Preface: IBM z15 Design and Technology

The digital transformation continues to have a profound effect on business. The 2020 global pandemic is creating and accelerating transformation of business activities, processes, competencies, and models. To succeed, businesses must embrace this digital transformation, adopting agile processes and new technologies to deliver services and experiences that customers and clients demand. This special issue of the *IBM Journal of Research and Development* showcases the design and technology innovations of IBM z15, a transformational system with an all new modular and scalable system design in an industry-standard form factor, which was co-created with our clients and business partners via IBM Design Thinking and Agile processes, enabling security, privacy and resiliency at scale to fuel your digital transformation and journey to the Cloud. We have a collection of 16 articles from our leading technologists and experts. The new capabilities covered in this Journal demonstrate the continued excellence of IBM technical innovations and leadership in Hybrid Cloud, Security, Data Protection, and Resiliency across the stack.

The first article in the issue, by Mayer et al., concerns IBM z15 in supporting users to deliver mission-critical workloads and services in a hybrid cloud environment. This article presents an overview of new capabilities introduced with IBM z15 that enable a seamless integration of the platform into private, public, and multi-cloud environments. This work is rooted in the broader IBM cloud strategy and point-of-view and based on ongoing efforts to identify and address critical pain-points for a key set of enterprise clients through applying Enterprise Design Thinking practices. By outlining the hardware, firmware, and software support added for IBM z15, the authors show how they have been able to integrate the IBM Z platform into the IBM Cloud infrastructure. They also discuss how cloud-native technologies and tooling enable users to access, deploy, and lifecycle manage z/OS resources and services for a seamless cloud experience.

Bradbury et al. describe the security model of IBM Secure Execution, the functionality of the hardware and ultravisor, as well as the required changes to the hypervisor in order to support protected virtual machines. With the growth of IBM Z and LinuxONE in both private and public clouds, customers are expecting their workloads and data to have the same levels of security, isolation, and privacy as running on-premise. In order to achieve these levels of trust, the IBM z15 and LinuxONE III provide the IBM Secure Execution for Linux facility that isolates customers' data from each other, as well as from cloud administrators, and requires no application changes. Unlike other solutions in the industry, IBM Secure Execution does not require remote attestation thus simplifying the deployment of applications into the protected environment. Also, unlike some other solutions in the industry, the integrity of data is protected end-to-end, that is, from the boot image on disk to memory as it is paged by the hypervisor and throughout execution. Isolation and integrity are provided by the hardware and trusted firmware known as the ultravisor.

Morris et al. emphasize that system security is a focus area for all IT infrastructure providers. New system features like pervasive encryption, the transition to cloud-based offerings, and the future for quantum-safe platforms demand increased cryptographic performance as well as more cryptographic agility. The new IBM 4769 Cryptographic Coprocessor addresses these trends. It brings performance improvements that match the requirements of the new IBM z15. A combination of newly available features allows IBM z15 to scale to greater than 5,000 Virtual Hardware secure modules (HSMs) per system and makes it suitable to support virtualized client environments such as cloud-scale datacenters. To meet the dense packaging and energy requirements of those data centers, the form factor and power consumption of the card were reduced significantly. The card also offers an expanded set of algorithms to support state-of-the-art as well as future workloads. For the first time, the user interface provides access to a selected set of quantum-safe algorithms. Infrastructure extensions add hardware-embedded, attestation-friendly trusted boot services that improve system resiliency by providing hardware-enabled measurements of the secure and trusted boot process. These extensions simultaneously simplify the security certifications built on them. In this article, the authors provide an overview of the IBM 4769 Cryptographic Coprocessor, highlighting security characteristics, internal hardware, form factor, and enhanced firmware.

Driever et al. consider Fibre Channel to be the premier enterprise storage transport, where an organization's most sensitive data flows over Fibre Channel links within and across data centers. Controlling access to and security of data within the enterprise can prove to be a formidable task, with increased complexity and management overhead encountered as the granularity of access control is increased. The authors discuss a new, easy-to-deploy innovation for Fibre Channel connections that ensures data is exchanged only between trusted servers and storage controllers, while also enabling the integrity and confidentiality of the data in flight between the trusted entities with negligible impact on transactional performance. This article explains how the components of IBM Fibre Channel Endpoint Security are configured to work together to provide protection from insider threats, requiring minimal steps to deploy, fully controlled via policy, and transparent to applications, middleware, and operating systems.

Digital Object Identifier: 10.1147/JRD.2020.3009463

Webel et al. discuss IBM z15 Selfboot and Secure Boot and describe the basic hardware and firmware concepts that are implemented and enabled for the IBM z15 central processor (CP) and system control (SC) chips. These chips contain hardware and firmware to serve Selfboot and Secure Boot needs. Selfboot initializes the CP/SC chips from hardware and firmware that reside in each chip module. This establishes a core root of trust, and also guarantees a boot time that is independent of the system configuration, which is key for large enterprise-class systems consisting of multiple drawers and chips. Secure Boot is built on this core root of trust and is used to authenticate the firmware loaded from system memory prior to execution of that firmware. Selfboot and Secure Boot also guarantee the integrity of the CP and SC chips by restricting hardware and memory accesses through debug or service interfaces during boot, runtime, and code update phases.

Surman et al. discuss in detail how System Recovery Boost on the IBM z15 server expedites both planned and unplanned restarting of partitions, including shutdown, initial program load (IPL), middleware startup, and client workload recovery execution that follows, to accelerate service restoration and mitigate the impact of any downtime. It does this by providing limited-duration "boost periods" that deliver significant usable additional processor capacity and parallelism. On sub-capacity machine models, it provides a boost in processor speed by running the general-purpose (GP) processors at full-capacity speed, for the boosting logical partitions (LPARs) only, and only during the boost periods. It makes all available processing capacity defined to the boosting images available to process any kind of work, "blurring" general-purpose processor and zIIP engines together during the boost period. System Recovery Boost also expedites and parallelizes processor reconfiguration actions that may be part of the client's overall restart and recovery process, as orchestrated by Geographically Dispersed Parallel Sysple (GDPS) automation. Optionally, System Recovery Boost provides the ability to add additional processor capacity from the client's unused "dark cores" via activation of a new type of temporary capacity record. All of this can be accomplished without increasing the client's IBM software billing costs or the processor consumption associated with the client's workload during these boost periods.

Saporito et al. discuss the design of the IBM z15 microprocessor, and how the latest-generation IBM Z processor provides enhanced performance and compute capacity as compared to its IBM z14 predecessor. This article describes some of the major improvements in both process and design, including out-of-order load-and-store sequencing, single-instruction multiple-data (SIMD) and floating point enhancements, a new modulo arithmetic engine for accelerating elliptic curve cryptography, a hardware sort accelerator, and a workflow that modernized the development of these features. Outside of the central processing unit (CPU), the cache sizes have increased on all levels, and each processor chip now contains 12 CPUs. System topology changes have been introduced allowing up to five drawers to exist in a fully populated system. The processor cache subsystem includes numerous improvements in the area of fetch, store, and cache management policies aimed at speeding up both traditional data serving workloads and highly virtualized environments alike.

Berry et al. explain how IBM z15 was designed in the same 14-nm high-performance GlobalFoundries technology as the IBM z14 and yet still added 20% more cores to the chip, doubled the L3 cache, increased the L2 cache by a third, while also adding a third PCIe port to the chip and an Elliptic Curve Cryptography engine into each core. This article discusses the many design, tool, and methodology enhancements required to increase the design content so significantly while maintaining the chip size and power limits from the previous z14 design. The authors also discuss other design and methodology improvements that were made possible via the deeper understanding of the technology and how to more fully leverage it in a second generation.

The IBM Z processor continues to improve over previous System Z processors, but for the first time, it does so without a technology improvement as the baseline enabler.

Klein et al. discuss the concept, implementation, and verification of DEFLATE compliant compression acceleration in z15 across both hardware and firmware. This article illustrates various challenges that result from incorporating complex data-dependent and data-intense functionality like DEFLATE as an architected instruction and discusses how solutions in hardware/firmware co-design have been applied to overcome these challenges. The IBM z15 processor chip contains a new hardware component to perform DEFLATE compliant compression and decompression. The Integrated Accelerator for zEnterprise Data Compression is based on a high-frequency DEFLATE pipeline and includes a hardware generator for Dynamic Huffman Tables. Accessible as an architected instruction, this engine has been designed for straight-forward exploitation by software and is easily available to any application in the problem state. A brand-new hardware/firmware integration model has been developed to provide this complex functionality without imposing restrictions on data patterns or data sizes and without impacting system responsiveness.

Sofia et al. discuss the integration of DEFLATE acceleration in z15 into the z/OS software stack in both synchronous and asynchronous mode and present the resulting performance for selected workloads in this article. IBM z15 replaces the former I/O attached accelerator for DEFLATE, zEnterprise Data Compression (zEDC) Express, with an on-chip accelerator that can be

synchronously accessed via an instruction. The integration of this new accelerator in the z/OS software stack has been designed to maintain a consistent user experience for software packages that used the previous technology, while still allowing the enhanced aspects of the new technology to deliver additional value. Two different access paths for DEFLATE have been created in z/OS to accomplish both goals. For user space programs that utilize the zlib application programming interface (API), z/OS directly executes the instruction synchronously, which avoids overhead and reduces latency. Authorized users continue to utilize existing infrastructure and have the Service Assist Processors (SAP) perform compression in an asynchronous fashion on their behalf. The SAP receives information about the requested task via a thin and efficient communication path to z/OS, invokes the instruction in a well-defined fashion, and returns the result to z/OS.

Somasundaram et al. discuss Partition placement by Processor Resource/System Manager (PR/SM)—in particular, the changes made to the PR/SM heuristic placement algorithm for z15 and how it surmounts the problems inherent for optimal placement of logical partitions. Every new machine generation of IBM Z brings with it an increase in number of physical processors and memory capacity. Some generations can also bring change in the physical configuration of the server. The z15, for example, can have from one to five drawers instead of a maximum of four on the z14. Another difference is that z15 has two fixed chips per node versus the two or three chips per node on z14. The logical partitions on the other hand can come in various configurations, including "Dedicated" logical partition, shared "Hiperdispatch = YES" logical partition, and shared "Hiperdispatch = NO" partition. Each of the partition types can request as many logical processors and memory as the machine generation will allow, which is usually less than the physical resources available on the machine. The optimal placement of logical partitions on the physical server, given its configuration, is an NP-Hard problem. Memory access latency and cache usage play vital roles in the performance of logical partitions, and it is imperative that placement is optimal. Moreover, on z15, the integrated facility for Linux (IFL) and internal coupling facility (ICF) processors can be moved from one chip to another, during reoptimization of partition placement, in addition to GP and zIIP processors that are already allowed to be moved, compounding the placement problem.

Guendert et al. focus on Sysplex Time Synchronization using the IEEE 1588 Precision Time Protocol (PTP) topic. Timekeeping, and highly accurate, precise time synchronization, is a key requirement for modern information technology systems. While true for several industries, this is especially true for industries involved in transaction processing such as the financial industry. As such, the IBM Z Sysplex needs highly accurate timing/timekeeping and synchronization technology to ensure data integrity, and to also provide the ability to reconstruct a database based upon logs. Recently enacted changes and new regulatory requirements, both in Europe and in the United States have brought increasing attention to time synchronization accuracy. These regulations spurred an interest, both from IBM Z and from our IBM Z clients, in the IEEE 1588 PTP being implemented in IBM Z. This article explains the history of PTP, PTP technology, the regulations that led IBM to introduce PTP to IBM Z, PTP's implementation on IBM Z, and IBM's involvement and leadership in the development of the PTP technology and standards going forward.

Valentine et al. present the innovative design basics of the IBM Z Hardware Management Appliance (HMA) feature, which has significant positive impact potential for both clients and IBM. IBM z15 supports the new IBM Z HMA optional feature that provides redundant Hardware Management Consoles (HMCs) and Support Elements (SEs) that run on redundant physical servers inside the Central Processor Complex (CPC) frame. This eliminates the need for having to manage one or more separate physical servers for HMCs outside of the frame. The authors describe the host HMC/KVM (Kernel-based Virtual Machine) and Virtual SE environment that is completely managed by IBM Z firmware as a true appliance. This article also illustrates the Firmware Integrity Monitoring environment for the host HMC/KVM extended to the Virtual SE to provide Secure Boot protection for firmware and continuous monitoring and utilization of a shared Trusted Platform Module (TPM). It describes how physical system errors are processed by Problem Analysis (PA) firmware running on the Virtual SE, thereby enabling online guided repair instructions running on the Virtual SE to be used for both HMC and SE detected errors.

McCain et al. discuss how the product development cycle is being transformed with "Artificial Intelligence" (AI) for the first time in IBM Z history. This new era of AI, under the project name IBM Z Development Transformation (zDT), has allowed the team to grow and learn new skills in data science. This transformation forces change structurally in how data is prepared and stored. In z14, there were incremental productivity gains with enhancements to automation with eServer Automation Test Solution (eATS) and a technology data analysis engine called zDataAssist. The introduction of AI significantly accelerated the efficiency in z15. In this article, the authors explain how Design Thinking and Agile principles are used to identify areas that are of high impact and feasible to implement: 1) data collection via System Test Event Logging and Analysis engine (STELA), Problem ticket management system (Jupitr), and Processor data analysis engine (Xrings); 2) problem identification, analysis, and management (AutoJup) along with Intelligent Recovery Verification Assistant (IRVA); 3) product design

documentation search engine (AskTheMachine); and 4) prototype microprocessor allocation processes Intelligent Commodity Fulfillment System (ICFS) using machine learning. This article details the approach of these areas for z15, the implementation of their solutions under the zDT project, as well as the results and future work.

Webel et al. discuss proactive power management in IBM z15. The IBM z15 processor power management enhances several on-chip power management techniques over z14 processor with a specific focus on reducing response time for voltage droop management. The IBM z15 processor puts a specific emphasis on proactive voltage droop management strategy to reduce conservative static guard band that is added to the supply voltage in order to protect against worst-case voltage droops. The z15 processor relies on selected events from the earlier stages of a deep pipeline processor as indicators to predict sharp changes in the power consumption over a short period of time. The early information of the selected events allows the processor to throttle the execution flow through the processor pipeline and prevents the sharp power change before it takes place and thus reduces the voltage droop. In z15, as one of the proactive schemes, both the digital power-proxies that are direct indicators of the processor activity and the Critical Path Monitors (CPMs) are combined to give an earlier and proactive indication of voltage droop events. This proactive indication provides enough time for the throttle actuation circuits to prevent the voltage droop. CPMs act as real-time timing margin indicators, and power-proxies act to serve as the activity monitors.

For the final paper, Kostenko et al. discuss improved data center density and energy efficiency, new system packaging, and modeling by the introduction of IBM z15. The authors note that IBM z15 is designed to meet the requirements of a range of data centers, while reducing costs through increased density, configuration flexibility, and cooling efficiency. z15 is a continuation and broadening of the physical transformation of the mainframe that began with IBM z14 ZR1/LR1, which introduced the new "true 19-in" frame. A maximum configuration z15 delivers greater than 30% additional compute capacity per watt than z14, and maintains approximately the same maximum system footprint, while enabling significant floor space reduction for most configurations. z15 introduces the choice of integrated 2N power using either intelligent power distribution units or bulk power, also supporting most data centers including hot/cold-aisle containment, raised-floor and non-raised-floor, and top and bottom-exit I/O and power. z15 supports the ASHRAE A3 (4th Edition) environment, providing efficiency advantages by reducing humidification requirements. z15 maintains the value of a system that is preconfigured/pretested before shipping. Innovations in packaging, I/O cabling, controls, and testing are put in the context of the latest data center trends. The capabilities of new tools to estimate power, weight, airflow, heat extracted to water for water-cooled systems, as well as 3D STEP and CFD models to aid in the planning for the system are described.

It has been a highly rewarding experience working with the authors and the world class IBM Z team to design, develop, and deliver IBM z15, a transformational system in every sense. We hope readers will enjoy and benefit from the content in this special issue. The details provided here describe the new design and technology innovations of z15 across the stack in security, availability, performance, accessibility, and configuration flexibility and convey a compelling message that IBM Z is the industry-leading platform for mission-critical hybrid cloud. We are looking forward to continuing the partnership with our clients in this digital transformation journey.

Miao Zhang-Cohen
IBM z15 Product Owner/Program Manager
IBM Systems
Guest Editor