

The IBM z15 processor chip set



[Anthony Saporito](#)

Senior Technical Staff Member
IBM Systems Hardware Development
IBM Poughkeepsie, NY
saporit@us.ibm.com

Trademarks

IBM, the IBM logo, and [ibm.com](http://www.ibm.com) are trademarks or registered 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 <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

AIX®, Bluemix®, CICS®, Db2®, DB2®, Distributed Relational Database Architecture™ DS8000®, FICON®, FlashCopy®, GDPS®, Global Technology Services® HyperSwap®, IBM®, IBM Watson® IBM Z®, IBM z Systems®, IBM z13®, IBM z13s®, IBM z14®, IBM z15®, TM Interconnect®, Language Environment® MVSTM OMEGAMON® Parallel Sysplex® Passport Advantage® PowerPC®, RACF®, Redbooks® Redbooks (logo) Resource Link® S/390®, System Storage™, System z® System z10®, System z9®, VIA®, VT AM® Watson™ WebSphere®, z Systems® z/Architecture® z/OS®, z/VM® z/VSE® z13®, z15™®, z9® zEnterprise®

The following terms are trademarks of other companies:

Evolution, are trademarks or registered trademarks of Kenexa, an IBM Company.

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

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

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

Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

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

VMware, and the VMware logo are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

Other company, product, or service names may be trademarks or service marks of others.

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 Sync 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 Sync 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.

The mainframe and enterprise assets are at the center of a digital enterprise



Virtually Limitless Scale

1.3 million CICS transactions are processed every second, every day. In comparison, there are **68,542 Google** searches every second globally¹

220+ billion lines of COBOL
COBOL accounts for more than **70%** of the business transactions that take place in the world today



You've likely used a mainframe today

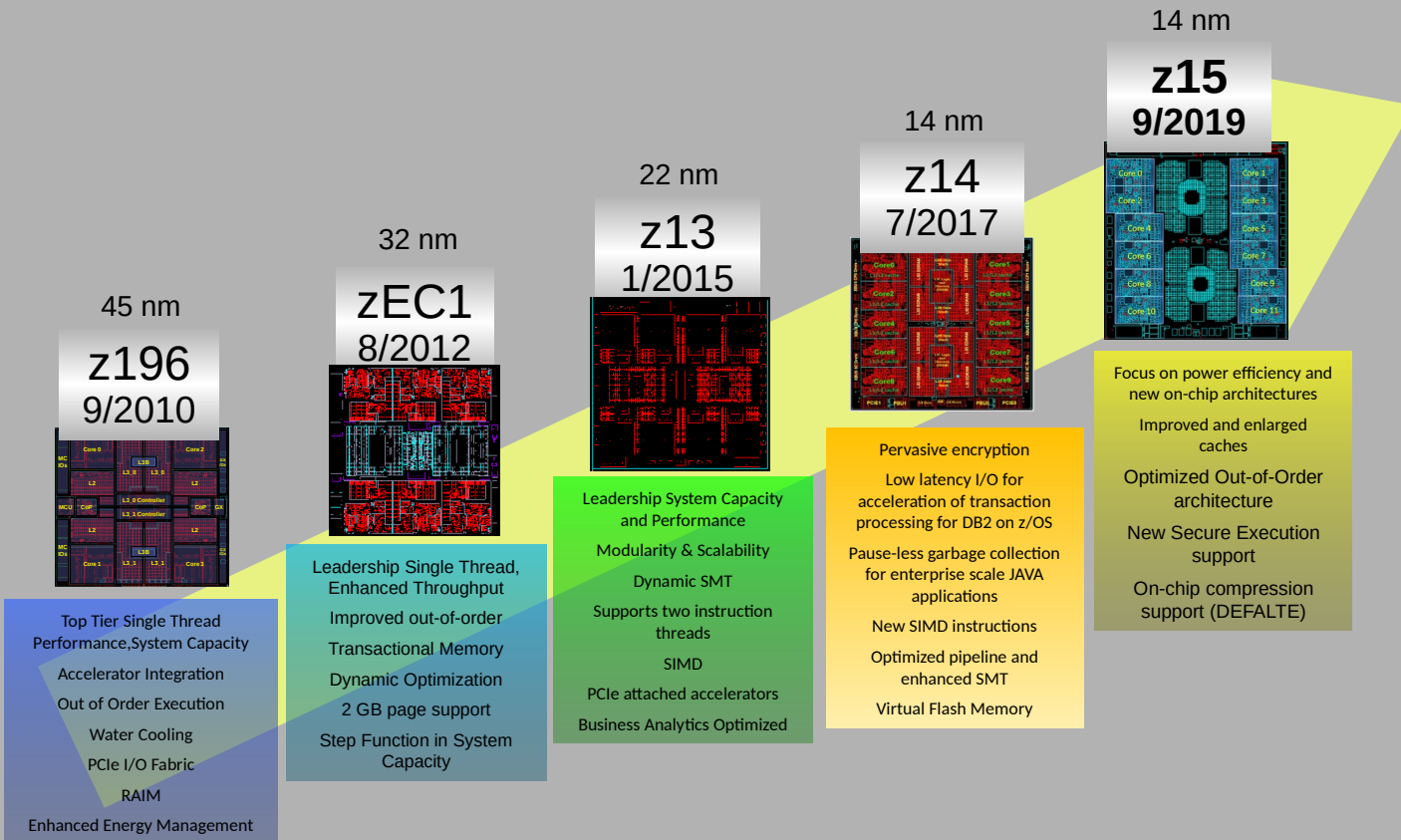
- 400 million retail transactions daily
- 80 million ATM transactions daily
- 1 million hotel night reservations daily
- >90% of all airline reservations.



3

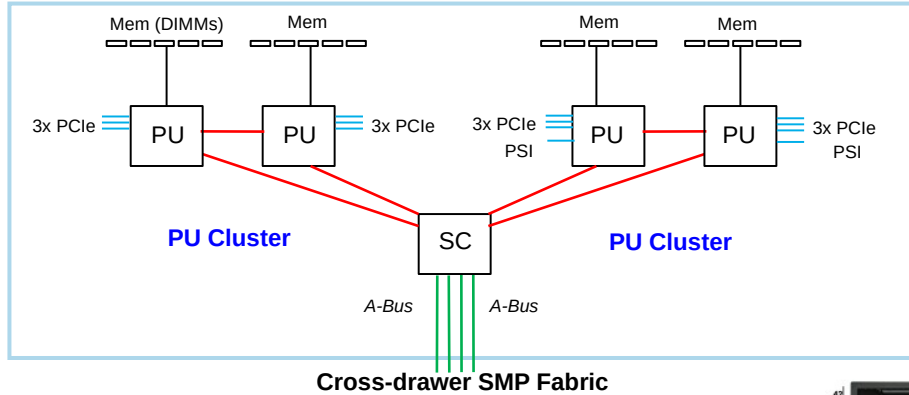
¹ IBM estimates based on real client usage. <http://www.statisticbrain.com/google-searches>

IBM Z – Processor Roadmap

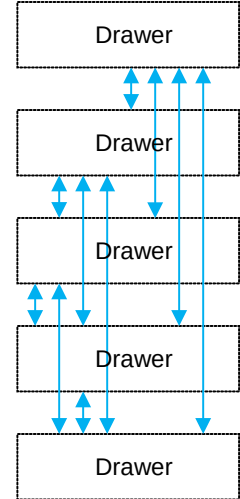


z15 Drawer & System Topology

Fully Populated Drawer



5 Drawer System Fully Interconnected



CP chip, 696 sqmm, 14nm, 17 layers of metal

- 9.2 billion transistors
- 12 cores, each 4+4MB I+D L2 cache
- Shared 256MB L3 cache

SC chip, 696 sqmm, 14nm, 17 layers of metal

- 12.2 billion transistors
- System interconnect & coherency logic
- Shared 960MB L4 cache

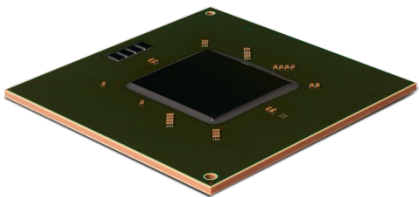
Max System:

- 20 CP sockets in SMP interconnect
- 240 cores (190 customer configurable)
- 40TB RAIM-protected memory
- 60 max PCI gen4x16 fanouts to IO/coupling
- 192 IO cards / 384 channels max

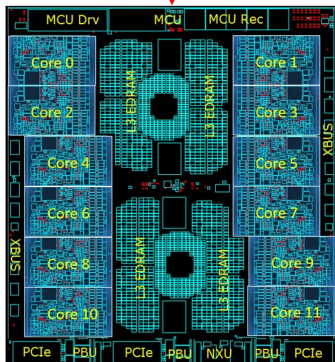


z15 Processor Chipset & Drawer Design

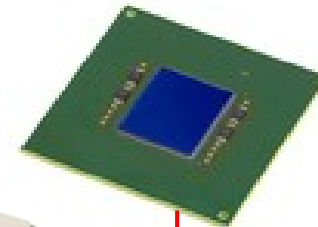
CP Chip



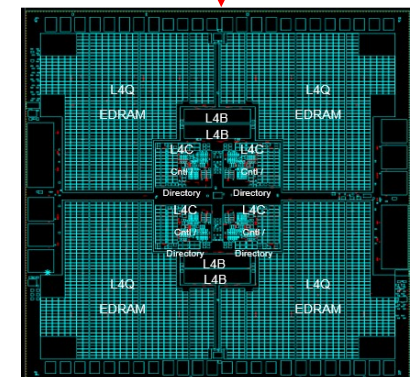
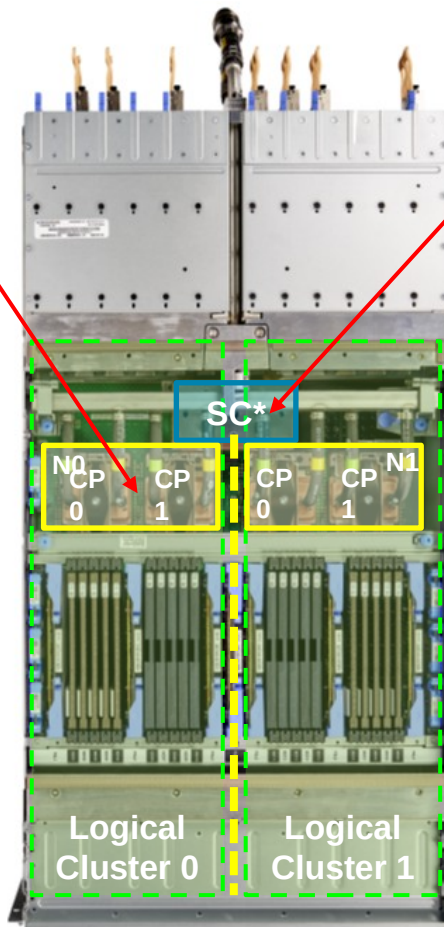
4x CP SCMs



SC Chip



1x SC



z15 Processor Design Summary

Micro-Architecture

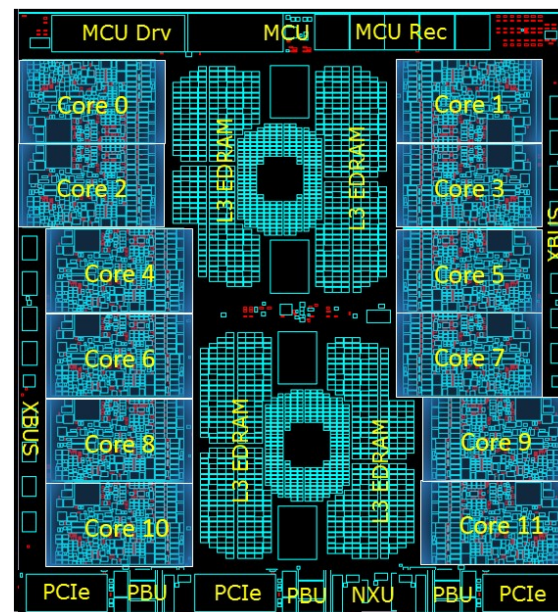
- 12 cores per CP-chip
- 5.2GHz
- More than 9.1 Billion Transistors
- Cache/TLB Improvements:
 - 128KB I\$ + 128KB D\$
 - L2 I/D\$ (4MB)
 - 256MB L3 Cache
 - 12 Concurrent L2\$ Misses
 - Enhanced D\$ hardware prefetcher
 - 512 entry 2-gig TLB2
- Pipeline Optimizations:
 - SHL/LHS avoidance improvements
 - Issue/Execution side swap on long running VecOps
 - Larger Global Completion Table
 - Larger Issue Queues
 - New Mapper design
 - BFU latency/throughput improvements
- Branch Prediction Improvements:
 - 16K enhanced BTB1 design
 - New Targe based PHT predictor
 - Improved call/return predictor

Architecture

- Secure Execution
- 38 new instructions for:
 - GPR based logical operations
 - Accelerators
 - Vector search & shifting
 - Vector load/store reversed
 - Vector 2x bandwidth loads
 - Conversions & more!

Accelerators

- On Chip Deflate (gzip)
- On Core Modulo Arithmetic (ECC)
- On Core sort/merge acceleration



z15 Processor Design Summary

Micro-Architecture

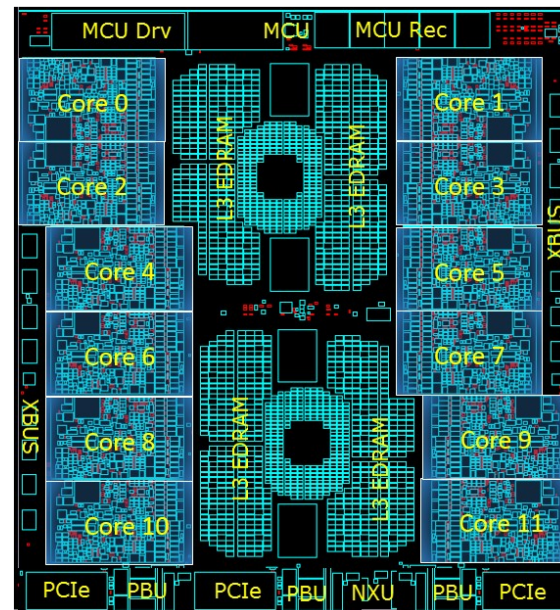
- 12 cores per CP-chip
- 5.2GHz
- More than 9.1 Billion Transistors
- Cache/TLB Improvements:
 - 128KB I\$ + 128KB D\$
 - L2 I/D\$ (4MB) (2x I\$ vs z14)
 - 256MB L3 Cache (2x vs z14)
 - 12 Concurrent L2\$ Misses (2x vs z14)
 - Enhanced D\$ hardware prefetcher
 - 512 entry 2-gig TLB2 (2x vs z14)
- Pipeline Optimizations:
 - SHL/LHS avoidance improvements
 - Issue/Execution side swap on long running VecOps
 - Larger Global Completion Table (25% > z14)
 - Larger Issue Queues (20% > z14)
 - New Mapper design (2x entries @ ½ area of z14)
 - BFU latency/throughput improvements
- Branch Prediction Improvements:
 - 16K enhanced BTB1 design (2x vs z14)
 - New TAGE based PHT predictor
 - Improved call/return predictor

Architecture

- Secure Execution
- 38 new instructions for:
 - GPR based logical operations
 - Accelerators
 - Vector search & shifting
 - Vector load/store reversed
 - Vector 2x bandwidth loads
 - Conversions & more!

Accelerators

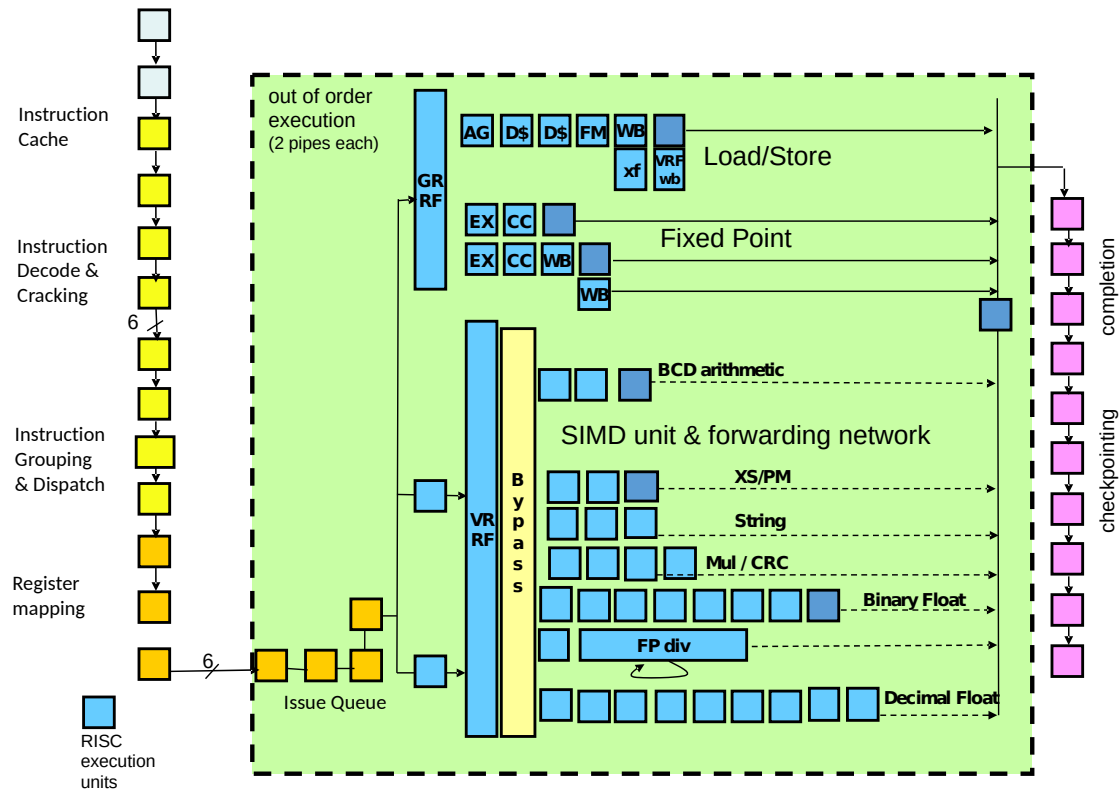
- On Chip Deflate (gzip)
- On Core Modulo Arithmetic (ECC)
- On Core sort/merge acceleration



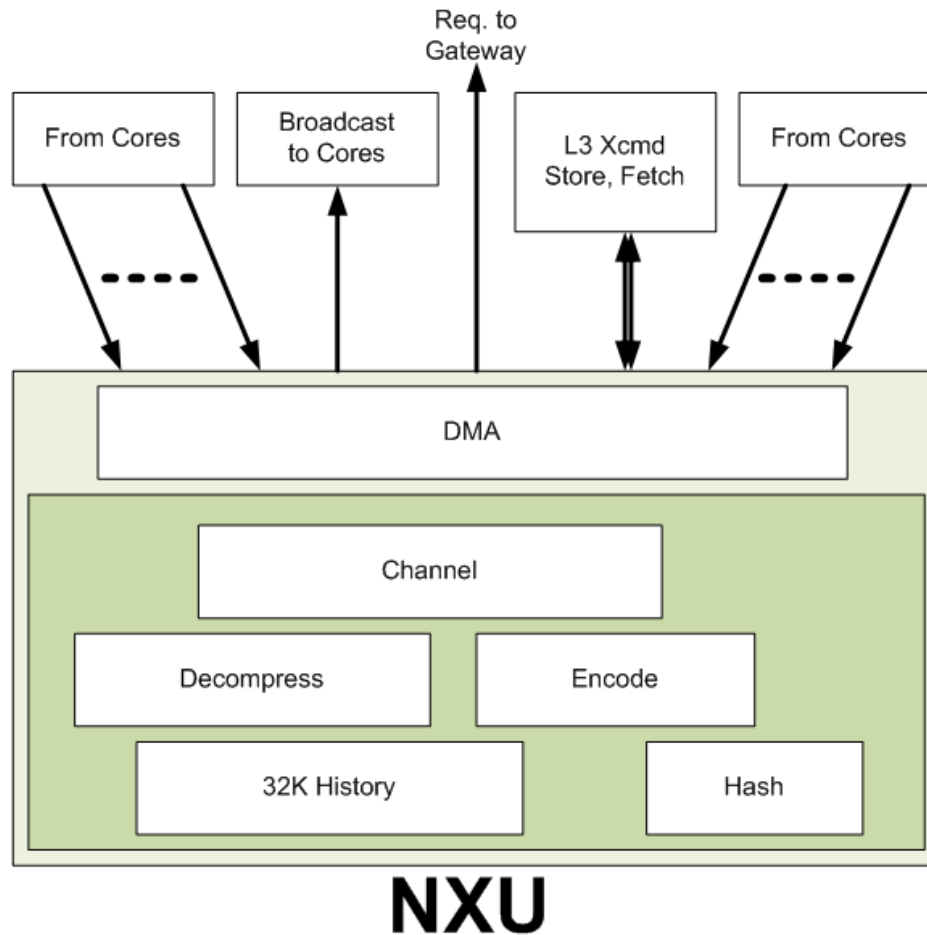
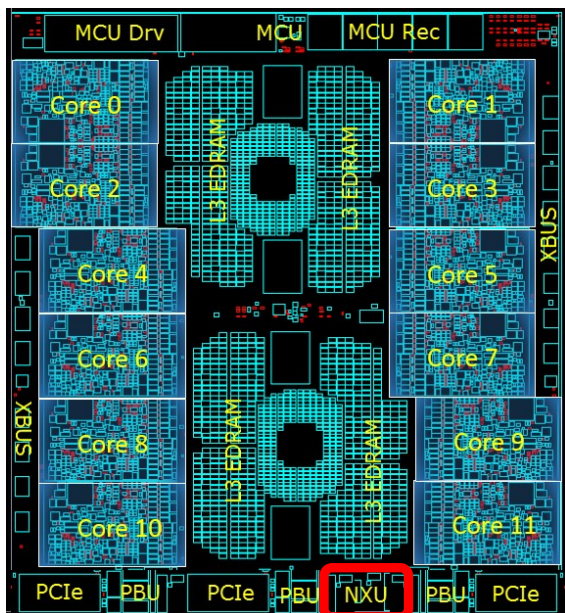
z15 Processor Pipeline

Deep high frequency pipeline

- Async branch prediction running ahead of instruction fetching
- 32B/cycle instruction fetch
- 6 instruction / cycle parse & decode
- CISC instruction cracking
- Unified OOO issue queue
- 2 LSU, 4-cycle load-use
- 4 FXU, 2 SIMD/FP/BCD
- In-order completion & checkpoint

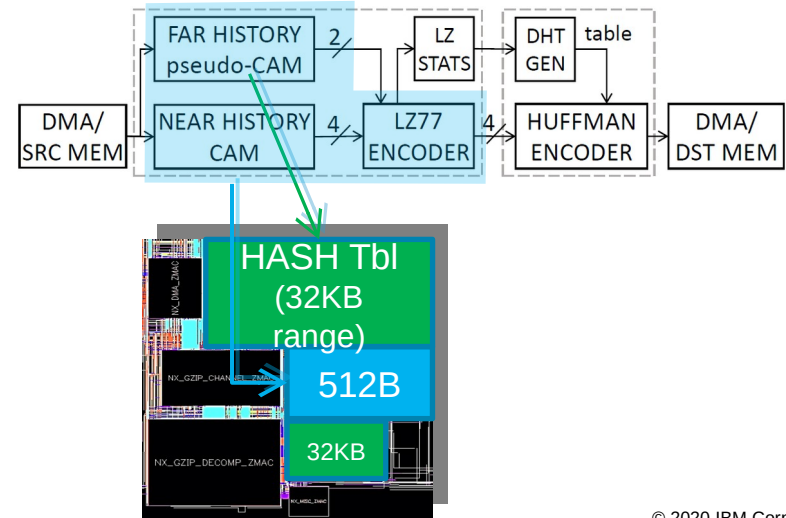
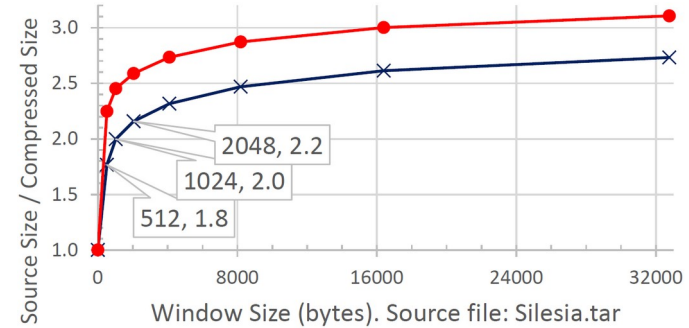


z15 Integrated Deflate Accelerator – Design Overview



z15 Integrated Deflate Accelerator - Hybrid LZ Encoder/Compressor

- Traditionally, two methods have been used for searching
 - Content Addressable Memory (CAM)
 - SRAM based hash table as a dictionary (pseudo-CAM)
- CAMs are precise, but area and power hungry
 - Thousands of comparators running in parallel
 - Custom circuits or random logic/latches (our design)
- Hash tables are imprecise and lossy, but area efficient
 - SRAM based
 - 14 to 16 times as many bits per unit area than CAM
 - SRAM port count, hash and bank collisions determine the precision and area (precision \square compression ratio)
- We recognized that compression ratio levels off with increasing window size
 - Locality of duplicate phrases
 - Far pointers use more bits than near pointers in Deflate
- Therefore, we use CAM for the 512B “Near” history and area-efficient hash table for the “Far” history, 513B to 32KB
 - Area/timing budgets didn’t permit using a 1KB or 2KB CAM



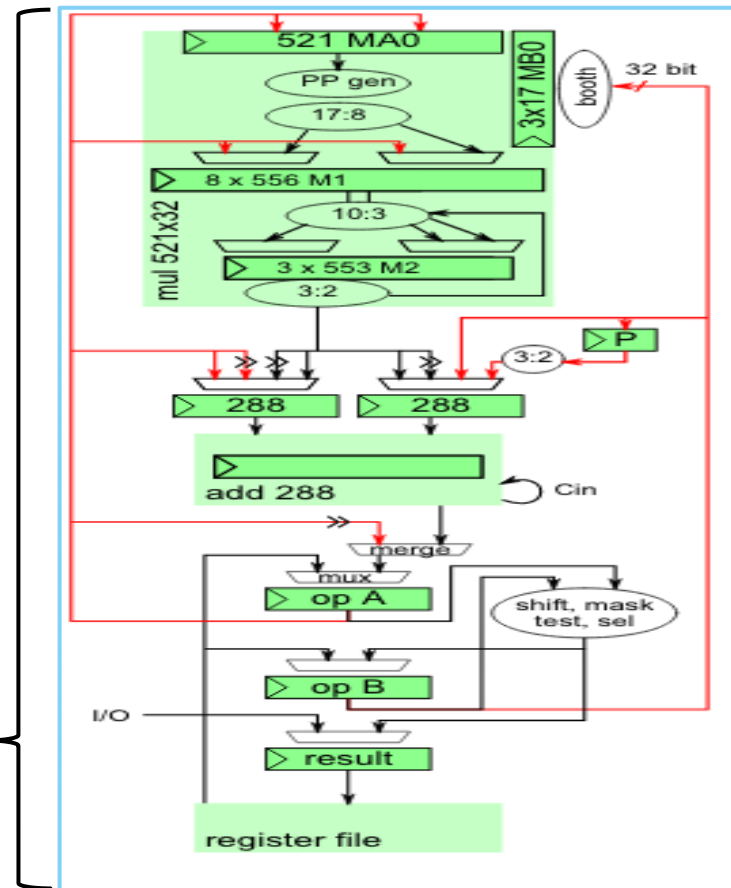
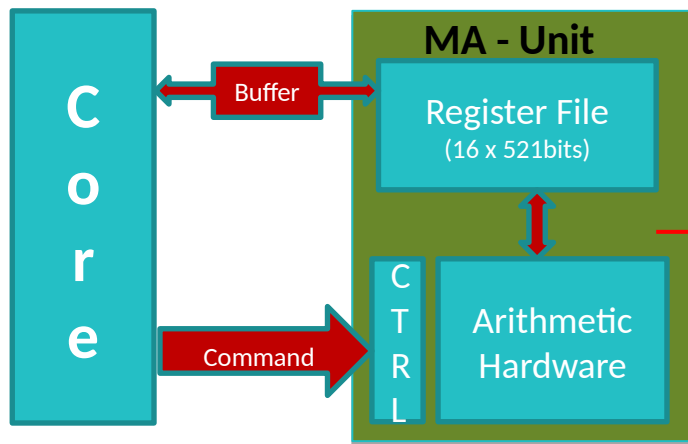
Modulo Arithmetic (ECC) Acceleration:

On-Core Elliptic Curve Cryptography Acceleration

- Software visible instruction for Sign, Verify, & Scalar Multiply operations
- Modular Arithmetic support for:
 - NIST P256, P384, & P521 Curves
 - Edwards 448, & 25519 Curves
 - Generic 521b P Curve

In-order, scalar, non-speculative execution

- All operations happen post completion
- No register renaming, branch wrong, etc...
- Firmware controlled, internal micro-instruction set driven
- A few DWs of input & result = dozens to hundreds of mod-p math steps



Modulo Arithmetic (ECC) Acceleration:

Internal Micro-Instruction Set

Modular Arithmetic

P256, P384, P521, X448, X25519, generic P

- Add
- Subtract under mask
- Halve
- Multiply
- (Multiplicative inverse)

Unsigned Binary Arithmetic

256b, 512b, 521b

- Add with/without carry
- Sub with/without borrow
- Multiply

Logical (521b)

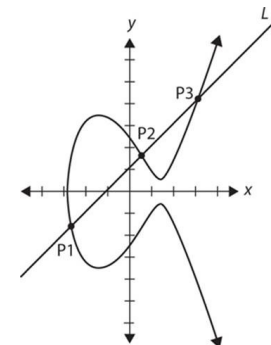
- AND, OR, XOR, NOT

Shift operations (521b)

- Shift by bit
- Shift by 32b word

Support

- Mask and Test (521b)
- Select (521b)
- Load
- Store
- Modify flags



Speedup

Curve	Speedup (PKS)	Speedup (PKV)
Prime 256-bit	21.8x	14.5x
Prime 521-bit	17.5x	11.4x

Improvement of Public Key Sign (PKS) and Public Key Verify (PKV) operations on z15 using the on-core Modulo Arithmetic Accelerator vs z14 using a Crypto Express6 PCI accelerator*

*The speedup is a combination of faster processing and the reduced latency of synchronous on core execution.

Sources:

IBM z14 Cryptographic Performance (<https://www.ibm.com/downloads/cas/K4AR1NLO>)

IBM z15 Cryptographic Performance (<https://www.ibm.com/downloads/cas/6K2653EJ>)

Secure Execution for Linux - Overview

Elevated privileges of an Operating System or Hypervisor can be abused by

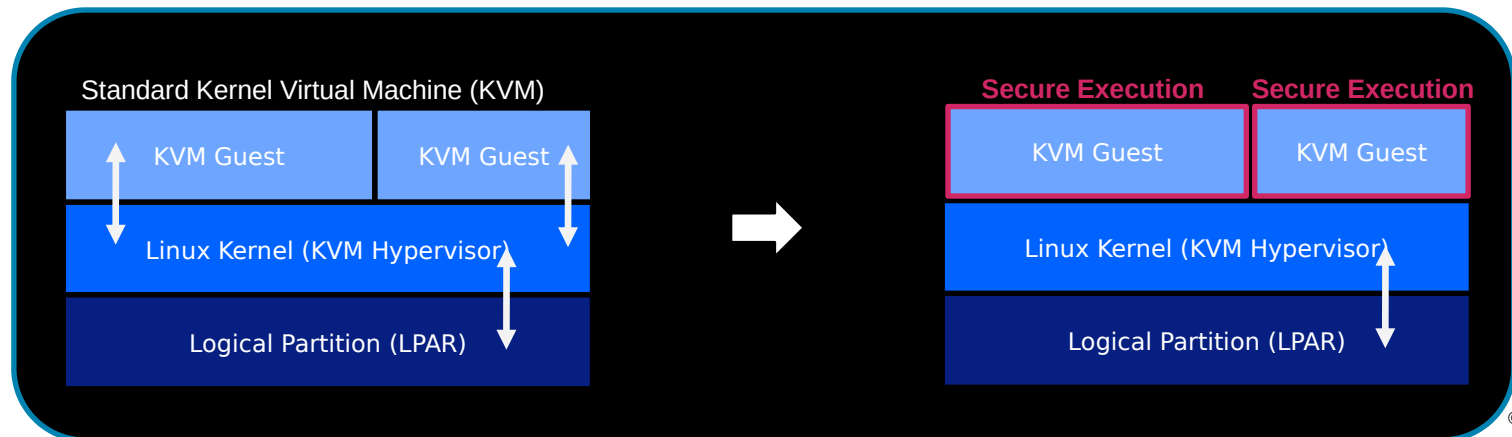
- Malicious system administrators
- Attackers that exploit flaws in system software

Secure Execution provides both horizontal and vertical container isolation.

- Specialized mode in the CPU/Memory
- Only the associated secure guest can see its data/execution state in the clear

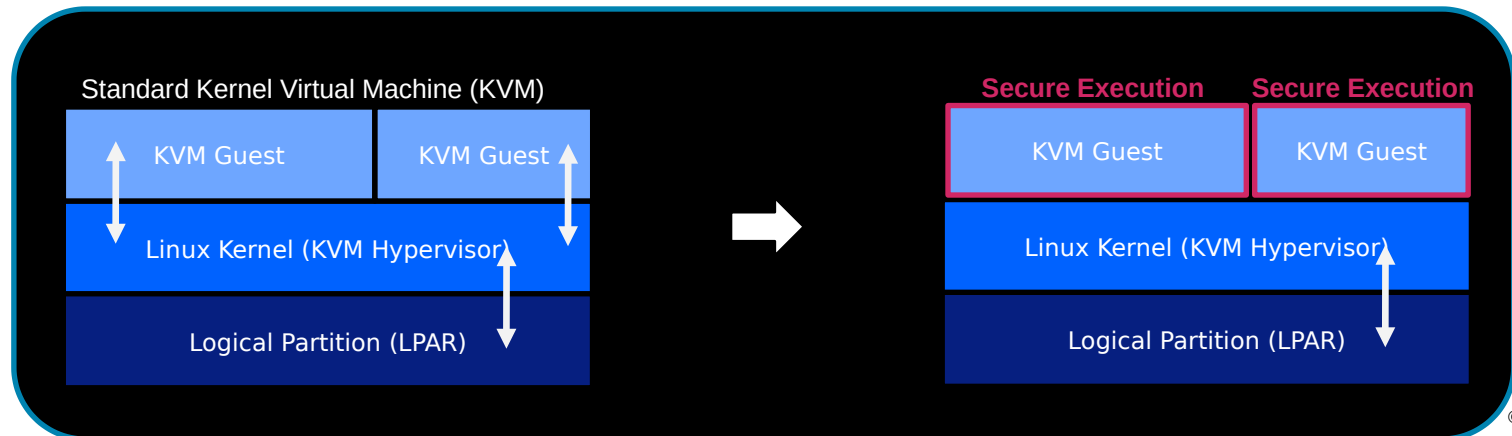
No changes are required to the container application code for exploitation

No additional restrictions exist for secured guests compared to non-secured guests



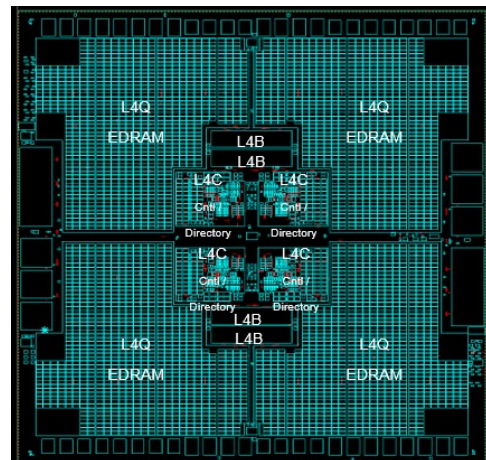
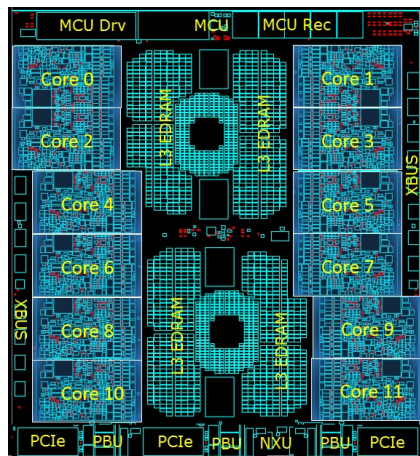
Secure Execution for Linux - Details

- KVM guest memory and execution state are protected by trusted hardware and firmware
- Secure memory can only be accessed in secure mode
- In secure mode, instructions are only executed from the secure memory of the guest
- Unique IDs distinguish and isolate secure memory images from each other
- A new trusted firmware layer called the Ultravisor sits between the hardware and hypervisor
- The Ultravisor encrypts memory blocks before export (paging), and decrypts them on import
- A saved off integrity hash and import/export count prevents using blocks that have been tampered with while paged out



IBM z15 – designed for massive scale commercial workloads

- Processor Chip w/ L3 cache + System Control Chip w/ L4 cache
- 14nm SOI technology, 5.2GHz water cooled enterprise server
 - CP: 9.2 billion transistors, 14.5 miles of wire
 - SC: 12.2 billion transistors, 13.5 miles of wire
- Up to 240 physical cores in 5-drawer shared-memory SMP
- 190 configurable customer CPUs, plus IO assist and firmware CPUs
- 14% single thread speedup & 25% capacity growth vs z14
- Micro-architectural and architectural enhancements for wide variety of workloads



Thank You!



[Anthony Saporito](#)

Senior Technical Staff Member
IBM Systems Hardware Development

IBM Poughkeepsie, NY

saporit@us.ibm.com