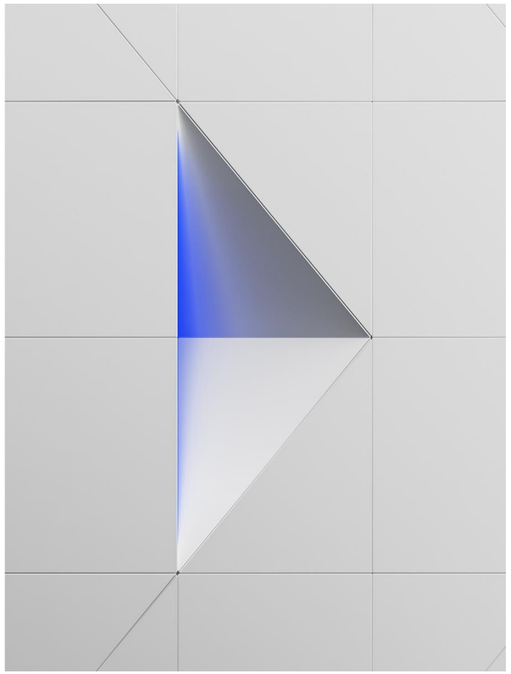
IBM Z Performance Hot Topics

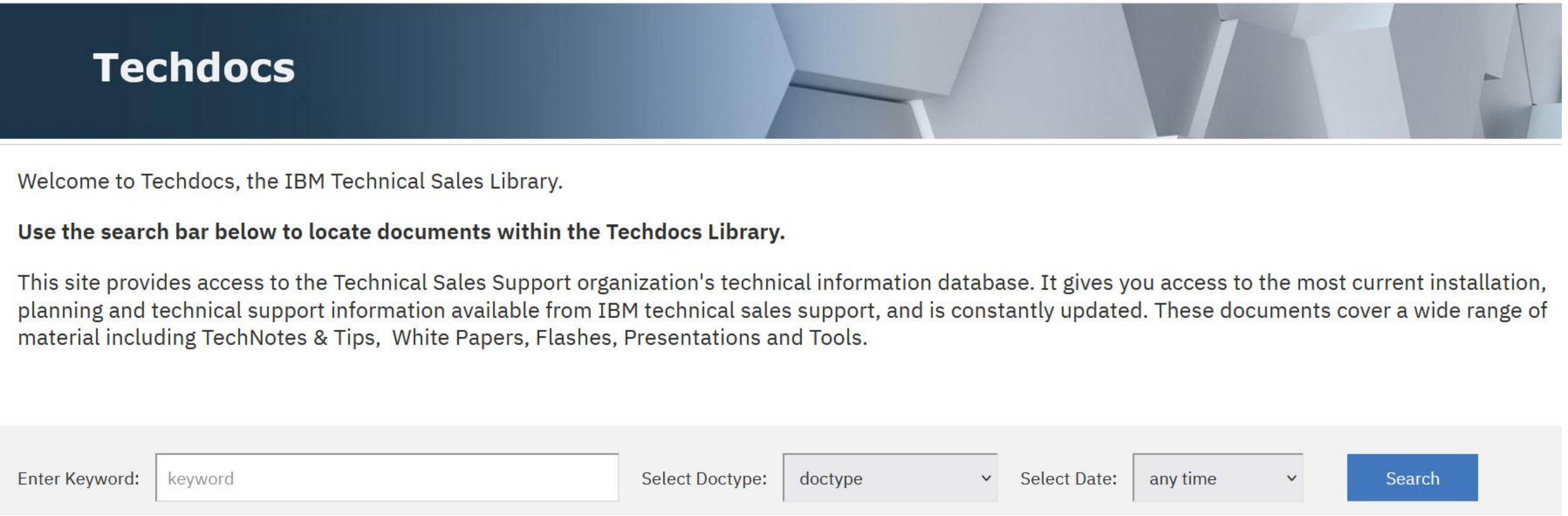
Bradley Snyder
Bradley.Snyder@us.ibm.com

Contents



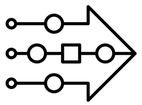
IBM Techdocs

https://www.ibm.com/support/pages/ibm-techdocs-technical-sales-library



IBM z17 makes more possible

Enhance your hybrid cloud with advanced AI where it matters most



Fueling innovation and growth with AI

- Deliver business growth with deep insights
- Unleash the power of generative AI on data*
- Drive greater impact from your most important data

Automating and transforming for efficiency

- Accelerate mainframe application modernization with generative AI*
- Streamline IT Operations with AI insight
- Boost productivity and automate tasks

Securing the most important data

- Proactively identify and mitigate cyber threats
- Quantum-safe protection
- Simplify compliance and improve productivity

* Upon IBM Spyre Accelerator availability

IBM z17 availability Dates – Driver Level 61

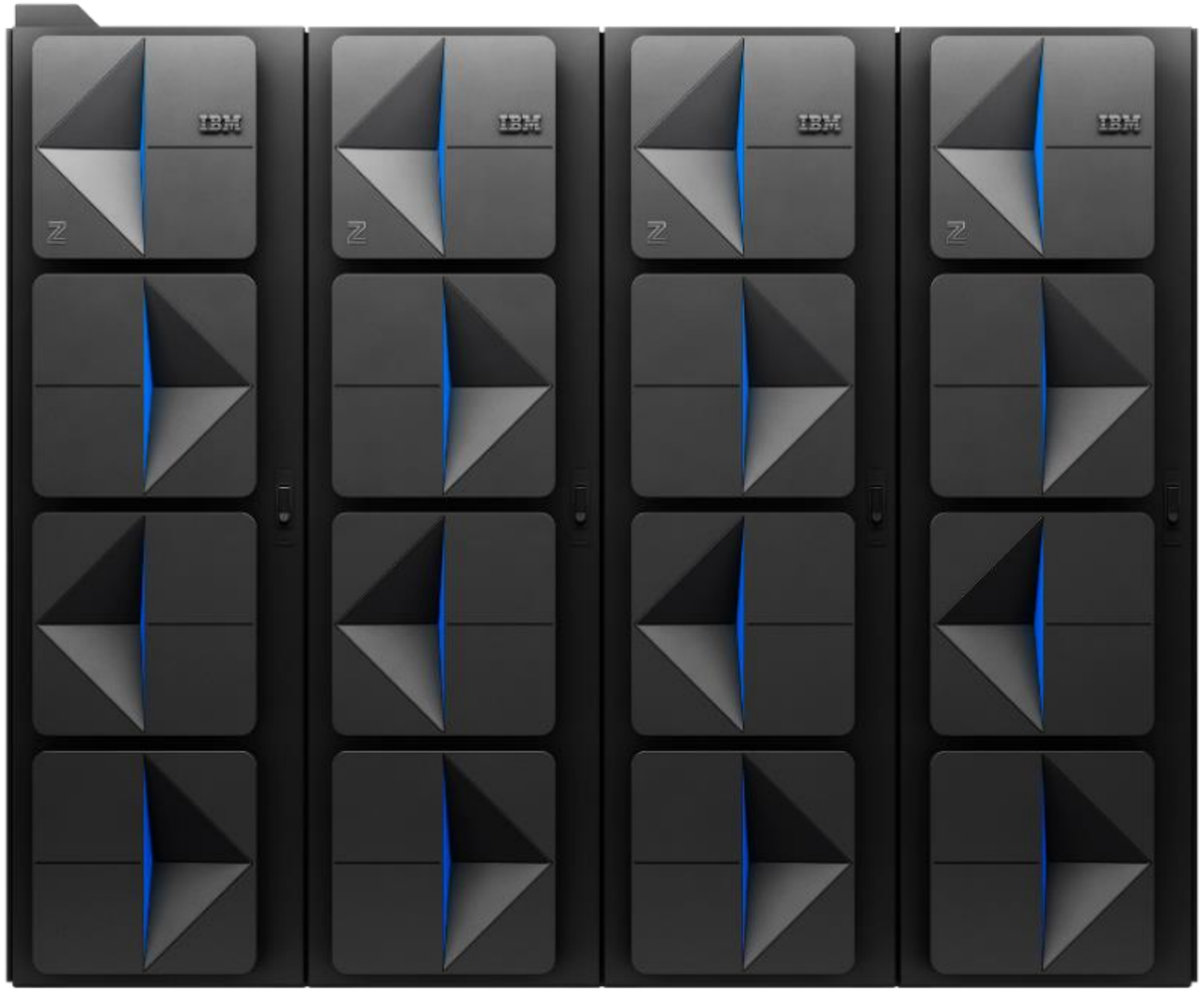
IBM z17

General Availability – Jun 18th, 20251

New features and functions for the IBM z17

Machine Type: 9175

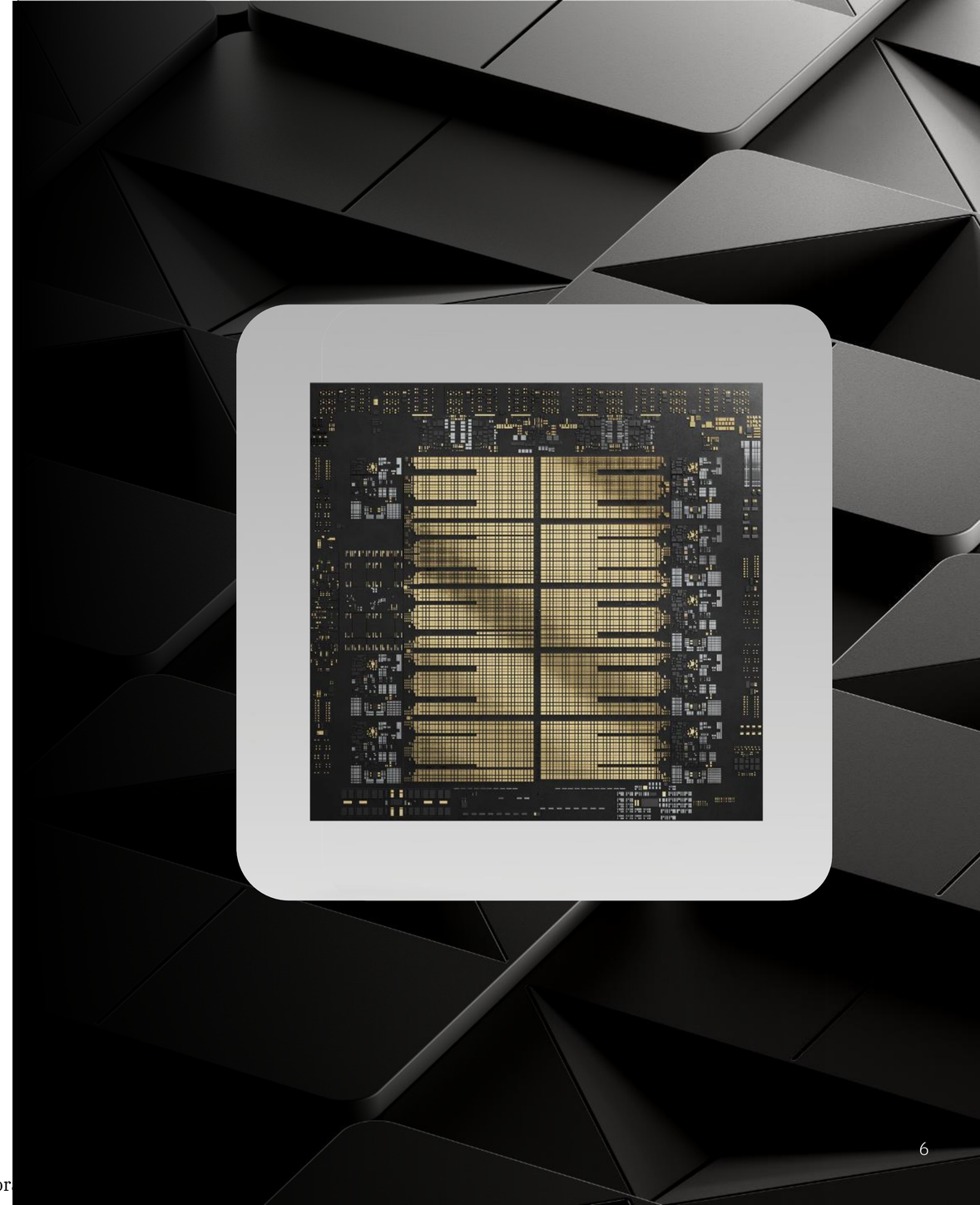
– IBM z17 Model ME1:
  • Features: Max43, Max90, Max136, Max183, Max208
– IBM z15 T01 water- or radiator-cooled upgrades to IBM z17 ME1.
– IBM z16 A01 upgrades to IBM z17 ME1.
– Field installed features and conversions on IBM z17 that are delivered solely through a modification to the machine's Licensed Internal Code (LIC)
– TKE 10.1 LIC (FC 0883)
– TKE HW (new order, w/4770 Cryptographic Adapter):
  • Rack mount: FC 0057
  • Tower: FC 0058
– System Recovery Boost (Stage 3) enhancements
– Validated Boot for z/OS
– ECKD support for Linux secure1 boot



NOTE: GA 1.0 Level: Driver 61 Bundle S04

IBM Telum II Processor

- 5nm technology, 5.5GHz
- 8 cores with 20% area reduction and improved microprocessor power management
- 40% more cache per core
- 24.1 Miles of wire per chip
- 43.0 Billion transistors
- **NEW:** On-chip Data Processing Unit (DPU): Increased I/O performance with 70% reduction in power for I/O management, RAS, reduced latency
- 2nd-gen AI Accelerator for high-speed inferencing with fine tuning
- 8x dedicated AI processing per core



More AI acceleration on IBM z17

In-transaction AI with encoder LLMs and multiple AI model techniques

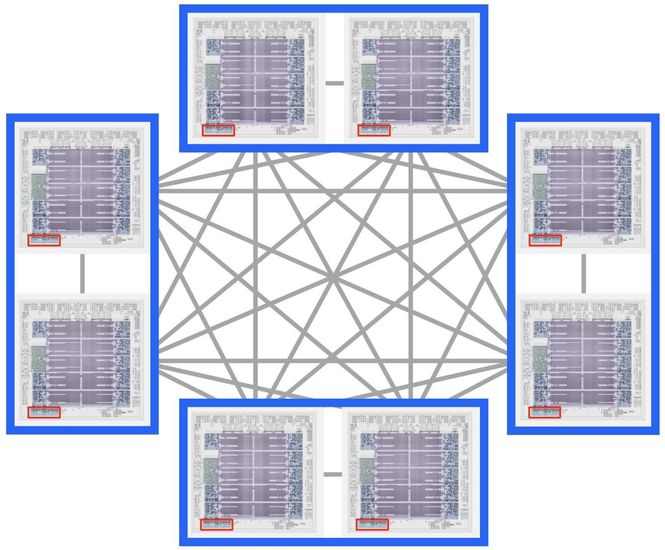
2nd Gen on-chip AI accelerator in Telum II

- Support for LLM compute primitives
- Improved quantization and matrix operations
- Improved AI processing over IBM z16

AI workload balancing during peak usage

In-drawer intelligent routing

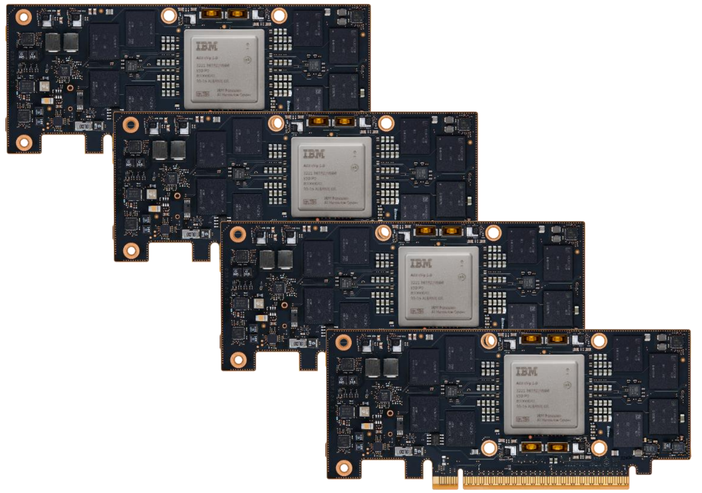
- Remote AI processing
- Up to 8x AI processing available



Optimize generative AI and LLM use cases

IBM Spyre Accelerator cards*

- 32 Gen AI-ready cores per adapter card
- Up to 48 adapter cards per system



*Available starting in 4Q25

IBM z17 At a Glance

Model ME1		
Feature	**Customer PUs**	**Max Memory**
Max208**	208	64 TB
Max183**	183	64 TB
Max136	136	48 TB
Max90	90	32 TB
Max43	43	16 TB

** Factory only models

Machine Type: 9175
One Model – ME1
– One, two, three, or four 19” Frames (A, B, C, and Z)
Five features
– Max43, Max90, Max136, Max183** and Max208**
Processor Units (PUs)
– 52 (59 for Max208) PU cores per CPC drawer
– 2 spares designated per system
– Dependent on the Feature - up to 43, 90, 136, 183, 208 PU cores available for characterization
  - Central Processors (CPs), Internal Coupling Facility (ICFs), Integrated Facility for Linux (IFLs), IBM Z Integrated Information Processor (zIIP), System Assist Processors (SAPs) and Integrated Firmware Processor (IFP)
  - 85 LPARs
– Sub-capacity available for up to 43 CPs
  - 3 sub-capacity points (4xx, 5xx, 6xx)
Memory
– 6 x 2 x 8 channel Reed Solomon RAIM Memory design
– System Minimum of 512 GB
– Up to 16 TB per drawer
– Up to 64 TB for System and up to 32 TB per LPAR (OS dependent)
  - 884 GB Fixed HSA, standard
  - 64/128/256/512/1024/2048 GB increments
– Virtual Flash Memory (512GB increments, Max 6TB)

I/O
– Up to 48 PCIe+ Gen4 Fanouts -- 2 port @32 GBps each and Integrated Coupling Adapters PCIe Gen3 -- 2 port @ 8 GBps per system
– 6 Logical Channel Subsystems (LCSSs)
  - 4 Sub-channel sets per LCSS
– Server Time Protocol (STP) Optional w/ PPS, NTP and PTP ETS support w/ connectivity direct to CPC (not to SE), plus security enhancements

IBM z17 system highlights

- 1-4 , 19” frames
- 208 customer cores
- IBM Telum II processor
- Improved I/O subsystem
- Less weight & floor space

11% more performance — Single thread performance vs IBM z16[2]

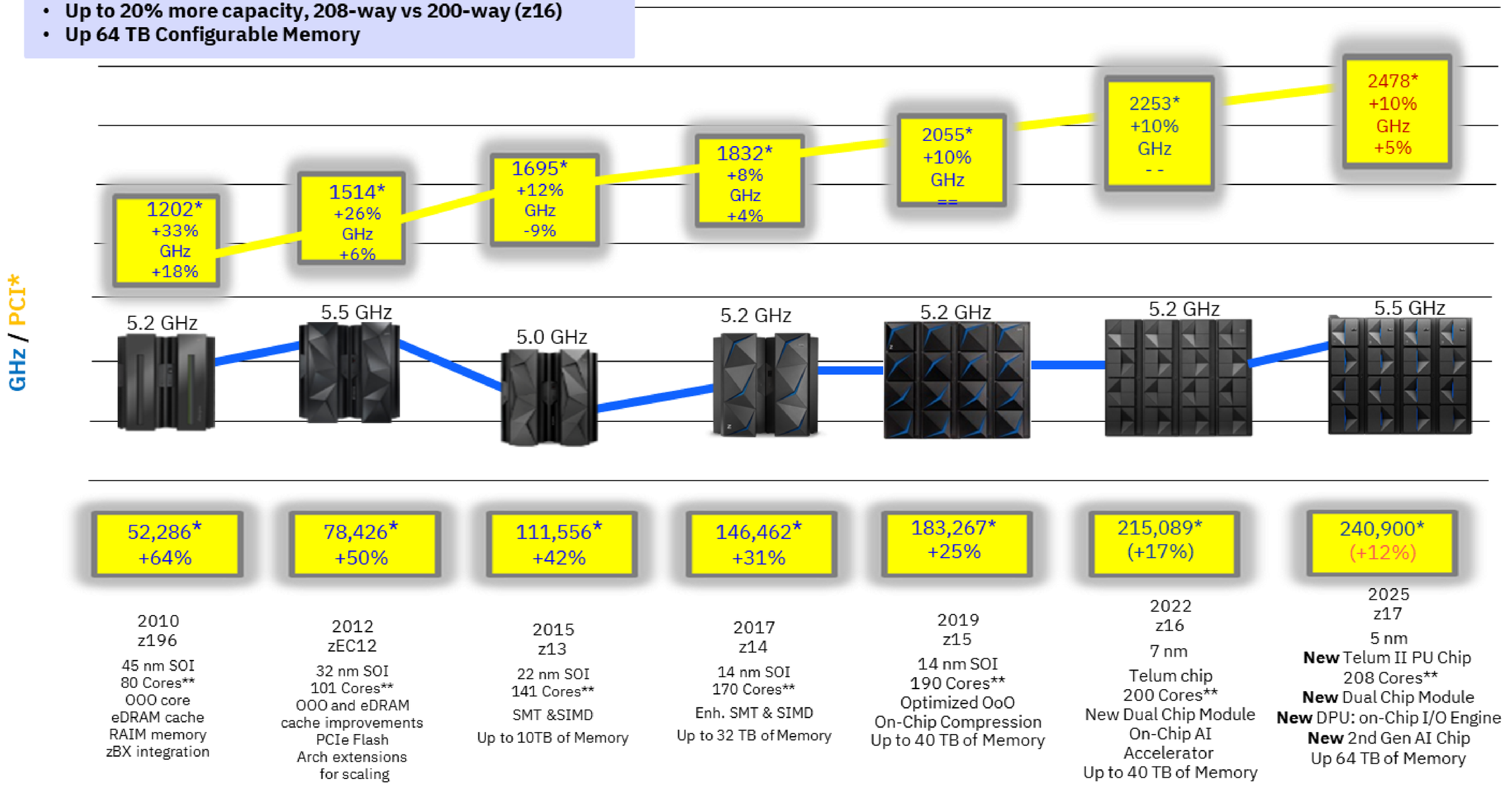
15-20% capacity growth — vs IBM z16[3]

60% more memory — Up to 64 TB memory

19% less power — than IBM z16[4]

IBM z17 Continues the CMOS Mainframe Heritage

- **Up to 20% more capacity, 208-way vs 200-way (z16)**
- **Up 64 TB Configurable Memory**



* MIPS Tables are NOT adequate for making comparisons of IBM Z processors. Additional capacity planning required

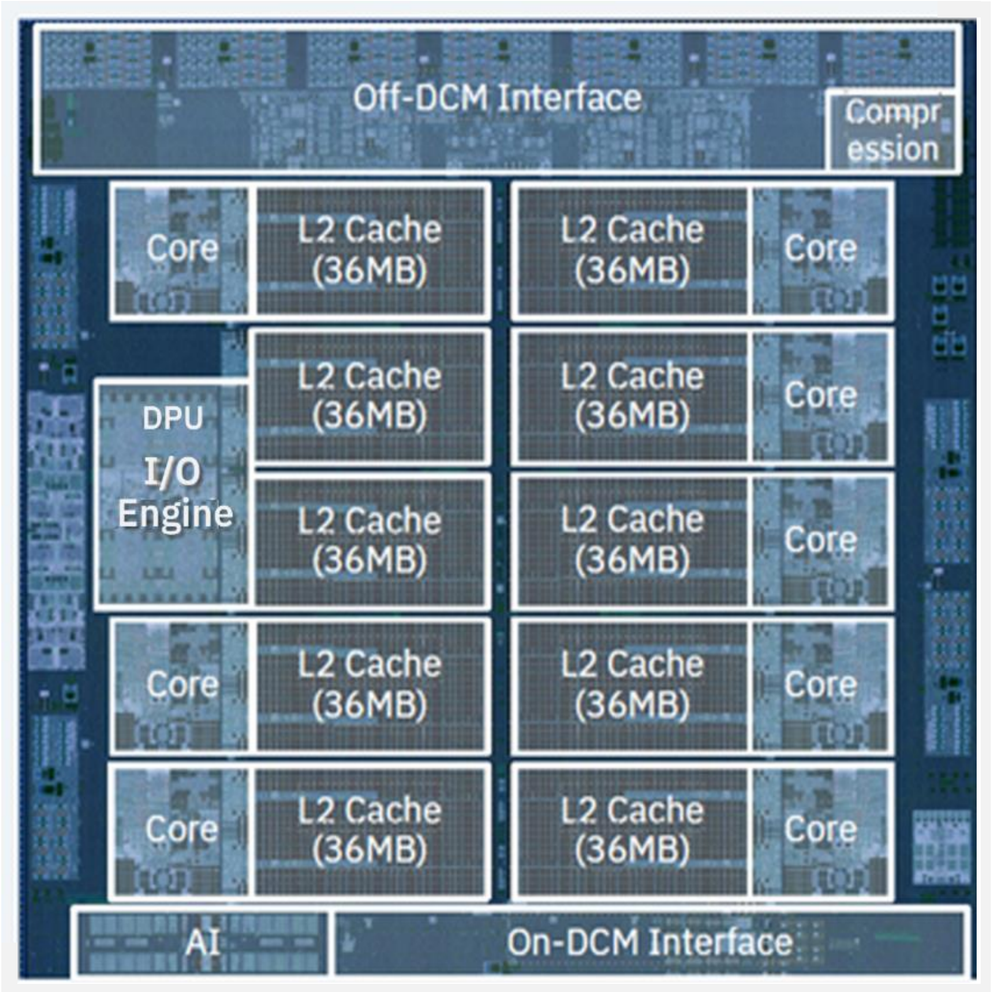
** Number of PU cores for customer use

IBM z17 Processor Design
I/O Engine (DPU)

Shift in I/O Adapter Management
– Designed to move I/O processing to firmware
– Supports new I/O Adapter innovations
– Improved RAS

One IO Engine per PU
– Uses the space of 1.5 IBM z17 cores
– Supports
  • FC/FCP/zHPF
  • OSA (EQDIO)
  • OSC

Legacy I/O Adapters management remains unchanged



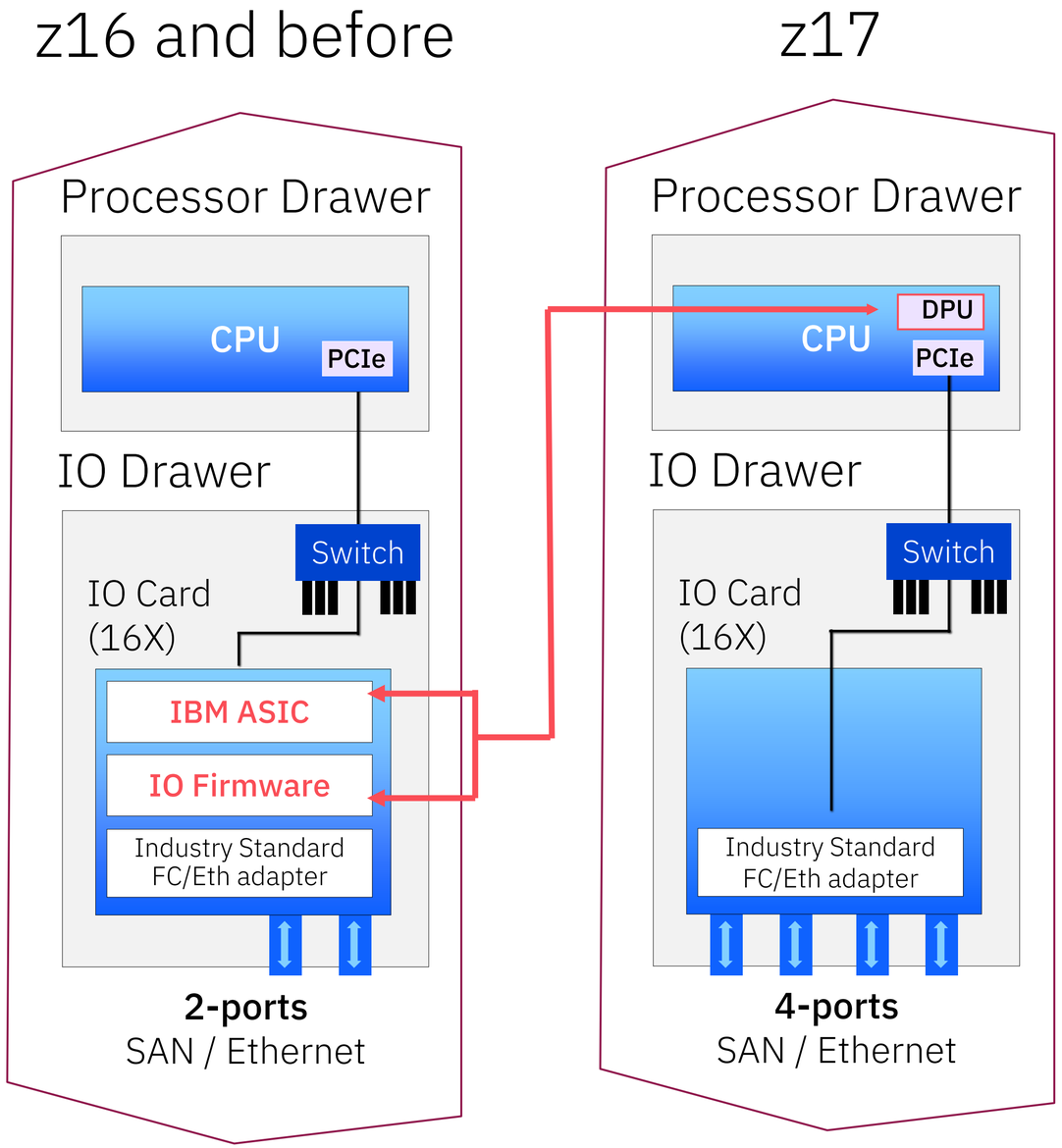
IBM z17 I/O Refactoring - Next-gen I/O Infrastructure

The I/O subsystem of IBM Z is built on a solid architectural basis that has provided a long and rich history of I/O (storage and networking) function, throughput and bandwidth improvements, while maintaining application backward-compatibility

Earlier I/O infrastructure is constrained in terms of being able to deliver:
- improved latency
- additional bandwidth capability
- enhanced functionality without changes to the current physical packaging and (in the case of networking) architecture

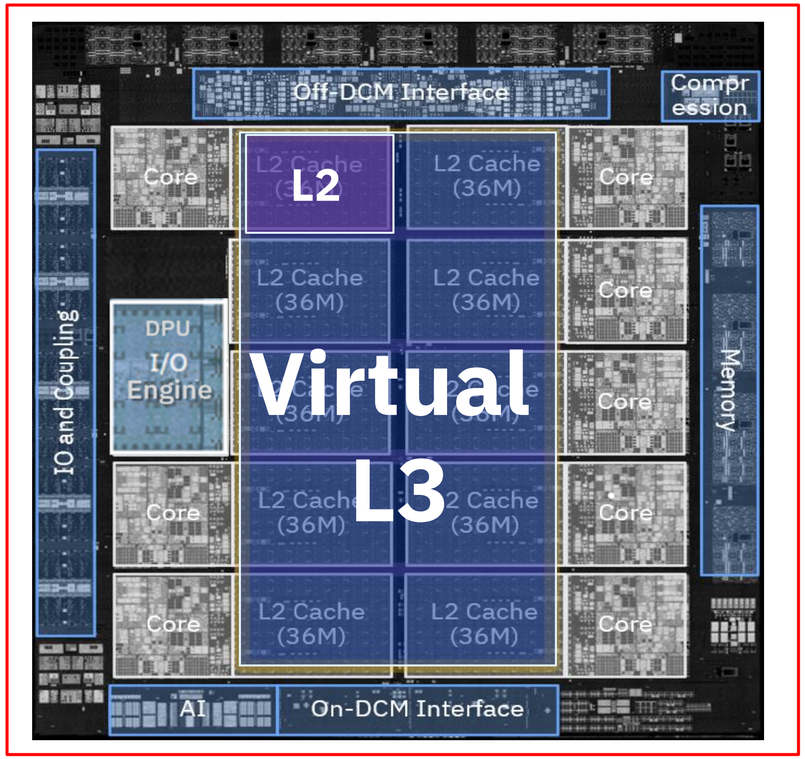
The IO firmware stack has evolved on ASICs platforms over the past 30 years. In our IBM z17 system, we are pulling that I/O ASIC functionality, across the PCI bus, and into the mainframe processor chip – much like the on-chip AIU engine in z16, and the on-chip compression acceleration engine in z15.

This will allow IBM Z to:
- Provide better I/O Performance/Latency/RAS
- Use higher I/O density (4-port FICON cards and converged network adapters)
- Be agile in delivery of new I/O feature function every generation. We always put out a new processor chip, so we get a chance to innovate on the I/O each time, versus once every ~10 years

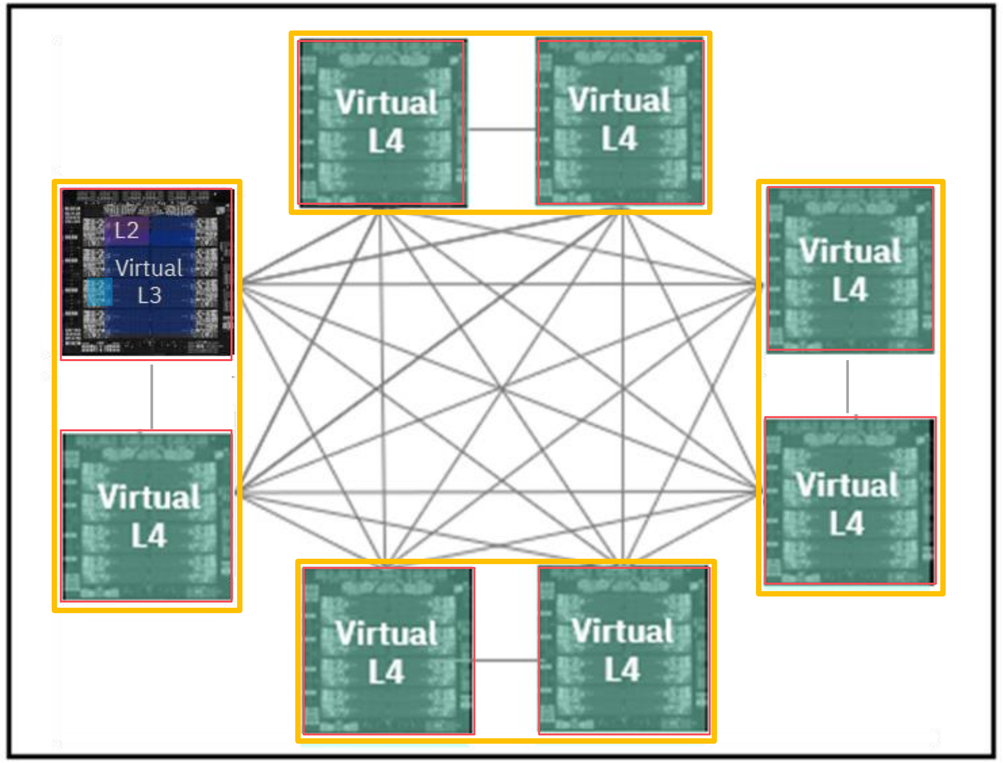


IBM z17 cache Topology Visuals from a single core's perspective

PU Chip



Drawer: 4x **DCMs** with 2x **PU Chips** each



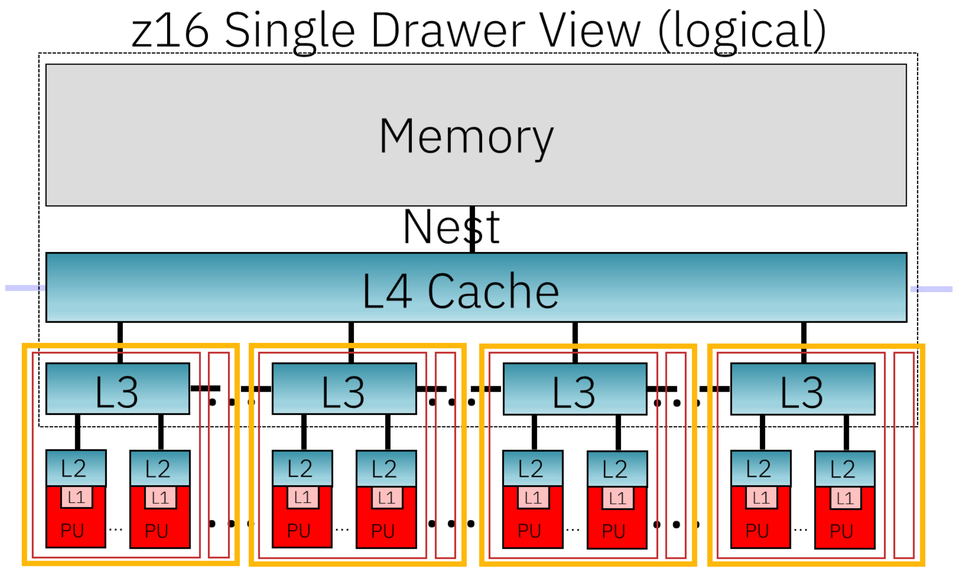
IBM z17 vs IBM z16 Hardware Comparison

IBM z16

CPU 5.2 GHz

Caches
L1 private 128k i, 128k d / core
L2 *private* 32 MB unified / core
virtual L3 up to 7x32 = 224 MB / CP chip
virtual L4 up to 8x32x7 = 1.75 GB / drawer

Topology
8 (core + L3)s / CP chip
2 CP chips / DCM
4 DCMs (64 engines) / drawer
4 drawers / CEC

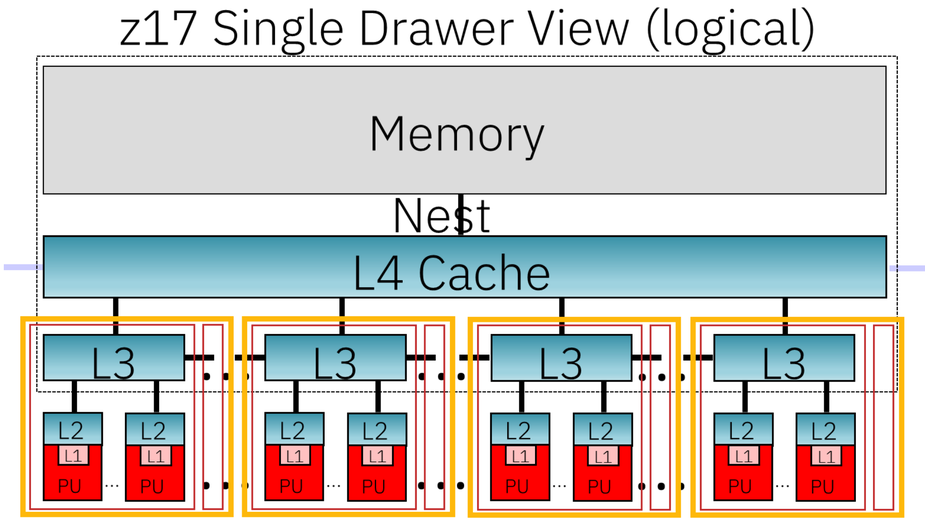


IBM z17

CPU **5.5 GHz**

Caches
L1 private 128k i, 128k d / core
L2 private **36 MB** unified / core
virtual L3 shared up to **9**x36 = **324 MB** /CP chip
virtual L4 shared up to 10x36x**7** = **2.52 GB** / drawer

Topology
8 (core + L3)s / CP chip
2 CP chips / DCM
4 DCMs (64 engines) / drawer
4 drawers / CEC



Processor Unit (Core) Locations: Customer, SAP, IFP and Spares

IBM z17		1st Drawer				2nd Drawer				3rd Drawer				4th Drawer			
Feature	Cust PUs	Cust PUs	SAP	IFP	Spare	Cust PUs	SAP	IFP	Spare	Cust PUs	SAP	IFP	Spare	Cust PUs	SAP	IFP	Spare
Max208	208	49	6	2	2	53	6	0	0	53	6	0	0	53	6	0	0
Max183	183	42	6	2	2	47	5	0	0	47	5	0	0	47	5	0	0
Max136	136	46	6	0	0	43	5	2	2	47	5	0	0				
Max90	90	43	5	2	2	47	5	0	0								
Max43	43	43	5	2	2												

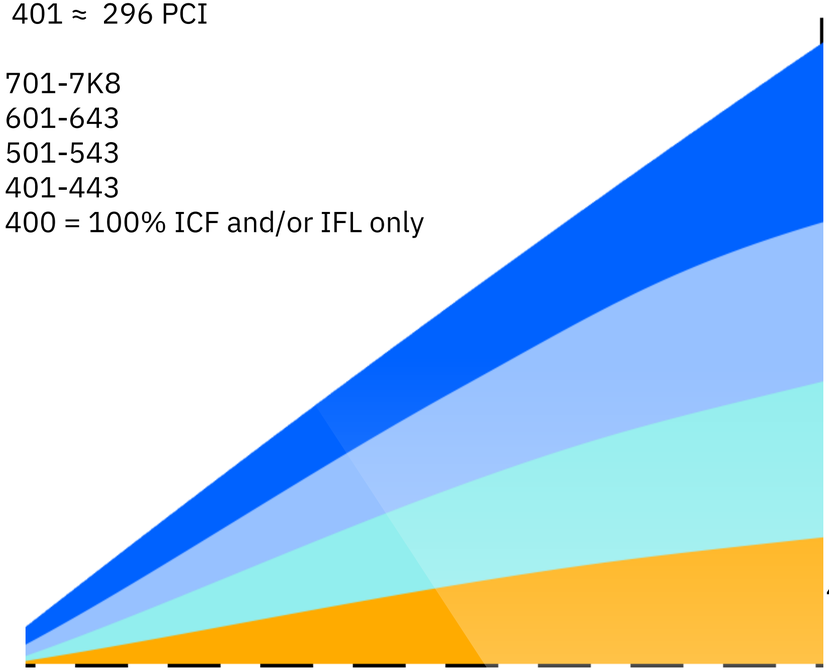
- PUs can be purchased as CPs, IFLs, Unassigned IFLs, zIIPs, or ICFs
  - Any un-configured PU can act as an additional Spare PU
  - CPs and zIIPs initial placement in 1st drawer working up
  - IFLs and ICFs initial placement in highest order drawer working down
- Upgrades available from any lower feature to any higher any models
  - Achieved via Concurrent Drawer Add from Max43 to Max125
  - Upgrade to Max183 and Max208 from any other feature are ***not supported***. Max183 and Max208 are factory built only.

IBM z17 Full and Sub-Capacity CP Offerings

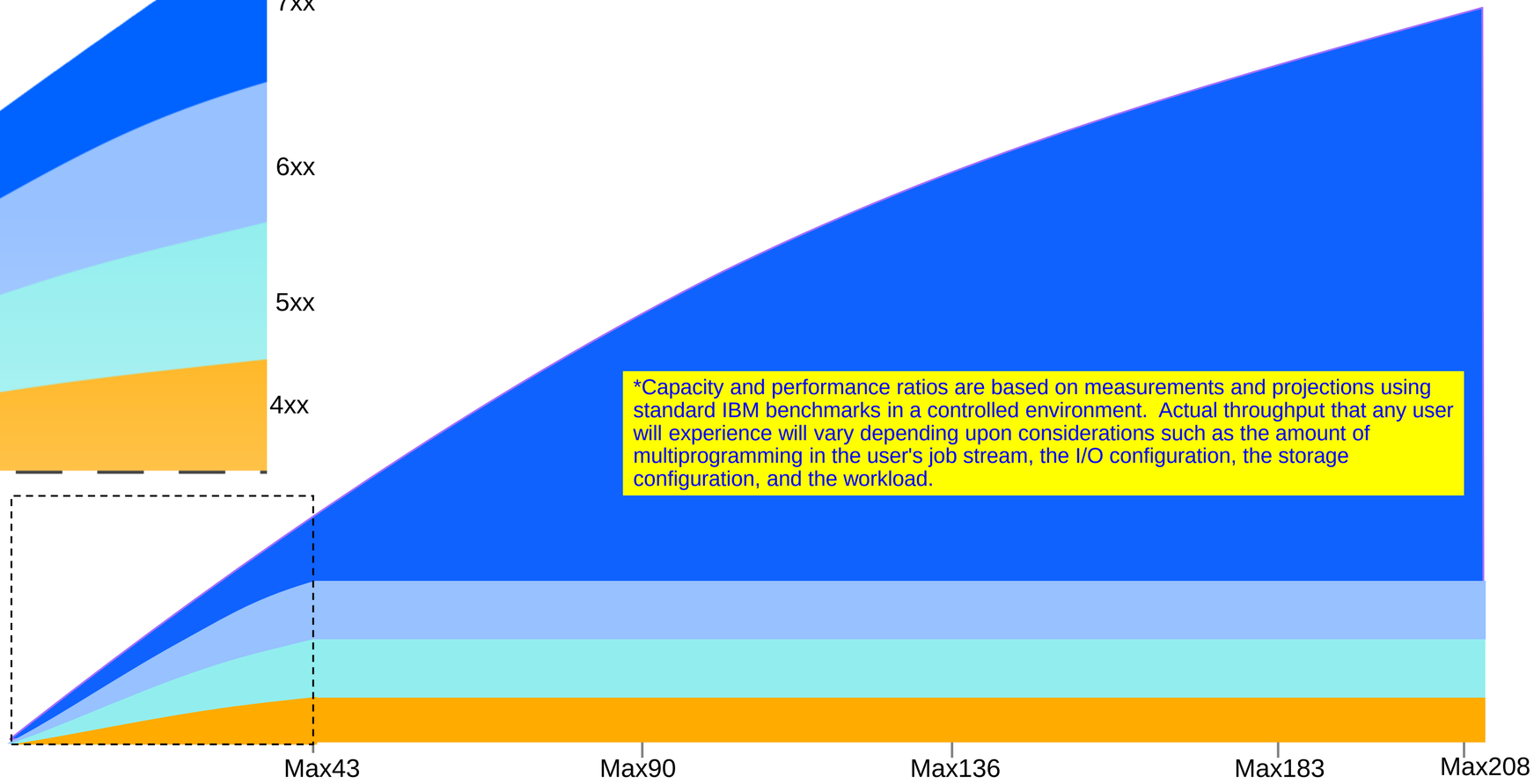
CP Capacity – Relative to Full Capacity Uni
701 ≈ 2478 PCI (IBM MIPS)
601 ≈ 1642 PCI
501 ≈ 1030PCI
401 ≈ 296 PCI

701-7K8
601-643
501-543
401-443
400 = 100% ICF and/or IFL only

Sub capacity CPs, up to 43 may be ordered.
If more CPs are ordered all must be full 7xx capacity.
All CPs on an ME1 CPC have the same capacity (except during Recovery Boost).
All specialty engines are full capacity.



MSU Sub Capacity



*Capacity and performance ratios are based on measurements and projections using standard IBM benchmarks in a controlled environment. 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.

Critical Migration actions for IBM z17

Any migration from a previous IBM Z system to IBM z17 requires a capacity study

- zPCR at a minimum, recommend zCP3000 study for complete analysis

---

CPU MF Counters MUST be enabled for every partition being migrated

- Enabled on both before and after
- Critical migration action for every IBM z17 candidate (z/OS and z/VM)

---

z/OS LPARs must enable SMF 98.1 records

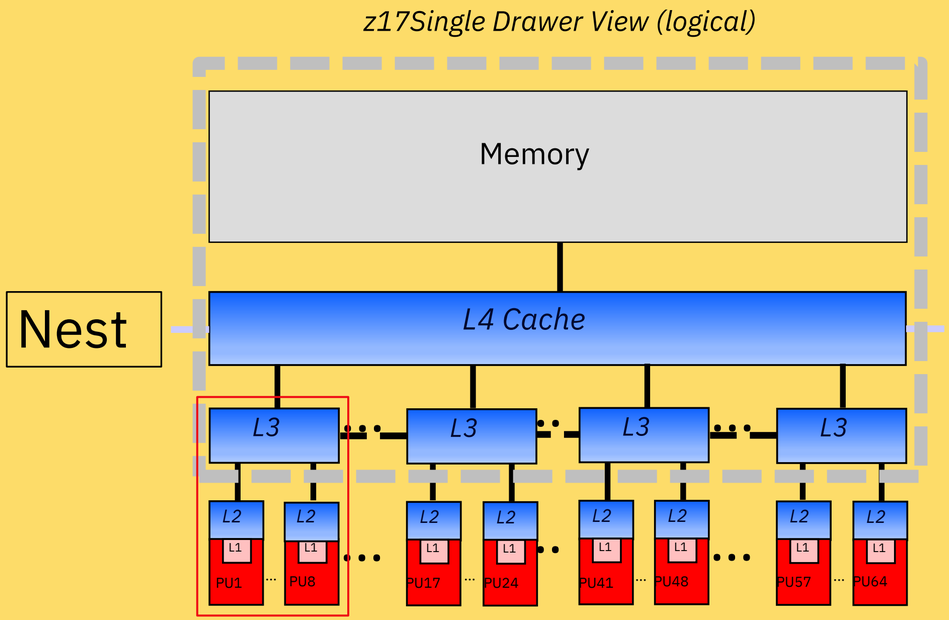
- SMF data must be saved on both before and after environments

IBM System Z Community Notes

CPU MF IBM Z Best Practice

– z/OS SMF 113 records introduced January 2009 with the IBM System z10

– The IBM Best Practice is to enable the CPU MF records on all IBM Z LPARs



z/OS z17™ Migration “Volunteers”

– Used to validate LSPR Workloads

– Opportunity to ensure your data is used to influence future server design

– **Looking for “Volunteers”**

- Need Before and After Data
- 3 days, 24 hours/day
- SMF 70, 71, 72, 99.14, 113.1 per LPAR
- Production partitions preferred
- Data as close to migration as possible

– No deliverable will be returned

If interested send note to:

stephanie.deluca@ibm.com

Second Sesson on May 5: CPU MF Update for IBM z17 by Stephanie DeLuca

IBM z17 operating system support

z/OS

z/OS V3.1 with PTFs
z/OS V2.5 with PTFs
z/OS V2.4 with PTFs for toleration only, Extended Support Required

z/VM

z/VM 7.4 with PTFs
z/VM 7.3 with PTFs

z/VSE

VSE[n] V6.3.1 (+PTFs) – 21st Century Software

z/TPF

- z/TPF 1.1 with PTFs

Linux on IBM Z

Minimum Distributions:

SUSE SLES 16.1 (Post GA)
SUSE SLES 15.6 (GA)
SUSE SLES 12.5 (Post GA)
Red Hat RHEL 10.0 (Post GA)
Red Hat RHEL 9.4
Red Hat RHEL 8.10
Red Hat RHEL 7.9 (Post GA)
Canonical Ubuntu 24.04 LTS (Post GA)
Canonical Ubuntu 22.04 LTS (Post GA)
Canonical Ubuntu 20.04 LTS (Post GA)

IBM cannot legally discuss IBM z17 ME1 exploitation prior to GA from distributors.

Officially Tested list here.

MIPS is MIPS is MIPS

The performance benefit of a new processor, this time the IBM z17 does not lead to a decrease in the MIPS requirement for your workload

Better performance claims represent one of two things
- Decrease in wall clock time
  - *What took 10 seconds of CPU time now takes 9*
- The ability to do more work in the same amount of time

The only way to decrease MIPS requirement is looking for/utilizing new technology
- If there is IIPCP time add capacity
- Recompile workloads to the latest architecture level of the hardware
- Identify SIIS, add memory if there is Demand Paging
- Etc.

IBM Washington Systems Center

MIPS is MIPS is MIPS is MIPS Except when it’s NOT

MIPS is MIPS is MIPS
Except when it’s NOT

All PCI/MIPS tables are derived by measurements done in Poughkeepsie with processors running at ~92% busy

Example: Workload running on IBM z16-720 and total CEC is only 25% busy
– Driving 5 CPs worth of work
Simple math says this workload is consuming 25% of 32,903 MIPS (8,226 MIPS)

Really this 5 CPs of work is really consuming ~10,000 MIPS

Processor	CPs	MIPS (assuming AVG)	MIPS per CP	Change in MIPS/CP
3931-720	20	32,903	1,645	
3931-705	5	10,129	2,026	23%

Processor Sizing

Lots of inputs needed for a processor sizing

Including but not limited to:
- Worklaod Type (from SMF 113 records)
- Base LPAR configuration
- Target LPAR configuration
- Busy of current box
- Expected busy of destination box
- System check needed to ensure no other bottlenecks

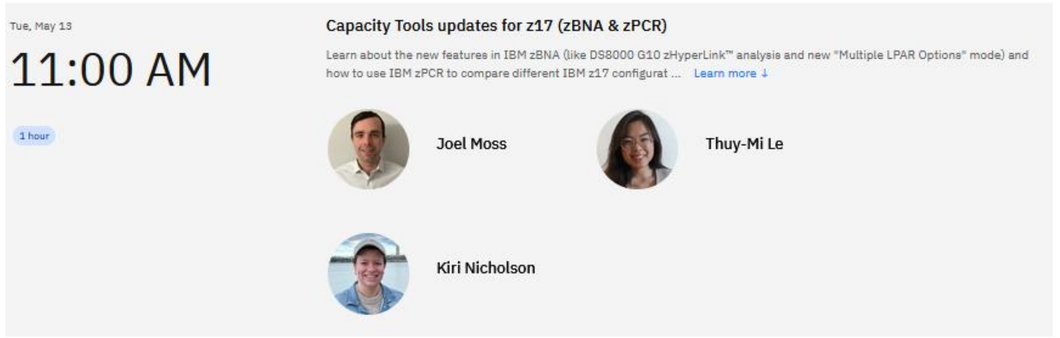
In depth work needed to be done with IBM or IBM BP to ensure proper sizing and expectations are set with any processor upgrade

Use of zPCR as a base, better to have full zCP3000 study done

Latest Versions of IBM zBNA and IBM zPCR Support IBM z17

zBNA Version 2.7.0

zPCR Version 2.7.0



Why CPU MF Data is Important

One of the main drivers for determining CPU Capacity
- Along with LPAR configuration

Load a configuration from EDF to get CPU MF information along with LPAR configuration
- EDF created from ZCP3KEXTR
- When loading, take notice of the ‘remove parked CPs’ option

z/OS-3.1 LSPR Data (04/08/2025)

LSPR Multi-Image Capacity Ratios

IBM Z General Purpose CPs

Values are applicable for z/OS; representative of z/VM, KVM, and Linux

Capacity basis: 2094-701 @ 559.792 MIPS for a typical multi-partition configuration

Capacity for z/OS on z10 and later processors is represented with HiperDispatch turned ON

IBM Z Processor	Features	Flag	MSU	LSPR Workload Category				
				Low	Low-Avg	Average	Avg-High	High
z17/700								
9175-701	1W	=	306	2,448	2,462	2,477	2,439	2,403
9175-702	2W	=	582	4,806	4,768	4,731	4,623	4,520
9175-703	3W	=	852	7,099	7,021	6,944	6,748	6,564
9175-704	4W	=	1,112	9,331	9,223	9,116	8,817	8,537
9175-705	5W	=	1,360	11,526	11,368	11,213	10,811	10,436
9175-706	6W	=	1,600	13,690	13,467	13,250	12,747	12,281
9175-707	7W	=	1,836	15,822	15,520	15,229	14,628	14,073
9175-708	8W	=	2,060	17,924	17,529	17,151	16,456	15,815
9175-709	9W	=	2,281	19,999	19,497	19,020	18,233	17,509
9175-710	10W	=	2,497	22,044	21,424	20,837	19,961	19,156
9175-711	11W	=	2,702	24,061	23,309	22,602	21,640	20,757
9175-712	12W	=	2,894	26,050	25,154	24,317	23,272	22,312

Proper Processor Planning

The actual capacity of any IBM Z processor is heavily dependent on the LPAR configuration and workload category of each partition

Key is to **set proper expectation** when comparing base to proposed configuration

MUST use **zPCR** (or **zCP3000** from IBM or IBM BP) to set proper capacity values for current and proposed processors

Partition Configuration	GCP Capacity	Change from Base
Base Value IBM z17 714	**30,707**	
Configuration 1	**31,099**	**+1.3%**
Configuration 2	**30,670**	**-0.1%**
Configuration 3	**27,387**	**-10.9%**

Upated LPAR Optimization Recommendation

Old recommendation
- Have at least two additional CPs defined to a partition above what's needed for its weight (fair share CPs) called 'moderate'
- Example, LPAR fair share is 7.2 CPs, define 10 LCPs for partition (8 VH/VM and 2 VL)
- 'minimal' configuration only has enough LCPs to meet weight

Allowed moderate amount of 'white space' for workload spikes when other partitions were not using their weight

New recommendation takes fair share CPs into account
- Large number of small partitions all getting 2 additional LCPs leads to too high a number of vertical low CPs for any configuration
- Some partitions may only see a single additional LCP with this optimization
- Partitions with fair share of at least 8 LCPs still recommend 2 additional LCPs

Goal is for smaller partitions to have at least 25% additional capacity above weight to handle spikes

zPCR Support for New Recommendation

Prior versions of zPCR had ‘Minimal’ and ‘Moderate’ options

- Minimal reduced LCPs to what is needed for weight
- Moderate was 2 additional LCP above weight requirement

New optimization routine takes weight (fair share CPs) into account

- Large number of small partitions all getting 2 additional LCPs leads to too high a number of vertical low CPs for any configuration
- Now smaller partitions may only see a single additional LCP with this optimization
- Partitions with fair share of 8 LCPs will always get 2 additional LCPs

Can still manually change the number of LCPs for any partition to fit your needs

New LPAR Optimization Examples

LPAR	Fair Share CPs	Original LCP definition	Prior Minimal	Prior Moderate	New Recommended Optimization
SYS1	0.6	1	1	3	1
SYS2	0.6	3	2	3	2
SYS3	1.5	4	2	4	2
SYS4	9.3	15	10	12	12

z/OS 3.2 Preview*

Updates to z/OS WLM Policy Adviser
- Updates for easier invocation, reading of current policy and importing of SMF data
- Updated color scheme for reading PI charts

Workload Resiliency Check
- Check to ensure consumption in SYSSTC does not reach above specified threshold
- Ensure system availability by not having high CPU consuming address spaces in SYSSTC

Workload Classification Pricing
- Workload marking and measurement in z17 and
- Tailored Fit Pricing for Hardware on zIIPs
- Expected for IBM and vendor products
  - Classification allows for different pricing for workloads and zIIPs
  - Clients not expected to classify executables
- Some stored in SMF record 1156
- Reporting handled in RMF 70 (1) and 70 (2)

* Planned, some features may change by GA

Claims/Disclaimers

1. The IBM z17 Telum II processor is designed to seamlessly scale peak AI workloads as each core on the chip can access each of the 8 integrated accelerators for AI. By allowing routing of inference requests to any idle IBM Integrated Accelerators for AI within the same drawer, the IBM Integrated Accelerator for AI can increase inference throughput by up to 7.5x as compared to IBM z16.. DISCLAIMER: Performance results are based on internal tests exploiting the IBM Integrated Accelerator for AI for inference operations on IBM z16 and z17. On IBM z17, each IBM Integrated Accelerator for AI allows any CPU within a drawer to direct AI inference request to any of the 8 idle AI accelerators on the same drawer. The tests involved running inference operations on 8 parallel threads with batch size of 1. Both IBM z16 and z17 were configured with 2 GCPs, 4 zIIPs with SMT and 256 GB memory on IBM z/OS V3R1 with IBM Z Deep Learning Compiler 4.3.0, using a synthetic credit card fraud detection model (https://github.com/IBM/ai-on-z-fraud-detection). Results may vary.

2 Single processor capacity of IBM z17 for equal n-way at common client configurations is approximately 11% greater than on IBM z16 with some variation based on workload and configuration. **DISCLAIMER:** Based on internal measurements. Results may vary by customer based on individual workload, configuration and software levels. Visit LSPR website for more details at: http://www.ibm.com/support/pages/ibm-z-large-systems-performance-reference

3.Within each single drawer, IBM z17 provides 20% greater capacity than IBM z16 for standard models and 15% greater capacity on the max config model, enabling efficient scaling of partitions. **DISCLAIMER:** Based on internal measurements. Standard models include IBM z17 Max43, Max90, Max136 and Max183 and IBM z16 Max39, Max82, Max125 and Max168. Max config models are IBM z17 Max208 and IBM z16 Max200. Results may vary by customer based on individual workload, configuration and software levels. Visit LSPR website for more details at: http://www.ibm.com/support/pages/ibm-z-large-systems-performance-reference

4. The anticipated typical IBM z17 system is estimated to reduce system power consumption by approximately 19% compared to a similarly configured IBM z16 system. DISCLAIMER: Based on an expected typical IBM z17 system configuration based on the actual historical average IBM z16 system configuration. IBM z17 is Max90 with 8TB memory, 82 active processors, 5 ICA-SR 2.0, and 3 PCIe+ I/O drawers with 43 I/O adapters. The IBM z16 is configured to provide the same hardware capability. Power consumption is based on the Power Estimation Tool, available at ibm.com/support/resourcelink/api/content/public/PowerEstimationTool-legacy.html. Results may vary.

Thank you!!

IBM Spyre™ Accelerator is
in tech preview until Q42025