ACHIEVING SUSTAINABILITY, SECURITY AND SCALABILITY WITH IBM LINUXONE

SPONSORED BY IBM

EXECUTIVE SUMMARY

In this research paper, Moor Insights & Strategy (MI&S) explores the challenges enterprise organizations face as they try to achieve the seemingly contradictory goals of digitally transforming business operations while reducing their carbon footprint, security vulnerabilities, and, ultimately, costs. This paper will demonstrate how the release of IBM LinuxONE Emperor 4 helps such organizations find success on their digital transformation journey.

SITUATION ANALYSIS

Many established businesses across industries need to transform. Business processes and operations that led to market leadership decades ago are outdated relative to today's digitally native competitors that can deliver products and services faster and respond to market needs with the agility and precision required for survival in the modern economy.

This market dynamic has led to seemingly every enterprise engaging in digital transformation projects. While each company is unique in its approach to transforming the business, the goals are universal – respond to the market's needs with the speed of these digital upstarts while reducing costs.

In parallel and ostensibly in contrast with these transformation projects is the rise of environmental, social, and governance (ESG) mandates that modern businesses have rightfully prioritized alongside growth and profitability.

Virtually every enterprise CIO with whom MI&S engages is in the throes of developing a plan with strategies to support these different objectives. Many are struggling to answer one question in particular, "How does an IT organization deploy underlying infrastructure that can support the business's diverse (and constantly shifting) needs and reduce its costs, power consumption, and datacenter space?"

DIGITAL TRANSFORMATION – SURVIVAL OF THE GREENEST?

"It is not the strongest of the species that survives, nor the most intelligent; it is the most adaptable to change." Often attributed to British naturalist Charles Darwin, this sentiment is perhaps more relevant than ever for businesses.

As mentioned, the specific goals of an organization's digital transformation project vary. However, the guiding principle – the north star, if you will – is survival. They must drive down costs and use technology to drive business efficiencies. The burden of achieving this goal rests squarely on CIOs and IT executives, who must rethink technology deployments and consumption.

While considering this survival, IT plays a significant role in a company's ESG goals. As it relates to the environment, the impact of IT is substantial. Consider these numbers uncovered in an IBM Institute for Business Value report on sustainability:[1]

- Collectively, datacenters account for 200-250 terawatt-hours (TWh) of energy – roughly 1% of all global energy consumption and 0.3% of carbon emissions.
- Further, the datacenter market is anticipated to grow by 30% through 2027.
- The dependence on workloads such as data analytics, artificial intelligence/machine learning (AI/ML), and security will further drive energy consumption due to the use of additional servers to run specialized software.
- Internet users doubled between 2010 and 2021, and internet traffic increased 15-fold.

The above demonstrates two critical roles IT plays in addressing climate change. The strategic IT decisions made around infrastructure can:

1. Have a tangible impact on the speed at which businesses can create and deliver products and services that have a measurable effect on climate; and
2. Impact IT's energy usage and carbon footprint and how much it contributes to good climate stewardship.

[1] "IT Sustainability beyond the Data Center" – IBM Institute for Business Value
https://www.ibm.com/thought-leadership/institute-business-value/report/it-sustainability

These are not hyperbolic proclamations designed to elicit an emotional response. Instead, they are realistic perspectives that have resonated so strongly that 86% of companies have a sustainability strategy.[2]

STEWARDSHIP INTERSECTS WITH PRAGMATISM

With so much focus on being good climate stewards, businesses are ultimately driven by the bottom line – making money. And sustainability goals have a strong alignment with the bottom line. To put this 250TWh estimate of datacenter energy consumption into greater perspective, this is the equivalent of the total power consumed by Australia per year. The rapidly rising cost of energy is making this more important than ever.

These numbers represent the cost associated with powering servers and IT infrastructure. This equation does not capture cooling, floor space, and other expenses. Considering the amount of idle time most x86 processors experience, this number represents a lot of datacenter inefficiency. Put more simply, wasted budget.

From a sustainability and resulting TCO perspective, the ideal server environment would consume less power per square foot and have far less idle time because of higher utilization.

SUSTAINABILITY SHOULD NOT BE LIMITING

It is reductive to think of corporate-led, IT-driven sustainability initiatives as constraints on digital transformation. As an IT organization considers how to modernize the application frameworks, tools, and associated development methodologies used to drive transformation, "thinking green" through the application development process can instead be somewhat liberating.

To support the digitized business, IT organizations transform operations to become more cloud-like in agility and function. This leads to a shifting perspective on infrastructure. The transformation further abstracts compute platforms and CPU architectures from workloads through these cloud and cloud-native technologies. Rather than focusing on core counts, instructions per clock (IPC), and artificial benchmarks, IT organizations look for consistency in (high) performance, reliability, security, and cost of ownership. Further, IT organizations must look at power consumption in the aggregate vs at the component or even rack level.

[2] Sustainability as a transformation catalyst – IBM Institute for Business Value
https://www.ibm.com/thought-leadership/institute-business-value/report/sustainability-transformation

With a sustainability mindset, cloud and cloud-native technologies become ever crucial for application performance efficiency. Mobility and the ability to shift to a hybrid-cloud model to support the needs of the business become necessary.

One of the keys to achieving power savings in IT infrastructure is scalability – the ability to use less (but more elastic) server technology and infrastructure to support the most demanding workloads, but also scale up for workload growth with minimal increases in energy usage.

Finally, organizations must consider securing their most important asset: data. How can a business run faster and more efficiently while simultaneously ensuring data privacy – whether in use, in flight, or at rest?

MI&S believes there is a server platform suited to this – IBM LinuxONE, an enterprise server optimized for Linux workloads.

LINUXONE EMPEROR 4 – SUSTAINABILITY AND TCO GAINS

LinuxONE can be an ideal compute platform for organizations looking to drive toward sustainability goals while simultaneously fueling transformation efforts. With LinuxONE Emperor 4, IBM has further extended its performance, efficiency, security, and reliability capabilities.

At the heart of LinuxONE Emperor 4 is Telum, IBM's newest processor developed for enterprise workloads. IBM designed Telum (and LinuxONE) to run at the highest utilization levels without sacrificing performance, reliability, or consistency of response time.

For x86 server IT administrators looking to compare server platforms, the default criteria tend to be around CPU cores and synthetic benchmarks such as SPEC.org integer performance testing. These comparisons are typically designed to measure the performance of a single workload. IBM LinuxONE is designed to run hundreds of workloads in a single system, so they don't measure the differences between Telum and x86; therefore, this paper will avoid the comparison. However, if one is looking for some analogous reference, IBM internal tests show that when running WebSphere and DB2 workloads, IBM LinuxONE Emperor 4 requires 16 times fewer cores than the compared x86 servers. If you scale this up to a complete IT solution this means when running this

workload, the IBM LinuxONE Emperor 4 Max 125 would be doing the work of about 2000 cores of the compared x86 servers.[3]

Instead, it's better to understand that from the compute complex to cache to memory and I/O, Telum is designed for the most demanding environments. Telum is also the CPU utilized in IBM's zSystems servers, which power the most demanding workloads in the enterprise.

Several on-chip accelerators are built into Telum and are designed to deliver real-world performance improvements, such as encryption acceleration, data compression, and AI inferencing. By developing dedicated accelerators, security, data management, and inferencing performance are greatly enhanced without taxing the core CPU complex. Contrast this with x86 CPUs, where additional CPU resources are needed for specialized functions such as encryption.

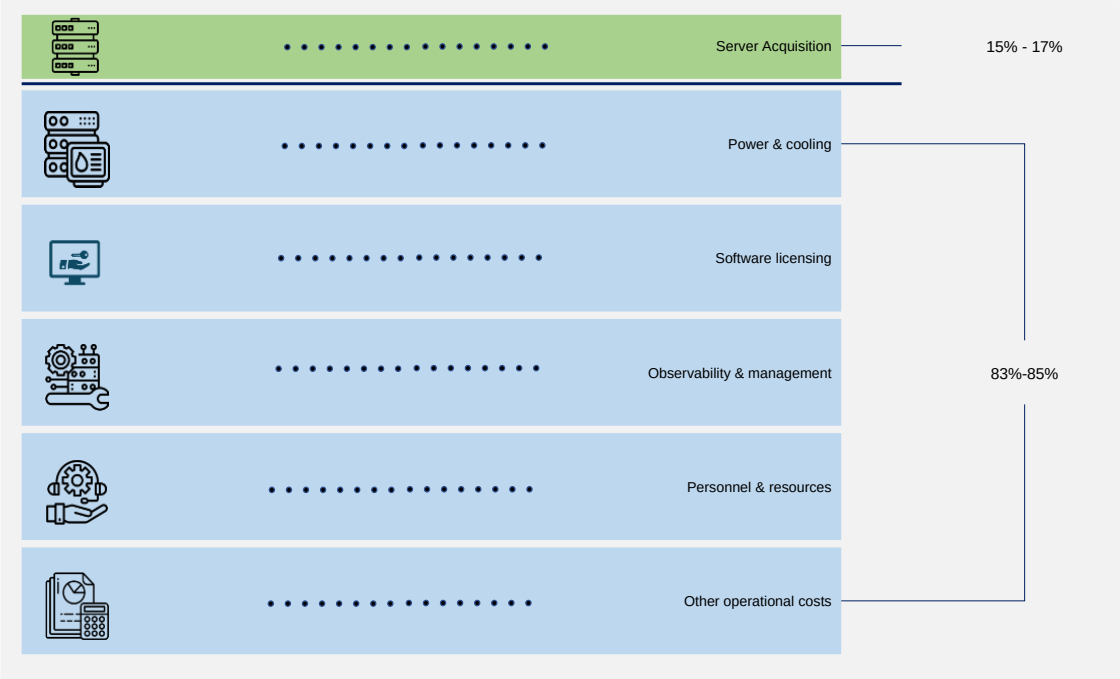
Sustainability, scalability, and security extend from Telum into system design, where IBM LinuxONE pairs Telum's high-performance cores with up to 40TB of memory, dedicated I/O processors, and the accelerators mentioned above.

[3] IBM internal tests show that when running WebSphere and DB2 workloads, IBM LinuxONE Emperor 4 requires 16 times fewer cores than the compared x86 servers. If you scale this up to a complete IT solution this means when running this workload, the IBM LinuxONE Emperor 4 Max 125 would be doing the work of about 2000 cores of the compared x86 servers.

Disclaimer
This is an IBM internal study designed to replicate a typical IBM customer workload usage in the marketplace. Results may vary. The core consolidation study targeted comparison of the following IBM LinuxONE and x86 servers: IBM LinuxONE Emperor 4 Max 125 system consists of three CPC drawers containing 125 configurable processor units (IFLs or zIIPs) and two I/O drawers to support both network and external storage. Lenovo ThinkSystem SR650 (2U) with two 2nd Gen Intel® Xeon® Platinum processors 2.1 GHz, 16 cores per CPU. Both the x86-based and LinuxONE solutions had access to the same storage array. The workloads consisted of a transactional application running on WebSphere Application Server and IBM DB2 simulating core online banking functions. The actual test results were extrapolated to the stated above x86 servers using IDC QPI metrics and IBM sizing methodology using the following assumptions on a typical IT environment of a banking client using x86 servers. The production IT environment has 16 x86 servers running at 50% average utilization. There are 48 x86 servers in the non-production IT environments: development (4 environments with 2 servers each, 8 servers total), development test environment (4 servers), system integration test environment (8 servers), performance test environment (16 servers), user acceptance test environment (4 servers), production fix test environment (8 servers). A typical average CPU utilization is 7% across all non-production environments. An equivalent LinuxONE Emperor 4 solution requires a single Max 125 server running at 85% average utilization across all IT environments separated using LPAR technology.

The result is scalable infrastructure that enables vertical and horizontal scale in a single platform. And this scalability has further enabled the ability to allocate capacity automatically, permanently, or temporarily based on priority – the very essence of the flexible, just-in-time concept that drives greater efficiencies.

FIGURE 1: THE ELEMENTS OF TCO



Hardware acquisition costs are a fraction of overall TCO
Source: Moor Insights & Strategy

WHAT DRIVES TCO?

While TCO has been a somewhat nebulous term contorted to fit marketing narratives, TCO itself is accurate. And significantly lowering costs associated with driving business services will be critical to the survival of IT. While many TCO has many elements, three of the larger contributors to this model center around energy consumption, manageability, and software licensing.

When considering LinuxONE as a pillar of digital transformation, MI&S believes organizations can achieve actual TCO savings while increasing business responsiveness to this everchanging economy, with power savings being an increasingly large contributor to the TCO equation. Part of what drives this belief is research conducted by IBM showing that an IBM LinuxONE Emperor 4 can reduce the

CO2e footprint by approximately 75% each year versus compared x86 servers running the same Linux workloads under similar conditions.[4]

While significant, energy savings is one element of the TCO model, another contributor to TCO is the cost of managing the servers. And this is a second area where IBM LinuxONE shines relative to its x86 counterparts.

From a management perspective – monitoring, observability, and management – the value of IBM LinuxONE is clear since it is a single centralized system rather than a distributed system. The consolidation of many server platforms alone makes the management value of IBM LinuxONE compelling. And the architecting of so many reliability capabilities only adds to IBM LinuxONE's differentiation.

IBM LinuxONE – Sustainability With Scalability And Security

Along with LinuxONE's ability to help IT organizations contribute to organization-wide sustainability goals, scalability and security are two additional areas of innovation worthy of spotlighting.

This discussion must start with recognizing IBM's longevity in a Darwinistic enterprise IT services market. While x86 dominates the datacenter and new entrants like the Arm-based silicon ecosystem begin to emerge, many enterprise IT organizations continue to run their most mission-critical workloads on IBM technology.

The reason for IBM's longevity is simple – reliability, security, and consistency of performance. Organizations know that these critical workloads running on IBM

[4] An IBM LinuxONE Emperor4 can reduce the CO2e footprint by approximately 75% each year versus compared x86 servers running the same Linux workloads under similar conditions

Disclaimer for Claim 2 (a-d)
IBM Machine Type 3931 [AF1] Max 125 model consists of three CPC drawers containing 125 configurable cores (CPs, zIIPs, or IFLs) and two I/O drawers to support both network and external storage. 39 Cascade Lake and Ice Lake x86 servers configured to provide same performance. Annual energy required for server operations was 143,962 KW for IBM Machine Type 3931 [AF2] with 125 cores and 566,448 KW for 39 x86 servers with 2072 cores, so the CO2e footprint would be 62.5 metric tons for IBM Machine Type 3931 and 245 metric tons for x86 servers. CO2e and other equivalencies that are based on the EPA GHG calculator (https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator) use U.S. National weighted averages. Results may vary based on client-specific usage and location.

technology will consistently run as expected with no interruption and no real risk of exploitation.

Further, because of the design of IBM LinuxONE, customers can achieve workload consolidation never thought of in the x86 market, on CPU and compute platforms designed to run at the highest utilization levels. While there is a wide range of estimates around the average utilization rates of x86 servers, MI&S consistently sees utilization below 30%, which means that 70% of the computational resources sit unused.

Contrast this number with the Telum processor used in the new IBM LinuxONE Emperor. This processor (and its predecessors) is designed, developed, and tested to consistently sustain utilization rates at 80% and above[5]. The very definition of server utilization efficiency. This design point enables IT organizations deploying virtualized environments on LinuxONE to achieve the levels of consolidation promised but never realized in x86 environments. For a deeper look at how LinuxONE helps IT organizations achieve greater levels of datacenter consolidation, read this previously published MI&S research report.

SCALABLE AND ELASTIC—REQUIREMENTS FOR THE DIGITAL ENTERPRISE

The successfully transformed digital enterprise is cloud-native and intelligence-driven. Applications made up of thousands of services access data from a seemingly infinite number of data sources (traditional SQL databases such as IBM's Db2 or Oracle, as well as NoSQL and unstructured data platforms) and synthesize it into intelligence that drives the business. These cloud-native workloads are often scale-out due to the mobile, distributed, and lightweight nature of container-based applications, but some of these scale-out workloads will rapidly grow and inherit some attributes of scale-up workloads.

This data environment evolution requires an underlying evolution of infrastructure, implementing scale-out servers to support the scale-up environments that support legacy virtualized and traditional database infrastructure. This evolution, however, comes at a floorspace, power, and management cost that can be significant.

[5] IBM LinuxONE delivers performance and cost benefits – Moor Insights & Strategy, white paper sponsored by IBM.

IBM's LinuxONE platform is unique in that it enables IT organizations to support their traditional scale-up and cloud-native scale-out environments simultaneously. And as cloud-native workloads become more "bursty" in nature, IBM LinuxONE can commit and decommit compute resources automatically, with no noticeable latency. MI&S has not seen any other server platform that can simultaneously support these environments with divergent performance characteristics and underlying platform requirements.

Further, given its support for major Linux distributions (Red Hat, SUSE, Ubuntu) and Red Hat's OpenShift container management platform, LinuxONE is an ideal hybrid-cloud platform for mission-critical workloads.

DESIGNED FOR COMPREHENSIVE SECURITY

And finally, through many IBM innovations, LinuxONE offers a wide range of advanced security capabilities. Pervasive encryption aided by hardware acceleration makes locking down data in flight automatic and undetectable from a performance perspective. Confidential computing isolates workloads at scale and can encrypt physical memory.

An excellent example of the forward-thinking nature of IBM can be found in its quantum-safe cryptography technologies newly available in IBM LinuxONE Emperor 4. Traditional public-key cryptography relies upon mathematical problems that are difficult to solve on classical computers. However, popular cryptographic schemes, such as RSA and ECC, can be easily broken by a sufficiently large quantum computer.

While one could argue quantum computing is still at an early stage of evolution, it is a technology coming to the market and will be here quickly. Much like AI, there is a lot of work and enablement being done in quantum that will make it appear faster than anticipated. In support of this, IBM LinuxONE Emperor 4 ushers in a suite of algorithms that are resistant to attacks by both classical and quantum computers.

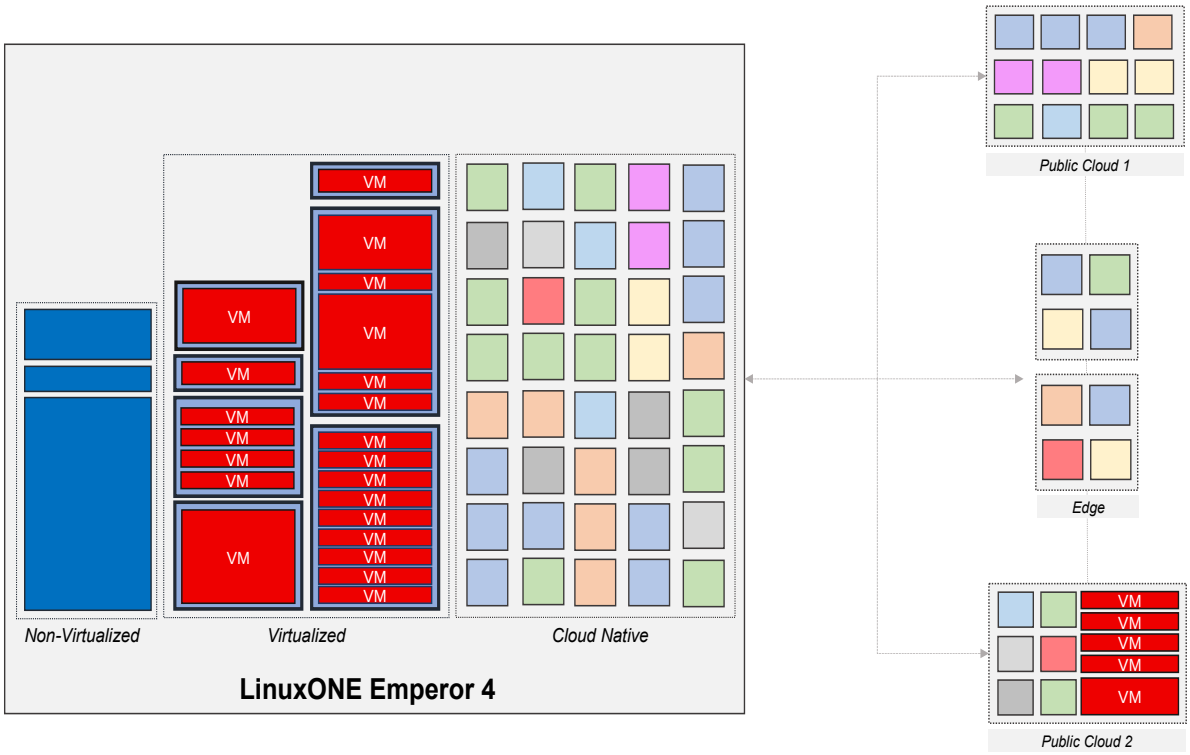
Because of this, MI&S believes IBM has a significant opportunity to expand its market footprint with the release of IBM LinuxONE Emperor 4.

THE MODERN ENTERPRISE THRIVES WITH LINUXONE

IT organizations undertaking digital transformation aim to adopt cloud-like practices to best serve the business. This means that applications and services are delivered on demand while IT professionals take a more consultive role within business units.

For IT organizations transforming with no real consideration for transforming infrastructure, this means acquiring, deploying, powering, and managing new specialized platforms to support new and emerging data management environments and applications. In other words, the very goals of transformation – do more with less – are being defeated.

FIGURE 2: POWERING THE MODERN ENTERPRISE WITH IBM LINUXONE



Source: Moor Insights & Strategy

LinuxONE Emperor 4 can drive a hybrid multi-cloud environment to the business with a fraction of the platforms, enabling the goal of expanding services while reducing complexity, cost, resources, and carbon footprint. Put more plainly, choosing LinuxONE can be good for the business and the planet.

CALL TO ACTION

The debate around climate change has ended, and its effects are felt more now than ever. The business world mirrors this global sense of urgency, with corporations of all

sizes, industries, and geographies making sustainability a top-tier business goal to be tracked.

Running parallel to sustainability is the need for organizations to transform. Modernizing and automating functions enables faster time to value and market and faster response to customer needs and competitive threats.

While achieving these sustainability and transformation goals, IT organizations are being asked to do more with less budget and fewer resources.

Some may look at this dynamic and see a litany of contradictions. As technology becomes the driving force of the business, more infrastructure, software, and resources are needed to deliver such capabilities.

In a view that runs counter to this narrative, MI&S believes companies that drive sustainability at the lowest levels of business planning can achieve cost savings and drive competitive differentiation while achieving the altruistic goals of being good stewards of this planet

IBM designed LinuxONE as a platform to enable the flexibility and power of the cloud while significantly reducing datacenter power and floorspace requirements. With IBM LinuxONE Emperor 4, the cost savings, operational agility, and rich security differentiate this platform more than ever.

Seeing LinuxONE as uniquely qualified to enable IT organizations to simultaneously achieve digital transformation and sustainability goals, MI&S strongly recommends enterprise IT organizations consider LinuxONE as a hybrid cloud platform enabling the business to compete in this digital economy.

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

IMPORTANT INFORMATION ABOUT THIS PAPER

CONTRIBUTOR

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

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

This paper can be cited by accredited press and analysts but must be mentioned in the 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

DISCLOSURES

IBM commissioned this paper. Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No firm employees hold 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 in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The views expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators, not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on the future, risks and uncertainties could cause actual results to differ materially. You are cautioned not to rely on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind 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.