IDC ANALYZE THE FUTURE

IDC VENDOR SPOTLIGHT

Sponsored by IBM

In today's heterogeneous datacenter, data streams of varying value flow from and to purpose-built platforms. These platforms are optimized for distinct workloads, both on-premises and in the cloud. High-value data requires a high-end platform with a mature operating system, such as IBM AIX.

The Importance of the Operating System When Modernizing on Hybrid Multicloud

August 2020

Written by: Peter Rutten, Research Director, Infrastructure Systems, Platforms, and Technologies

The Changing Datacenter

Trends That Are Driving Datacenter Transformation

The global digital transformation (DX) wave is causing a reorientation of IT staff and IT budget toward a digital, customer-facing, revenue-generating, and central function of the business. This holds true for any business, small or large, and regardless of industry. A significant portion of DX is about shifting resources toward the public cloud to improve flexibility, scalability, and programmability while running IT as an operational expense (opex) model. This is an important element of a business' ability to respond to market fluctuations, consumer trends, and demand peaks and valleys. IDC has seen that workloads that *can* go to the public cloud *will* go to the public cloud.

However, the massive amounts of data that continue to be managed in the world's enterprise datacenters -- in siloed systems, o private clouds, or in hybridclouds -- are less often talked about because of the spectacular growth that cloud service providers have experienced. For example, in a recent IDC survey, large business (i.e. 1,000+employees) that plan to adopt SAP in the next 12 months said that they intend to run 62% of their databases, 69% of their business applications, and 65% of their business analytics applications on-premises (source: IDC's Special Study Infrastructure Adoption Trends for SAP and SAPHANA2019). The worldwide volume of this data not only is larger than that of any cloud-based data but also is significantly more valuable to the companies that generate, manage, and analyze the data for operational and competitive insights. This is the data that differentiates a business, drives a business' competitiveness, enables customer focus and, ultimately, helps steer a business through economic ups and downs.

This trove that every business "sits" on consists of customer data, transactional data, supply chain data, competitive data, operational data, parts data, technical data, and more. Most of the workloads that generate and manage this data remain on-premises for multiple reasons, even as other types of workloads that operate on less sensitive data move off-premises. Some of the reasons for remaining on-premises are proximity requirements of the data near other facilities such as plants, security, data privacy regulations, transactional performance, existing skill sets with a platform (including operating environments), and the existing communications infrastructure in a specific region, which may be subpar for low-latency on-premises-to-cloud transmissions.

AT A GLANCE

KEY TAKEAWAYS

As digital transformation (DX) continues to reorient IT strategies, businesses are shifting resources toward the public cloud. However, massive amounts of data continue to be managed in the world's enterprise datacenters. Why? Because data is not homogeneous.

Efforts to maximize capabilities have triggered a growing trend toward designing the right infrastructure for the right workload — and that extends to the operating system. Thus, operating systems, such as IBM AIX, have transformed to meet today's IT requirements and operate as a hybrid cloud.

The creed that "data is the new oil" misrepresents something critical: Data is not that homogeneous. Ten thousand pictures of birds have little value, and no one would have any qualms about putting them in a public cloud, even if these pictures enable a machine learning (ML) algorithm to learn how to differentiate between birds. Ten thousand pictures of different parts that make up a new model passenger jet have an entirely different value. Data *dissimilarity* rather than the presumed data sameness is the more relevant creed, and treating data differently according to its value is critical.

Data dissimilarity goes hand in hand with the trend toward datacenter heterogeneity. The days of homogeneous sprawl that can process every workload reasonably well are over. Purpose-built processors, purpose-built storage, purpose-built servers — the efforts to maximize capabilities have triggered a growing trend toward *the right infrastructure for the right workload*. That does not mean, however, that such purpose-built solutions, on-premises or in the cloud, should not have a common plane that allows applications and data to move freely between them and manage them in an orchestrated way; there will be no return to the old days of siloed IT infrastructure. It means that every workload is optimized on the most suitable IT infrastructure and in the most suitable deployment scenario while remaining fully open and connected.

The Right Operating System for the Right Task

The right IT infrastructure for the right workload extends to the operating system. We have come full circle in how we regard operating systems — from the lone rule of IBM z/OS to the fragmentation around Unix to the proprietary nature of Windows and ultimately to the openness of Linux. We now live in a post–operating system wars world in which these operating systems peacefully coexist and will continue to do so as they become increasingly interoperable, thanks to abstraction layers such as containers, Kubernetes, and Red Hat OpenShift and to automation tools such as the Red Hat Ansible Automation Platform.

Mainframes and IBM z/OS will not disappear. Indeed, they are thriving because they are the *right IT infrastructure for the right workload* for thousands of global companies: extremely fast and secure transaction processing without the slightest interruption. Unix-based platforms, such as IBM AIX, will not disappear; they too represent the *right IT infrastructure for the right workloads*: critical databases that require fast and secure processing with the highest possible availability and reliability. Windows will continue to be important because of its ecosystem of Windows-based clients, and Linux is the fastest-growing and, by now, largest operating system in the datacenter thanks to its community-based evolution.

Looking at operating systems as "battling" each other for market share is outdated and fairly useless; each operating system has an important role to play in the datacenter. As data streams grow and are differentiated, and as the way businesses leverage these data streams is transformed, different operating systems provide optimal suitability for different tasks. In the modern datacenter, data flows unimpeded from and to connected, purpose-built "stations" that are deployed on-premises, in a cloud, or at the edge.

Data of the Highest Value

One distinct purpose-built data "station" in this constant data stream is the mission-critical platform that runs a business' relational database management system (RDBMS). For production workloads, this is the station where core enterprise data of the highest value is collected, protected, queried, and leveraged for business functionalities. It is no secret that, even today, the most suitable platforms for high-value data are highly optimized platforms for transactional and query performance running RDBMS, with superior reliability, availability, and serviceability (RAS), security, and data protection. However, again, the workloads that generate or act upon this data must be liberated from any siloed environments and made available as part of a hybrid cloud.

IDC identifies such systems as having Availability Level 4 (AL4), which means that they have true fault tolerance. Several server vendors built such platforms, and together they make up a server market that is as steady as the platforms themselves. The vendors are IBM with systems that run z/OS, Unix, and/or Linux; HPE with systems that run Linux, Unix, or Windows; and Oracle with a platform that runs Linux or Unix. These systems should not be mistaken for the systems of the past that were siloed, hard to manage or program, tough to connect to, and mostly proprietary. Today, these systems are as modern as any cloudlike environment. They run open source languages; are mobile, web, and API enabled; and can participate in hybrid clouds, in some cases with Kubernetes and OpenShift. Most importantly, they increasingly play a critical role in the digital transformation of the enterprise, unlocking an organization's most valuable data for new revenue-generating opportunities.

The Benefits of Purpose-Built Platforms for High-Value Data

Purpose-built platforms for high-value data are usually highly optimized across the entire infrastructure stack as their manufacturers develop and integrate all the layers of the stack. Such end-to-end systems, in which everything from the processors to the I/O modules to firmware to the operating system is designed and fine-tuned to seamlessly work together, generally perform better on multiple attributes.

They are designed to deliver exceptional transaction performance, which is one of their fundamental workloads. They operate with faster cores, bigger caches, more memory, lower latency, and higher bandwidth I/O than general-purpose servers. Typically, they scale up to as many as 16 sockets, allowing for tremendous transactions-per-second performance with flexibility to take on surges.

Because of their mission-critical nature, they have been built with extensive RAS features in the hardware, middleware, and operating system, as well as with vendor-built applications. They deliver security through hardware-based encryption of data at rest or in flight and typically have several overlapping security solutions that make them practically impenetrable.

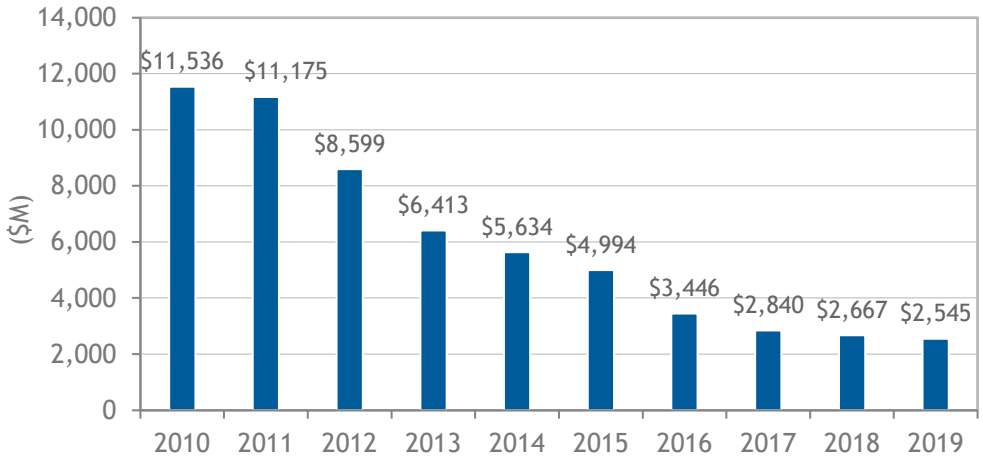
So far, though, these are the features for which these platforms are generally known. What is less well understood is that in the past six or seven years, they have been reconceptualized to enable them with the characteristics that businesses expect from a modern cloud platform. Their redesigns allow them to become fully integrated with the rest of the datacenter so that their applications and data are no longer siloed and can be leveraged in new and innovative ways for DX.

They support most of the popular open source software and have full API enablement, allowing end users to interact with the data they process in a multitude of new ways. They are web and mobile enabled, and developers can use various popular open source languages to develop on them. IDC is also seeing these systems increasingly being used as part of agile development processes and participating in organizations' DevOps. Some of them support containers, run Kubernetes, and can be operated as a hybrid cloud with OpenShift. Further, they have been equipped to run artificial intelligence (AI) and AI-enabled applications and are considered extremely suitable for high-end AI inferencing. Indeed, their transactional performance gives them a distinct advantage with this particular AI workload.

Unix-Based Platform Trends

Over the past 20 years, Unix has transitioned from being a multipurpose operating system, dominating the server market, to becoming a special-purpose operating system as the engine of several high-value data platforms. With this shift, its market size has contracted significantly, but IDC is now seeing the Unix-based server market stabilize at around $2.5 billion annually (see Figure 1).

FIGURE 1: ***Worldwide Unix Server Revenue, 2010–2019***



Source: IDC, 2020

Together, AIX (IBM), Solaris (Oracle), and HP-UX (HPE) are several flavors of Unix that continue to support mission-critical databases such as Oracle, DB2, and SQL servers. AIX dominates the market with 53.3% of worldwide Unix-based server revenue (source: IDC's Worldwide Quarterly Server Tracker, 4Q19), not just in terms of revenue but also with respect to how far out the operating system's future road map extends. All three vendors have made their systems available with Linux but have typically found that, except for HPE, customers' preferences remain largely with Unix for these specific workloads.

The reasons for this loyalty are manifold — in-house skill sets for optimally running a mission-critical platform are not easily replicated, and modernizing while remaining on a Unix platform is a lot less disruptive than replatforming, which can consist of a multiyear exercise that costs millions or even tens of millions of dollars, depending on the size of the system. Moreover, the ROI of a replatforming effort can be unclear. Businesses ask themselves the following question: We will have dedicated enormous effort and significant financial resources to switching operating systems, but what will we have gained? In general, today's Unix users understand the value of modernizing with the platform to make it part of their digital transformation efforts.

Foundational AIX Strengths and Dynamic Modernization

IBM's operating system, AIX, is a cornerstone of mission-critical computing — extremely reliable and backward compatible — but at the same time it is designed to be a *dynamic* operating environment. The operating system evolves in synch with the larger IT technology progressions to ensure that AIX-based platforms not only play well with internal and external systems but also drive DX opportunities that arise from an organization's core enterprise data and workflows.

The foundational strengths of the AIX operating system on IBM Power Systems are performance, reliability, availability, scalability, and security. Also noteworthy is the investment protection of the platform due to its long release cycles, strong binary compatibility, and ability to live migrate current and prior AIX versions to new Power Systems hardware platforms as they are released.

How dynamic AIX on Power Systems platform is can be illustrated with the most recent optimizations and modernizations that the platform has undergone. AIX now runs in the IBM Cloud, Google Cloud, Skytap Cloud on Microsoft Azure, and various other cloud service providers' clouds. This enables businesses to run AIX workloads in the public cloud without having to rewrite them, thus providing businesses with a fast path to cloud infrastructure capabilities and the ability to innovate their applications.

With regard to AI, AIX users can install open source Python-based ML packages such as NumPy, SciPy, and Sklearn by using the Python and BLAS libraries that are available in the AIX toolbox. IBM has worked closely with the open source community to contribute the code needed to deliver these capabilities on AIX. AIX users can also explore scoring and inferencing on PowerVM-based systems by using Watson Machine Learning Community Edition (WML CE). Apart from this, AI models generated with H2O Driverless AI can be leveraged natively on AIX.

IBM has fully embraced open source technologies for AIX, constantly updating the AIX open source toolbox with new or updated packages — there are now more than 300 open source packages available on AIX. Recently, about two dozen open source packages were added to the AIX Toolbox, and almost three times that many were updated. Furthermore, IBM has been contributing to cloud automation communities, such as Ansible, enabling clients to leverage the same tooling and skills across heterogeneous environments.

PowerVC is the face of virtualization and private cloud management on IBM Power Systems, which is built on top of Open Stack and makes AIX interoperable with other cloud orchestrators, such as VMware vRealize. Also, recently AIX has been enabled to integrate with cloud-native apps, containerized solutions via the IBM Cloud Paks on Red Hat Open Shift. IBM Cloud Paks are enterprise-ready, containerized software solutions that give businesses an easy and secure way to extend or move business applications to the cloud. Each IBM Cloud Pak includes containerized IBM middleware and common software services for development and management on top of a common integration layer. With this integration, clients can leverage containerized and VM-based workloads running in AIX.

The IBM Power Systems Portfolio

The AIX operating system should be viewed as an integral part of the Power Systems hardware and software ecosystem, designed together to deliver an extremely performant and secure environment that allows businesses to modernize their core enterprise workloads and adopt innovative capabilities that drive DX. For example, IDC classifies the enterprise-class platform Power Systems 980 running AIX as one of the very few platforms on the market today that achieve true fault tolerance while playing a transformative role in the datacenter.

IBM Power Systems are purpose-built platforms for extremely data-intensive workloads. The combination of the IBM POWER9 processor with large caches, very large memory, and low-latency I/O gives the Power portfolio a unique role in the datacenter for workloads that process massive volumes of data at high speed, whether in the form of database transactions or for the purpose of training an AI model with a machine learning algorithm.

Businesses often run these systems side by side. For example, they may be running AIX workloads side by side with SAP HANA; they may be running IBM Watson Machine Learning Accelerator on a scale-out system such as the Power Systems AC922 with NVIDIA V100 GPUs for AI training next to their AIX workloads; or they may be running open source databases side by side with Oracle or SAP workloads. In general, the different workloads that are running on different IBM Power Systems platforms integrate well with each other. This is also true for enterprise-class systems that run on the AIX operating system.

Key Power Systems Software Components

The following IBM Power Systems components play an important role in enabling businesses to leverage their enterprise-grade Power Systems platform with AIX for secure, highly available, cloud-based workload modernization:

- **IBM PowerVM.** IBM Power Systems server workloads are virtualized, mobile, and fully cloud enabled with PowerVM, which was recently enhanced with multiple new features, including compression and encryption of Live Partition Mobility (LPM) data, meaning that when an active partition is migrated from one Power Systems server to another, which occurs with zero downtime, the data will be automatically encrypted and compressed — an important security and performance feature.
- **IBM PowerVC.** PowerVC is the virtualization management tool that is built on OpenStack, simplifying the management of virtual resources in Power Systems environments; the software has recently been improved with multiple new features, including an export/import capability to share VM images across datacenters.
- **IBM PowerSC.** PowerSC is the platform's security portfolio, simplifying security and compliance management and featuring compliance automation, malware intrusion detection, patch management, and more; it has been enhanced with various features or even new offerings, including multifactor authentication (MFA) enablement, another important security feature. In general, security on IBM Power Systems with AIX is achieved with a comprehensive solution that includes the processor, firmware, and hypervisor, as well as the countless security features of the operating system itself, to protect data at all levels.

- **IBM PowerHA and VM Recovery Manager HA and DR.** PowerHA is a high-availability technology that helps provide near-continuous application availability and improves service reliability. It is a key contributor to IBM Enterprise Power Systems being characterized as fault tolerant (AL4) by IDC and has been improved with various features such as enhanced failover metrics and cross-cluster verification (e.g., to compare a development cluster with a test cluster); VM Recovery Manager (VMRM) is a simplified HA/DR solution based on VM replication and restart that is operating system agnostic and includes application monitoring agents for DB2, Oracle, and SAP HANA.
- **Cloud Management Console (CMC).** The Cloud Management Console provides a complete view of performance, inventory, and logging of on-premises and off-premises Power Systems infrastructure. CMC is hosted on the IBM Cloud, thereby freeing businesses from having to maintain software to monitor their infrastructure, and it helps simplify management of hybrid cloud deployments and monitoring and managing of the infrastructure.
- **Enterprise Cloud Edition 2.0.** Enterprise Cloud Edition brings together all the key components of a simplified cloud management infrastructure on top of PowerVM, including PowerSC, MFA, PowerVC, CMC, VMRM, and Aspera. It enables rapid deployment and management of a private cloud, simplified security and compliance management, simplified high availability, and accelerated large file transfers across clouds. Enterprise Cloud Edition 2.0 can be purchased with AIX 7.2 built in.

New Licensing Models

In addition, IBM has introduced flexible licensing models for AIX. Sometimes, businesses do not want to purchase perpetual licenses with a high initial investment even if they are combined with lower ongoing costs in terms of subscription and support. Instead, they prefer a pay-per-use model. IBM has been providing such a model for about two years with AIX monthly licenses, allowing businesses to evenly distribute the investment across the usage period. Businesses can order these AIX Monthly Packs in the same way as perpetual licenses. An organization could, for example, order a three-month pack if it is planning a POC or a short-term project.

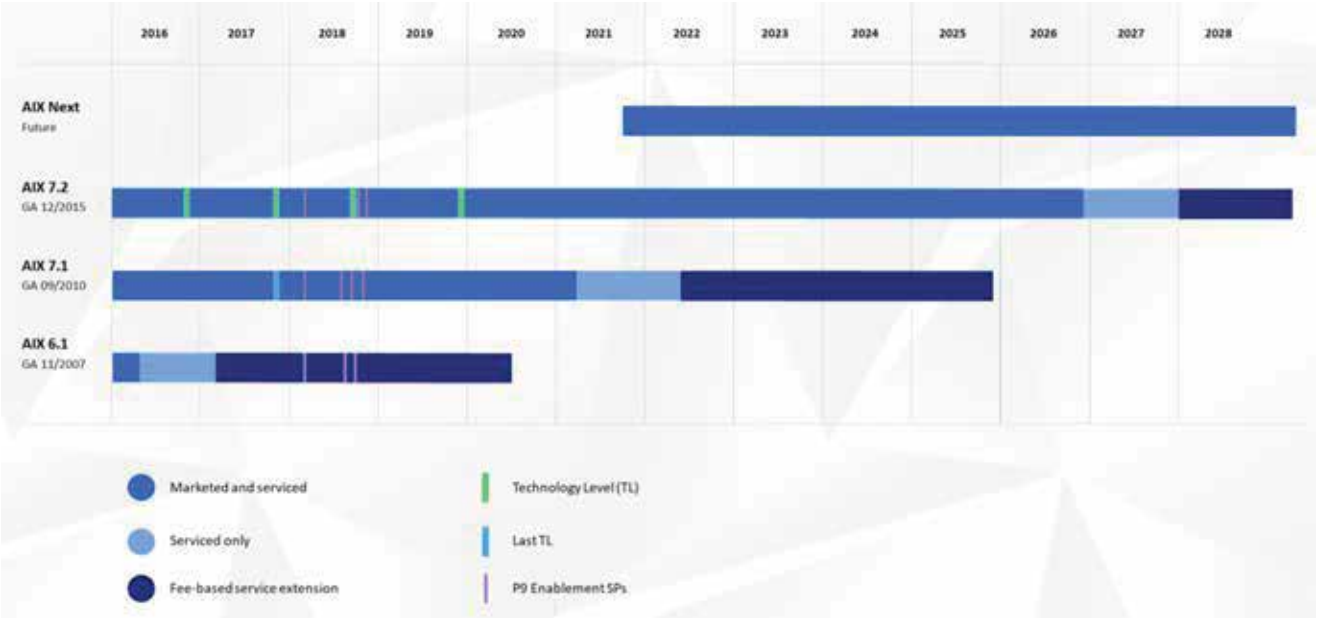
The Monthly Packs can also be used on the IBM Cloud. A business might have a system in its datacenter with a perpetual license that's running a stable workload, but for a dev/test exercise, the business could purchase a three-month package, use it on a VM on the IBM Cloud, and afterward decide whether it wants to purchase another package or switch to the perpetual license. The Monthly Packs are a fairly unique offering in that other operating system vendors do not directly allow short-term licenses, although some do provide shorter-term licenses via partners.

In addition, there are several options for AIX as you go: leveraging elastic Capacity on Demand, Enterprise Pools 2.0, or the new Power Private Cloud offering with Dynamic Capacity or consuming AIX by the minute on the IBM Cloud. These offerings make it easy for a business to quickly expand its hybrid cloud infrastructure while balancing the flexibility of public cloud infrastructure and the security and reliability from on-premises datacenters.

The Next 10 Years: An Active AIX Road Map

The AIX road map that IBM has released as a statement of direction shows anticipated general availability of the next version of AIX sometime in 2021 with a fully marketed and supported life span beyond 2030 (see Figure 2).

FIGURE 2: ***AIX Release Road Map***

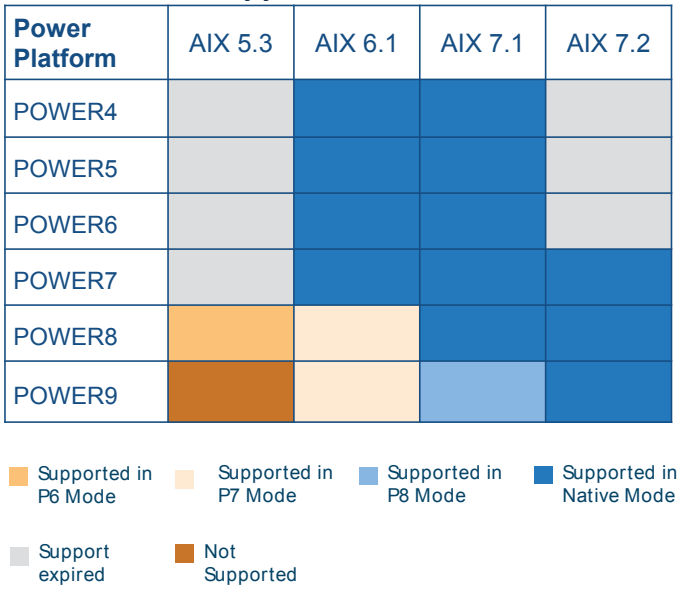


Source: IBM, 2020

With this road map, IBM demonstrates that AIX is an extremely active and future-oriented Unix platform. Organizations that depend on the operating environment should have no qualms about continuing to rely on it.

As mentioned previously, AIX and Power are fully integrated. Moreover, the AIX operating system is designed to leverage the IBM Power Systems hardware features. The matrix in Figure 3 shows which versions of AIX run on which hardware platforms, and in which modes, to take advantage of the hardware features.

FIGURE 3: ***AIX Support Matrix***



Source: IBM, 2020

Challenges for IBM

There is an important role for AIX to play in the datacenter and the cloud, which are becoming increasingly purpose driven while sharing an open, hybrid cloud platform for sharing applications and data. Purpose driven means that a workload must run on the most suitable infrastructure/operating system platform to perform tasks as well and as efficiently as possible. For the most mission-critical workloads, those that run RDBMS with extremely high transaction rates and that have to be much more available and much more secure than average infrastructure, Enterprise IBM Power Systems with AIX is the right platform. Few would dispute this.

However, the challenge for IBM is not to prove that AIX is stable and secure but to prove that IBM Power Systems with AIX also is modern, open, manageable, cloudlike, and flexible. IDC believes that is the case and that a slew of enhancements in the operating system have transformed it into a platform that plays well with an organization's entire datacenter/cloud infrastructure. These improvements have occurred on AIX 7.2 over a period of multiple years, so the marketplace may have been unaware of some or even most of them. IDC believes that AIX will continue to drive the Unix market forward and wipe out any remaining preconceptions with a new AIX release in 2021, extending the overall road map to 2030.

Conclusion

The modern datacenter is founded on purpose-built platforms that are designed to run certain workloads really well while operating on a shared plane with the rest of the datacenter. This is especially true for the platforms that run an enterprise's mission-critical workloads such as the core database. These mission-critical platforms, a significant portion of which run on Unix, are typically designed as an end-to-end infrastructure stack, integrating all the components to optimize their performance, availability, security, and manageability.

Unix is ideal for mission-critical workloads, and modern Unix implementations, such as IBM's AIX, are completely in synch with today's IT requirements. They operate as a hybrid cloud, support open source software, and run on automation platforms such as Ansible. Businesses that have modernized their platforms while staying on Unix have proven to be extremely loyal to the operating system, and IDC is seeing this market stabilize, crystallizing into a significant purpose-built datacenter segment for many enterprises' highest-value data processing.

IDC is seeing the Unix market stabilize, crystallizing into a significant purpose-built datacenter segment for many enterprises' highest-value data processing.

About the Analyst

Peter Rutten, *Research Director, Infrastructure Systems, Platforms, and Technologies*

Mr. Rutten focuses on high-end, accelerated, and heterogeneous infrastructure and their use cases, which include supercomputing, massively parallel computing, artificial intelligence (AI) and analytics, and in-memory computing. His research on high-end servers includes mission-critical x86 platforms, mainframes, and RISC-based systems as well as their operating environments (Linux, z/OS, Unix).

IDC Custom Solutions

IDC Research, Inc.
5 Speen Street
Framingham, MA 01701, USA
T 508.872.8200
F 508.935.4015
Twitter @IDC
idc-insights-community.com
www.idc.com