BUILDING THE DATACENTRIC ENTERPRISE

HOW IBM LINUXONE 5 DRIVES SECURITY, PERFORMANCE, AND COST EFFICIENCY

SUMMARY

We live in a datacentric world that generates data at an unprecedented scale. This is a world in which data drives efficiency through automation across every business function of the modern enterprise and improves how organizations interact with their partners and customers by leveraging AI.

Data has become the most critical asset to the modern business. At the same time, the threat landscape grows more complex, with more bad actors employing more advanced tactics across more attack vectors. With the need for post-quantum cryptography (PQC) on the horizon, there is a real fear that even today's security platforms will fall short of fully protecting the data estate.

An enterprise's need to be more agile in response to the market and to deploy greater measures to support security and resilience can create tension. Enterprise IT organizations already taxed with modernization efforts are seemingly unable to deliver on these critical initiatives simultaneously. Increasing that tension even further is the need to be cost-efficient. Secure platforms that deliver performance without compromise come at a cost, as does the use of the public cloud, which is exceedingly high as data generation increases across the enterprise.

How, then, does an enterprise achieve agile and performant yet secure operations while driving down costs? Is this even possible?

This research brief, sponsored by IBM, will explore this tension and how IT organizations can achieve data environments with the greatest levels of performance and operational agility while increasing an organization's security profile. Further, it will consider how IBM's LinuxONE 5 platform can bring the power of IBM's security and AI capabilities in a cloud-native environment to the most mission-critical, datacentric workloads that populate the enterprise — while also achieving cost-efficiency.

DATA DRIVES THE WORLD

Data is the key driver of strategic decision-making and operational efficiency in the modern business landscape. Organizations generate vast amounts of data and harness historical (old) and real-time (new) data to understand their business, customers, and market comprehensively. The synergy between old and new data is crucial, enabling businesses to leverage insights from past trends while adapting to current market dynamics.

This intelligence-driven approach becomes more effective as more real-time data is harmonized with recent and not-so-recent historical data to create insights. This is a crucial reason why organizations are capturing and utilizing data at an unprecedented rate. In 2025, 495.89 terabytes of data will be generated daily (181 zettabytes annually), a tenfold increase from a decade ago.[1] It is safe to assume this number will continue to increase at an accelerated rate in the coming years.

However, it's not just the sheer volume of data that is transformative; it is how that data is put to use. Advanced open data platforms enable organizations to extract deeper insights and drive impactful outcomes, leveraging the ability to access all data through a single entry point with a shared and open metadata layer. Fueled by this data access, discriminative AI has long been delivering operational efficiency to enterprises. For example, visual recognition for manufacturing faults has decreased product defects considerably over the decades.

---

[1] Petroc Taylor, "Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2023, with forecasts from 2024 to 2028," Statista, 21 November 2024

FIGURE 1: DATA DRIVES THE AI REVOLUTION
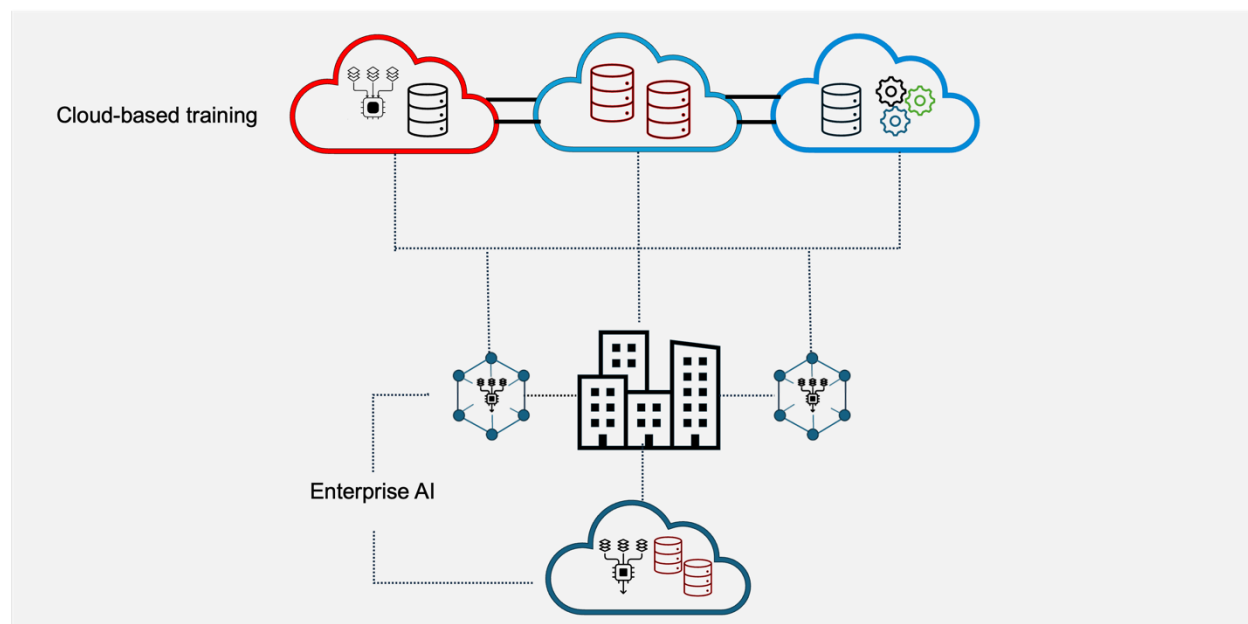
Data is driving the AI revolution across the enterprise.
Source: Moor Insights & Strategy

Now, generative AI is thought to deliver even greater operational efficiency as it drives new levels of automation across business processes and customer interactions. Improving a business process can mean anything from streamlining supply-chain inefficiencies to lowering drug-discovery costs for cures to the most vexing diseases.

THE ADVANCED THREAT LANDSCAPE

As the data estate becomes more diverse and distributed, the increased value of enterprise data has led to a threat landscape that is more perilous than ever for CISOs and their teams. Not only is data being generated at an increasing rate, but it is also being shared around the globe within organizations, between partners, and with customers.

As expected, this has led to increased cyberattacks that show no sign of slowing down. These numbers from an IBM-sponsored report from industry security firm Ponemon[2] provide a more detailed reflection of the current security threat landscape:

- The average security breach in the U.S. cost the affected company $4.88 million in 2024, a 10% year-over-year increase.

---

[2] IBM, Cost of a Data Breach Report 2024

- It takes an enterprise IT organization 292 days on average to discover and contain a breach associated with stolen credentials.
- Over half of breached companies in 2024 faced IT security staffing shortages, a roughly 26% increase from 2023.

Other studies paint an even more dire picture:

- Cybercrime will cost the world economy $10.5 trillion annually by the end of 2025, with crime involving blockchain or cryptocurrencies (cryptocrime) accounting for $30 billion.[3]
- Phishing is the primary attack vector for malware and ransomware, representing approximately 80% of reported cybercrimes.[4] While this number is stunning, the use of AI to hyper-personalize these attacks to mimic trusted individuals will make phishing even more dangerous in the future.
- While we think most cyberattacks are related to the more conventional "hacker breaches network through phishing" vector, the supply chain has proven to be a real vulnerability as well. In 2024, roughly 15% of data breaches were tied to supply chain vulnerabilities (including third-party software).[5]

Just as AI has shown great promise for driving efficiencies across the enterprise, it also shows great potential for disruption through multimodal attacks, automated attack chains, and AI-driven ransomware.

Finally, while quantum computing is still nascent, the need for PQC should be top of mind for all enterprise IT and security leaders. Harvest now, decrypt later (HNDL) attacks are real, with hackers stealing encrypted data today and waiting to use Shor's algorithm and other methods to break the encryption that has protected data for decades.

While these facts provide a grim view of the security landscape, enterprises can take intelligent measures to protect their operating environments and the data generated, used, and stored across their data estates. The foundation for any holistic security strategy begins in silicon and creates a chain of trust up through the operating stack.

---

[3] Cybercrime Magazine, "90+ Cyber Crime Statistics 2024: Cost, Industries & Trends," 2 February 2024
[4] Nivedita James Palatty, "90+ Cyber Crime Statistics 2025: Cost, Industries & Trends," Astra, 6 February 2025
[5] Gary Smith, "100+ Data Breach Statistics and Trends for 2025," StationX, 10 December 2024

IBM LinuxONE — Uncompromising Security

As mentioned previously, the security challenge with data in the enterprise is not simply that it continues to grow (That reality actually presents very little challenge). Rather, the challenge is the utility of data and its increased *importance* to the modern enterprise. With this in mind, Moor Insights & Strategy (MI&S) sees IBM's LinuxONE 5 as arguably the market's most secure server hardware platform.

With LinuxONE 5, security begins in silicon with the Telum II processor. As LinuxONE is powered, its secure boot feature uses a specialized processor in dedicated silicon, employing quantum-safe algorithms to ensure that the server and its firmware are booted into a secure and pristine state. The vertically integrated stack enables applications running in this modern cloud environment to benefit from the natively delivered security in silicon and hardware.

Contrast LinuxONE 5 with legacy x86 servers that populate many datacenter racks. Even the most secure servers require a handshake between the hardware platform, its various sources of firmware, and a third-party operating system — be it a Windows Server or Linux distribution. This multivendor handshake introduces inherent vulnerability. Further, the workloads, applications, drivers, and APIs running in these operating environments introduce vulnerabilities not encountered in a LinuxONE 5 operating environment, where all components are curated.

Once LinuxONE 5 boots into this secure operating environment, IBM's Confidential Computing is the key enabler of the secure cloud or third-party-operated environment. In this environment, LinuxONE deploys a given pod of containers based on an encrypted artifact, and the system itself protects the memory that houses these containers and their associated data. Because this pervasive encryption and key management is based in hardware, software applications suffer little or no performance tax while gaining the benefits of operational simplification — specifically because running them does not depend on environment configurations or applied policies. This extends the concept of secure boot to the solution layer.

Quantum Resistance and Algorithmic Support

Many see quantum computing as a future state of the computing industry. However, while it may be another five years before quantum computing is reliable enough for mainstream adoption, its effect is already being felt as MI&S sees more frequent HNDL attacks. Worse, there is a surprising lack of awareness about these attacks among enterprise IT organizations.

The algorithms used to support encryption-key management are essential for PQC. The [National Institute of Standards and Technology (NIST)](#) has identified four algorithms it considers to be quantum-resistant, three of which were developed by IBM. It is not hyperbolic to say that IBM is at the forefront of PQC security, and it has brought its learnings into the LinuxONE design. In particular, LinuxONE 5 stands out as a secure compute platform for the modern enterprise thanks to its vertically integrated security. Its Crypto Express 8S security module is designed to support NIST's most stringent post-quantum requirements (FIPS 140-3 Level 3) and [CNSA (Commercial National Security Algorithm Suite) 2.0](#).

Of course, it is easy for vendors to make claims around security. MI&S believes a better indicator of a platform's real-world value can be seen in which organizations trust their most critical workloads and environments to that platform. In LinuxONE's case, the underlying architecture handles 70% of the world's financial transactions. These organizations (in banking, healthcare, insurance, and so on) are some of the largest data generators anywhere, operating under some of the tightest regulations and scrutiny.

PERFORMANCE FOR MODERN WORKLOADS

It's worth exploring how IBM LInuxONE 5 can deliver a fully performant operating environment for modern workloads, such as AI and analytics, with every transaction and action secured. Even in the face of a growing threat landscape, performing transactions faster, responding to customer queries first, and delivering value immediately remain critical. For many business leaders who speak with MI&S, time-to-value has become the key metric by which they judge themselves and the IT organizations that support their business directives. In fact, time-to-value competes with maintaining secure operations as the most critical imperative.

Interestingly, LinuxONE 5's performance begins in the same place as its security capabilities: in silicon. The Telum II processor represents a significant leap in performance over its predecessor.[6] Even at a 5.5 GHz clock speed, Telum II consumes 20% less physical space and 15% less power while delivering a 20% increase in performance per socket.

---

[6] Christian Jacobi and Elpida Tzortzatos, "[New Telum II Processor and IBM Spyre Accelerator: Expanding AI on IBM Z and IBM LinuxONE](#)," IBM, 26 August 2024

WHEN A CORE IS NOT A CORE

IBM's Telum II processor and cores differ significantly from the x86 CPUs powering legacy servers from Dell, HPE, and others. What IBM refers to as Integrated Facility for Linux (IFL) processors power LinuxONE 5. These are high-performing, power-efficient processor cores with Linux-optimized microcode that cannot be equated with legacy distributed computing cores. Internal IBM testing demonstrates that a single IBM LinuxONE Emperor 5 Max 136 can do the work of up to 2,944 cores in an x86 solution.[7]

These higher-performing processors don't just improve performance; they are designed to do more. Whereas MI&S still sees enterprise organizations running x86 servers at a 30% to 40% utilization rate, enterprise-grade servers such as LinuxONE are designed to run at utilization rates up to 80%. Because of this, enterprise organizations can achieve significant consolidation ratios when migrating from legacy x86 to a cloud operating environment such as Red Hat OpenShift on LinuxONE 5.

---

[7] To convey more specifics, IBM requested that MI&S include this documentation of the test:

CLAIM: IBM internal tests simulating a complete IT solution running containerized WebSphere Liberty and EDB Postgres workloads show that a single IBM LinuxONE Emperor 5 Max 136 can do the work of up to 2,944 cores of the compared x86 solution.

DISCLAIMER: IBM internal performance tests for the core consolidation study targeted a comparison of the following servers. IBM Machine Type 9175 MAX 136 system consisting of three CPC drawers containing 136 configurable processor units and six I/O drawers to support both network and external storage. The x86 solution used a commercially available enterprise server with two 5th Gen Intel Xeon Platinum 8592+ processors, 64 cores per CPU. Both solutions had access to the same storage. The workloads consisted of a containerized online transaction processing (OLTP) WebSphere Liberty v25 application running on Red Hat OpenShift Container Platform (OCP) v4.17, and an EDB Postgres for Kubernetes v1.25 on the same OCP cluster simulating core online banking functions. Both solutions used Red Hat Enterprise Linux v9.5 and KVM. Results may vary.

FIGURE 2: CONSOLIDATION WITH LINUXONE 5
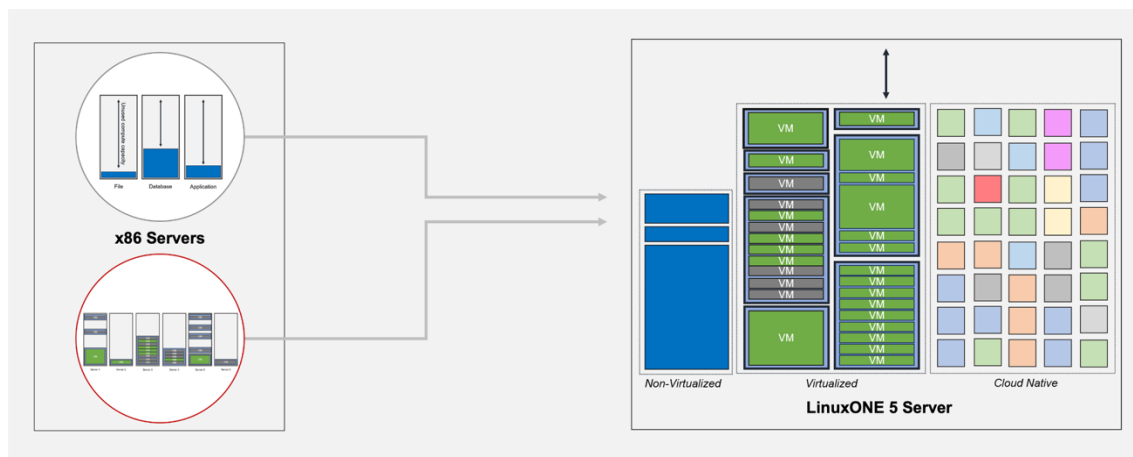
LinuxONE 5 enables considerable consolidation ratios on a hybrid cloud platform.
Source: Moor Insights & Strategy

Performance is not simply about the CPU, but also about how quickly data can move from memory to the CPU or an accelerator and how quickly that data can be processed. With Telum II, IBM integrated a data processing unit (DPU) directly into the compute complex, extracting even greater performance for data movement across the system.

IBM LINUXONE AND AI

The second generation of IBM's dedicated on-chip AI accelerator drives faster inferencing by delivering four times as much processing power, combined with greater quantization and matrix operations. When looking across the entire LinuxONE system, in-drawer intelligent routing enables the availability of all accelerators to any process or inference call, regardless of its origination. This unique capability happens only through this vertical integration, where the OS can fully exploit its underlying resources. And, with IBM LinuxONE, this acceleration can also take advantage of Confidential Computing to further ensure the privacy of operations.

As with security, every server platform vendor will make claims about performance, especially in this era of analytics and AI. However, also like security, the true measure of performance in the real world is to look at who uses a product — and for what. IBM Z, which shares an underlying architecture with IBM LinuxONE, has been deployed for over 60 years for workloads that require the fastest response times to generate, move, and use data across the globe 24 hours a day. Today, LinuxONE is powering these highly demanding workloads for banks, government agencies, and university research departments. MI&S sees this as perhaps the greatest endorsement of IBM LinuxONE.

TELUM II AND SPYRE — THE FUTURE OF AI

LinuxONE can already deliver highly performant and cost-efficient inferencing of machine learning, deep learning, and encoder large language models (LLMs) with its on-chip AI accelerator inside Telum II. Beyond this, the expected release of the Spyre accelerator chip in Q4 2025 will bring a fit-for-purpose AI accelerator to drive LLMs on IBM LinuxONE.

The generative AI supported by the Spyre accelerator is designed to handle encoder-decoder AI models for applications such as chatbots, document summarization, and agentic AI that assists with critical database transactions. Both structured and unstructured data is supported to enable a wide array of enterprise-level functions. The Spyre accelerator chips are packaged via PCIe adapter cards, with clusters of eight cards creating a robust, scalable environment for enterprise AI. Up to six clusters (48 total Spyre cards) can be configured on a LinuxONE 5 system. The cards are designed to be power-efficient, targeting 75W per adapter.

What MI&S finds compelling about Telum II and Spyre is the ability to deliver acceleration for workloads that benefit from using multiple AI models, such as money laundering or fraud detection, where deep learning combined with generative AI can deliver more accurate results faster. Overall, the addition of generative AI also enables summarization, recommendations, and agentic AI to run applications like contract generation, financial document processing, and image screening with automated agent functions. The ability to drive this multiple-model support from a single accelerator with minimal power consumption should make this a valuable tool for financial services organizations.

BUILDING THE CLOUD ON-PREMISES

The cloud is the great abstraction layer. Perhaps more precisely, *cloud-native* is the great abstraction layer, enabling enterprise IT organizations to deploy and move applications, workloads, and data across platforms regardless of underlying server types or architectures.

The public cloud has long removed architectural affinity from enterprise organizations, which focus more on container platforms and development environments such as Vert.x, Node.js, and Golang. The focus has shifted from cores and memory capacity to the number of instances and environments optimized for the cloud-native platform.

This is precisely what IBM has designed in LinuxONE 5: a hybrid cloud machine that seamlessly connects the enterprise to the public cloud, delivering maximum performance and consolidation most securely, significantly reducing operational costs. LinuxONE is the cloud platform that delivers the efficiency and point-and-click simplicity of the cloud without the runaway costs that usually accompany datacentric workloads.

While many IT teams have seen their x86, Unix, and mainframe systems siloed, LinuxONE bridges that gap. It supports everything from cloud-native, datacentric applications to tooling like Ansible for managing infrastructure to modern DevOps practices with CI/CD pipelines using tools like Jenkins and Git. LinuxONE is not some add-on platform used to support only the most sensitive data with the highest performance requirements; it is an integrated data platform that can support both legacy and modern, cloud-native workloads.

FULLY REALIZING THE PROMISE OF THE CLOUD

In addition to bringing the simplicity and agility that the cloud offers, LinuxONE can deliver on the often unfulfilled promise of cost savings that originally drove a lot of cloud adoption. Indeed, deploying LinuxONE 5 can result in significant TCO savings.[8] This may seem a bit incredible to x86 consumers who acquire a single server for considerably less capital outlay. However, when considering the full cost of deployment — from capital expenditure for hardware to software licensing to management, power, and other indirect costs — that initial capital expense is quickly accounted for. This is especially true in datacentric environments where traditional database licensing costs, along with private cloud costs based on physical core counts, represent the vast bulk of the IT budget.

While each organization is unique, MI&S strongly recommends that enterprise IT organizations perform a complete TCO analysis when considering LinuxONE versus x86-based server platforms. Consider licensing, power costs, additional accelerators, and the utilization rates of each platform.

CALL TO ACTION

Data drives the modern business, and the modern business drives the world. In this era of AI, data is increasingly vital to feed the AI models and workloads that deliver outcomes. This has led to an increased focus on capturing and using data everywhere

---

[8] IBM, TCO and CO2e Calculator for IBM LinuxONE and Linux on IBM Z

at all times and ensuring that the unique compute requirements of each application and workload are being met.

The evolving distribution of data and workloads across the public and private technology estate has led to a threat landscape that has continued to increase in complexity at an alarming rate. This trend is only exacerbated by bad actors' use of AI and (imminently) quantum computing.

Despite the threat landscape, many IT organizations that speak with MI&S believe that there must be some tradeoff between security and the ability to utilize data to its fullest potential in the most performant way. In other words, there is a concern that employing hardened security will stifle performance, agility, and openness — besides being cost-prohibitive.

MI&S has found that IBM's LinuxONE platform delivers the performance required by enterprise IT organizations without sacrificing security. For the modern enterprise deploying containerized, cloud-native workloads spanning the on-prem data estate and the public cloud, LinuxONE 5 is a compelling, cloud-architected platform. Further, the introduction of LinuxONE 5 delivers more features that drive end-to-end, data-heavy workloads — from Telum II to integrated DPUs and built-in AI acceleration, frameworks, and tools. In short, it delivers high performance and security to the enterprise in a modern cloud-like experience.

Because of these factors, MI&S suggests that enterprise IT organizations modernizing their environments and operationalizing AI consider IBM LinuxONE 5. Begin with a thorough performance evaluation and follow with a TCO comparison that includes all cost factors. We believe LinuxONE 5 will compare favorably.

For more information, please visit: www.ibm.com/linuxone.

IMPORTANT INFORMATION ABOUT THIS PAPER

CONTRIBUTOR
Matt Kimball, Vice President and Principal Analyst, Datacenter Compute and Storage

PUBLISHER
Patrick Moorhead, CEO, Founder, and Chief Analyst at Moor Insights & Strategy

INQUIRIES
Contact us if you would like to discuss this report, and Moor Insights & Strategy will respond promptly.

CITATIONS
Accredited press and analysts can cite this paper, but it must be cited in context, displaying the author's name, title, and "Moor Insights & Strategy." Non-press and non-analysts must receive prior written permission from Moor Insights & Strategy for any citations.

LICENSING
This document, including any supporting materials, is owned by Moor Insights & Strategy. This publication may not be reproduced, distributed, or shared in any form without Moor Insights & Strategy's prior written permission.

DISCLOSURES
IBM commissioned this paper. Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

DISCLAIMER
The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of this document's publication date. Please remember that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.