific instruction scheduling
 - No new instructions used (does not require a z15 to run)
 - Alias -march/-mtune=**z15** available in GCC 9.3 and LLVM 10.0.0 or later

Libraries & Debuggers

GNU C Library Support

- Hardware capability flags indicate z15 facilities
- New instruction MVCRL (move right to left), provides 1.5x speedup for memmove when used for array insert operations
- New instruction Vector Substring Search, providing 2.5x speedup for strstr and memmem functions
- Included in glibc 2.30
- RHEL 8.1, SLES 12 SP5, SLES 15 SP2, Ubuntu 19.10, or later

GNU Debugger GDB

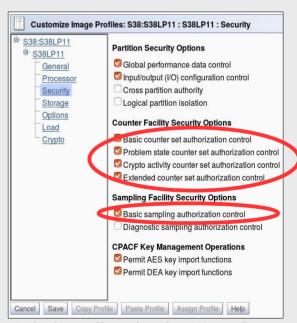
- z15 instruction set support for record and replay (reverse debugging)
 - Upstream in GDB 9.1
 - Available via TCM 2019 for SLES, Ubuntu 20.04

Valgrind - Memory Debugger

- Full instruction set translation to and from intermediate language required
- z13 support available with RHEL 8.0, SLES 12 SP5, SLES 15 SP1, Ubuntu 19.04
- z14 support available as of Valgrind v3.17

Perf – Performance Profiling

- Includes support for IBM Z hardware sampling and counter facilities
- z15 support available with RHEL8.0, RHEL7, SLES15, SLES12 & Ubuntu19



Libraries with IBM Z Optimizations

• Algebra libraries serve as back-ends for various math frameworks like *Numpy*, *R*, and *Octave*.

- OpenBlas

- z13 double precision support (dgemm) upstream since 0.2.20
 ⇒ RHEL 7/8, SLES 12/15, Ubuntu 18.04
- z14 single precision support (sgemm) upstream in 0.3.10
 ⇒ future distros

libAtlas

z13 and z14 support in Ubuntu 19.10

Note: *OpenBlas* and *libAtlas* implement the standard *blas* library interface and can be used interchangeably.

Eigen

- CPU compute back-end for Tensorflow
- Source-only package: Recent versions will be pulled in during build
- IBM z13 full support since 3.3.0
- IBM z14 support upstream since 2017 will be part of upcoming release
- z15 alignment hints added by compiler

Anaconda

- Popular Data-Science platform
- Enables installation of most popular AI tools and frameworks independent from Linux distribution







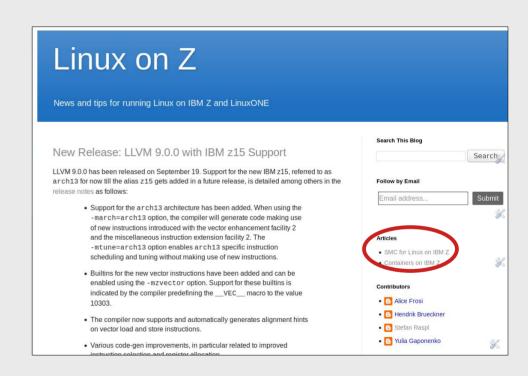


Miscellaneous

Staying Up-To-Date

Blogs

- Very latest news from the development team
 - 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 on Z development team



References

Documentation

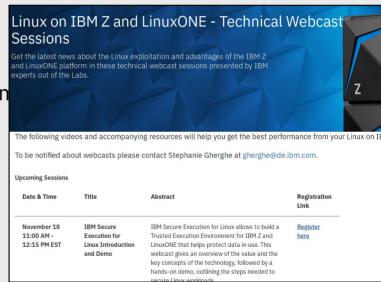
- Linux on Z and LinuxONE on IBM Documentation https://www.ibm.com/docs/en/linux-on-systems?topic=linux-z-linuxone
- Videos explainers
 https://www.ibm.com/docs/en/linux-on-systems?topic=linuxone-video-explainers

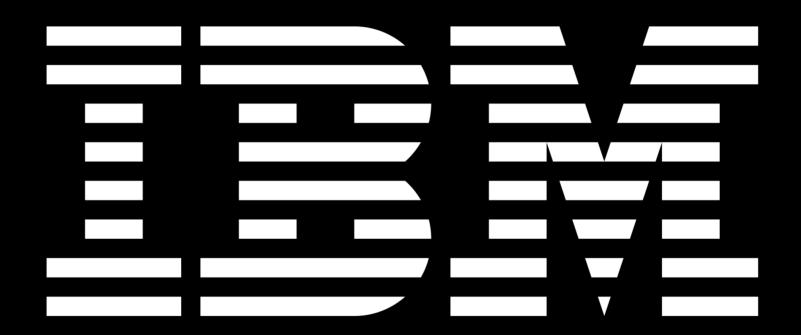
Webcasts

- In-depth sessions right from the Linux on Z development tean
- Recordings available http://ibm.biz/Linux-on-IBMZ-LinuxONE-Webcasts

Blogs

- Primary places for news and updates
 - Linux on Z, including containers: http://linux-on-z.blogspot.com/





Tag Legend

Supported distributions



- for RHEL $\langle x \rangle$ Update $\langle y \rangle$, e.g. 7.4 for RHEL7.4
- for Ubuntu x.y, e.g. for Ubuntu 16.04 LTS
- Suppored environments
 - usable for systems running in LPAR mode
 - (z/VM) usable for guests running on z/VM
 - (KVM) usable for guests running on KVM