SPOTLIGHT

Sponsored by: HPE and Intel

In the healthcare industry, high availability and security are paramount to supporting the clinical and operational requirements of provider organizations and privacy of patients, especially as organizations start to process and analyze ever-larger amounts of data.

In Healthcare, Mission-Critical Platforms Are a Crucial Part of Digital Transformation

July 2023

Written by: Peter Rutten, Research Vice President, Performance-Intensive Computing, and Mutaz Shegewi, Research Director, Provider IT Transformation Strategies

Mission-Critical Digital Transformation in the Healthcare Industry

Modern analytics environments provide the crucial underpinning for firms transforming themselves into data-driven enterprises. Being data driven enables innovative enterprises to compete more effectively in the digital economy via analysis of data coming from core enterprise applications and emerging initiatives such as the Internet of Things (IoT), robotics, next-generation security, and next-generation supply chain automation. As firms seek to create and deliver digital offerings and experiences, insights obtained from data are paramount in making key business decisions.

In the healthcare industry, core provider software solutions include those that support clinical documentation and workflow (e.g., electronic health records [EHRs]), medical imaging and archiving (e.g., PACS/RIS, VNA), and revenue cycle management. Healthcare providers including hospitals and medical centers (acute/inpatient care); outpatient centers, clinics, and physician groups/practices (ambulatory/outpatient care); and long-term care hospitals, inpatient rehabilitation facilities, skilled nursing facilities, and home health agencies (post-acute/long-term care) continue to prioritize goals that aim to drive better outcomes (e.g., quality, safety, access), facilitate technological advancements, and remove any systemic barriers to care.

AT A GLANCE

KEY STATS

- 58.3% of healthcare providers are increasing their overall organizationwide IT spending on clinical documentation and workflow over the 2022–2023 period.
- 94.1% of healthcare organizations have begun or are well underway in applying health IT to advance their digital transformation initiatives.
- Key areas include clinical workflow optimization, advanced analytics, virtual care, IoT, mobility, the move to the cloud, and "patient of one" initiatives (e.g., 360-degree views and longitudinal health records).

A move to a future-ready real-time organization includes the convergence of care-centric transaction processing and data-centric analytics systems. To speed up analytics, healthcare organizations are deploying in-memory and memory-centric databases, infusing data analytics platforms with high-performance technologies, and using a highly available and secure conduit for data movement between the various application tiers.

In the healthcare industry, high availability and security are paramount to supporting the clinical and operational requirements of provider organizations and privacy of patients. Data continues to grow exponentially in healthcare as

organizations want to make the shift from being data rich to data driven. Healthcare is also frequently targeted by rogue actors due to the richness and value of its data, particularly in electronic health records. The stakes are high from a reputational and legal standpoint. Health Insurance Portability and Accountability Act (HIPAA) violations, for example, can significantly damage an organization's reputation, in addition to the enforcement of hefty penalties. Ensuring high availability and security helps organizations put information into context, while infusing resiliency into business, clinical, and operational functions.

Mission-critical real-time online transaction processing (OLTP) and online analytical processing (OLAP) platforms are designed for hosting core relational database management systems (RDBMSs) and modern data management systems. These platforms combine unique characteristics around transactional and analytical performance, reliability, availability, and security.

Today's mission-critical platforms have been standardized, modernized, opened up, and cloudified, but they have also maintained their unique characteristics: unmatched transaction processing and analytical performance. The move to the cloud off these platforms has therefore been slow, and IDC has even seen a return to private cloud on premises, referred to as "repatriation," for transactional and analytics applications.

IDC does not expect growth in public cloud adoption to slow, but a significant portion of organizations will leverage a private cloud to modernize their large installed base of noncloud applications. Many of these noncloud applications will be mission-critical applications, including core databases, business processing, or apps that execute Big Data and analytics and artificial intelligence (AI).

In the healthcare industry, multiple events trigger cloud adoption, including IT budgets being constrained or reduced, data growing beyond the capacity of existing systems, the need to keep up with compliance across regions (e.g., HIPAA, GDPR), and the ability to leverage functionality and services only found in cloud offerings. The greatest benefits of the cloud include its ability to help drive innovation and digital transformation as well as to simplify and standardize IT infrastructure and application platforms. The move to the cloud will increase as a result of the next-generation clinical documentation use case that seeks to transform electronic health records from a traditional system of record for routine clinical data entry to a system or platform for engagement and value-added activities such as patient engagement and population health.

Benefits of Mission-Critical Platforms in the Healthcare Industry

Business processing, decision support, and analytics have never fared well in horizontally scaled environments. These are demanding workloads that require maximum resources to process multiple terabytes of data. And when these resources — lots of processors that are close together and lots of flat RAM that is globally addressable for in-memory computing — are packaged in a single system, the benefits, compared with scale-out environments, are significant.

The healthcare data explosion will exceed 10ZB by 2025. Although healthcare and life science organizations currently manage on average 21PB of data (i.e., 25% less than the industry average), at the same time, they retain data almost 20% longer. These data retention trends are due to regulatory mandates even though regulations differ around the world. Healthcare organizations tend to retain patient and financial data the longest, while typically holding operational data for shorter periods. Regulations require some data to be retained for decades (e.g., health records). In contrast, other data types are retained for as little as 30 days (e.g., video surveillance data).

The large memory footprints of scale-up systems allow for large and growing databases to be completely held in memory, eliminating the latencies of disk access. Latency is also much reduced because of the use of interconnects that allow for dynamic scaling rather than the complex and extensive networks needed to connect nodes in a scale-out environment. Power consumption and cooling costs are significantly lower, as are software licensing costs.

In the healthcare industry, lower cost is important because of dwindling profit margins and the reality that all organizations must continue to navigate the aftereffects of the global COVID-19 pandemic combined with the impact of everything from inflationary pressures to supply chain constraints and pervasive labor and skills shortages. Healthcare organizations also face escalating cybersecurity risks, shifting care models, and an ongoing need to strive for service excellence while addressing broader systemic challenges. Therefore, organizations need to invest across a wide range of areas to bolster their health IT portfolio and be able to contend with challenges.

Scale-up systems are also suitable for consolidation projects, as they are easier to implement and more efficient to manage and operate than scale-out clusters. They take up a smaller footprint and provide greater reliability and availability.

IDC has seen scale-up x86-based servers evolve to the highest levels of availability, meeting service levels that are demanded by mission-critical workloads for the data-driven organization. They deliver extreme reliability, availability, and serviceability (RAS), even reaching availability level 4 (AL4) in IDC's High Availability framework.

In terms of economics, many of today's scale-up systems are nothing like legacy scale-up systems — they leverage the same standardized components (memory, processors, and storage) that scale-out servers are built with, albeit with their own specifications rather than the proprietary components from the past.

The idea that scale-up systems are too costly is simply no longer valid, especially if they are available on a consumption model basis, as many are.

Digital Transformation Trends in the Healthcare Industry

Digital transformation in healthcare can no longer be ignored. In fact, it is necessary for organizations to survive, considering preexisting and emerging challenges. Healthcare digital transformation is not only a matter of modernizing existing health IT systems but a shift toward platform thinking that combines multiplied innovation to enhance experiences for all stakeholders while improving performance.

Compared with other industries, healthcare is often considered a technology laggard. However, because of mounting pressures, healthcare is digitally transforming — and rather quickly. In the United States, for example, the ongoing shift in reimbursement models from volume- to value-based programs has driven increased demand for new revenue streams as these programs aim to improve care for individuals and populations while reducing costs. Organizations require quicker and more convenient access to clinical, claims, population health, social determinants of health, and other sources of data (e.g., consumer, research, and patient-generated health data). The impact of the COVID-19 pandemic has led to the realization that organizations must modernize as a matter of preparedness for future potential crises and to cater to shifting expectations of care-seeking consumers.

An aging healthcare workforce combined with workforce shortages, particularly in specific medical disciplines (e.g., radiology and primary care), coupled with the rise of consumerism, further drives the need for service model and care offering transformation (e.g., virtual care and teleradiology) powered by digital capabilities. In addition, patients with one

or more chronic conditions, aging populations, and growing healthcare expenditures further emphasize a need for healthcare organizations to undertake digital transformation.

In the healthcare industry, electronic health records (e.g., Epic) give providers access to a digital version of a patient's record. The next generation of these clinical documentation solutions aims to build on the collective learnings and aspirations of the industry moving forward to support new models of care, more integrated care pathways, digital patient engagement and experience, value-based reimbursement, revenue cycle optimization, and reduced workforce burden (i.e., physician burnout).

AI is playing an increasingly important role in the healthcare industry, as healthcare data is analyzed and acted on by AI models for better diagnostics, patient care, hospital management, and so forth. In diagnostics, AI can play a role by having trained models from large amounts of healthcare data support medical professionals. To avoid training AI models on small sets of data from just one hospital or medical institution (because that data cannot be shared for regulatory reasons), organizations have begun to embrace swarm learning — a technique whereby data is analyzed locally and only the results are shared in an AI model that is accessible to the larger network of partnerships. This protects patient confidentiality, data privacy, and data security.

Life-saving care and critical decision making run central to healthcare. Therefore, having timely, accurate, and secure data on hand will impact the quality and safety of care. Yet, in many healthcare settings, an efficient and seamless flow of real-time patient information cannot be realized. Furthermore, healthcare settings and their respective systems face a growing number of new clinical, operational, and regulatory requirements that introduce new challenges and risks. Therefore, having access to the right information at the right time for the right person, at all points of care, will improve outcomes and reduce risks. Yet the explosion of data, combined with multiple sources of truth and an ever-expanding set of storage locations, has made the realization of excellence in care quite challenging.

The Value of HPE Mission-Critical Scale-Up Servers for Electronic Healthcare Records

The original HPE Superdome Flex, powered by Intel Xeon Scalable processors, was born out of Hewlett Packard Enterprise's (HPE's) efforts to deliver a modular, standards-based mission-critical system with maximum flexibility, performance, and reliability. The platform has seen strong adoption since its introduction, covering businesses in all geographies and spanning industries from telecommunications to banking, healthcare, manufacturing, the public sector, education, and more.

Midsize to large HPE Superdome Flex systems (12–32 sockets) are running in hundreds of production environments, although, notably, the majority of shipped units are 4- and 8-socket systems. This reflects both buyer desires to equip for growth and the market requirement for scale-up environments of a smaller size. To address this need, HPE introduced the HPE Superdome Flex 280 model that starts at a lower 2-socket entry point and offers more granular scaling.

Recently, HPE introduced a third and next-generation member to this platform family: HPE Compute Scale-up Server 3200. The Scale-up Server 3200 features 4th Gen Intel Xeon Scalable processors and is aimed at digital-first modernization for mission-critical applications.

In the healthcare industry, equipping for growth is important given the evolution toward more integrated, dynamic, and proximate care models (e.g., home, retail, urgent, and virtual care). Being equipped with the right capabilities will eliminate any guesswork from common scenarios such as the need to add users and facilities, application upgrades, data

migrations, business continuity, consolidation scenarios, and the alignment of clinical and operational actions to outcomes improvement. Finally, evolving interoperability standards for Fast Healthcare Interoperability Resources (FHIR), microservices, X12, HL7, and API are homogenizing data, in addition to hybrid cloud and multicloud solutions coming to scale for enterprises to further unlock and utilize data value.

HPE's mission-critical scale-up servers have a modular design that enables firms to start small and scale up as their needs grow. They are built with a multisocket, multicore x86 architecture.

The New HPE Compute Scale-up Server 3200

The new HPE Compute Scale-up Server 3200, featuring 4th Gen Intel Xeon Scalable processors, is the follow-on to the HPE Superdome Flex model. It is a powerful single-system solution that uses a modular, scale-up architecture to handle critical databases such as SAP HANA, Oracle, and SQL Server and workloads such as electronic healthcare records, core banking systems, and enterprise resource planning. With up to 60 cores per processor, more than double compared with the prior generation, and with built-in accelerators, these processors are designed to significantly boost performance across the fastest-growing workloads, including AI and data analytics.

With the HPE Compute Scale-up Server 3200, businesses can start with 4 processors in a 5U chassis and scale in increments of 4 — up to 16. The chassis connect to one another with HPE technology to form a high-bandwidth, ultralow-latency fabric via cabling. This "glued" architecture enables the HPE Compute Scale-up Server 3200 to scale as a single server beyond the 8-socket upper limit of Intel's design.

The system can deliver up to 32TB of shared DDR5 memory using industry-standard 128GB DIMMs. With 16 DIMM slots per socket and a data transfer rate of up to 4,800MTps (megatransfers per second), this server offers twice the memory bandwidth compared with prior generations. The server's flexible I/O subsystem is based on PCIe 5.0 and delivers four times the I/O bandwidth that the previous model provided. Businesses can choose either 16, 12, or 6 I/O slots per 4-socket chassis, depending on their workload requirements.

The Large HPE Superdome Flex

For midsize to large environments, the HPE Superdome Flex model features 2nd Gen Intel Xeon Scalable Processors and scales from 4 sockets to 32 sockets in 4-socket increments and from 768GB to 48TB of shared memory in a single system. The system features a modular scale-up architecture and a 5U 4-socket chassis building block. Within each chassis, there are two ASICs utilizing eighth-generation HPE technology that connect the chassis to form a high-bandwidth, ultralow-latency fabric via cabling. This is the same "glued" architecture utilized by the new HPE Compute Scale-up Server 3200.

In addition, HPE Superdome Flex can be carved up into nPars (electrically isolated hard partitions) to physically separate workloads within a common platform. When fully configured, the HPE Superdome Flex model supports up to 128 Gen 3 PCIe card slots. These slots can be used for external storage connectivity; hardware accelerators like GPUs, 32Gb Fibre Channel cards, Mellanox InfiniBand, Ethernet cards, and NVMe cards; and other peripherals. The Superdome Flex model features 2nd Gen Intel Xeon Scalable processors.

The Lower Entry Point HPE Superdome Flex 280

The HPE Superdome Flex 280 enables businesses to scale and provision to the exact level needed while equipping for growth without having to overprovision. For small to midsize environments, the HPE Superdome Flex 280 utilizes a 5U chassis and scales seamlessly from 2 to 8 sockets in 2-socket increments (up to 4 sockets per chassis) as a single system.

The server is designed to provide 64GB to 24TB of shared memory. The chassis utilizes a "glueless" architecture and leverages Intel Ultra Path Interconnect (Intel UPI) links to connect.

The HPE Superdome Flex 280 model features select 3rd Gen Intel Xeon Scalable processors and supports up to 28 cores per processor, driving enhanced performance, throughput, and CPU frequencies compared with previous-generation processors. The 4- to 8-socket 3rd Gen Intel Xeon Scalable processors deliver multisocket core count density, with up to 224 cores per platform in an 8-socket configuration. Up to 6 Intel Ultra Path Interconnect channels increase the platform scalability and improve inter-CPU bandwidth for I/O-intensive workloads. When fully configured, the HPE Superdome Flex 280 model supports up to 32 PCIe 3.0 card slots.

Across the Three Platforms

Multiple Operating Environments

As HPE does not modify Linux for HPE Superdome Flex, compatibility can be expected with any peripherals running under standard SUSE Linux Enterprise Server (SLES), Red Hat Enterprise Linux (RHEL), and Oracle Linux distributions. HPE Superdome Flex, HPE Compute Scale-up Server 3200, and HPE Superdome Flex 280 support multiple additional standard operating environments and virtualization technologies, including Windows Server, VMware, and KVM. As the platforms run on standard, unmodified Linux, they support all the certified stacks from Red Hat and SUSE, including containers and container management software such as Docker and Kubernetes.

Reliability, Availability, and Serviceability

The three platforms feature many RAS capabilities at all levels of the IT stack. Some key features are *Firmware First,* which ensures error containment at the firmware level, before any interruption can occur at the operating system layer; *self-healing capabilities* that provide several mechanisms to avoid unplanned downtime when faults do occur; and the *Analysis Engine,* which reduces human error through predictive fault handling. Additional processor, memory, and platform RAS features are also supported. Furthermore, the platforms support *HPE Serviceguard for Linux,* a highly integrated software suite that delivers fully automated, proven high availability (HA) and disaster recovery (DR) capabilities at the application layer. In addition, the HPE Compute Scale-up Server 3200 has the following new RAS features:

- I/O slot–enhanced Downstream Port Containment support and HPE PCIe isolation capabilities
- Redundant interchassis management network
- New automatic deconfiguration capabilities for the I/O subsystem

Security

HPE's mission-critical scale-up servers have been designed with a security strategy to minimize threat exposure through several mechanisms. These include an isolated manageability subsystem that prevents the host operating system from accessing the management network and hardware that prevents the host from updating system and management firmware directly, eliminating a major source of vulnerabilities. The security strategy also limits the number of common industry interfaces (which add security vulnerabilities) to only those that customers need. To help secure today's most sensitive data and workloads, Intel Xeon Scalable processors also come with a suite of innovative, hardware-based security technologies, ensuring a critical base of protection. The HPE Superdome Flex 280 model adds Silicon Root of Trust protection implemented directly in HPE-controlled specialized hardware to enable detection of potentially

compromised firmware and prevent its execution. In addition, the HPE Compute Scale-up Server 3200 has several new security features, including the following:

- HPE's Silicon Root of Trust (In HPE Superdome Flex 280, HPE provides a Silicon Root of Trust, but that does not leverage the HPE iLO ASIC; the iLO ASIC allows HPE to now add it earlier in the manufacturing process, increasing security.)
- HPE Trusted Platform Module (TPM) for the Rack Management Controller (RMC) (In HPE Superdome Flex 280, there is one TPM for the system; the HPE Compute Scale-up Server 3200 has added one TPM for the Rack Management Controller.)
- Standard chassis intrusion detection
- Secure Boot defaults to "on"

As-a-Service Consumption

Aligned with HPE's strategy to offer the company's entire portfolio through a range of subscription, pay-per-use, and consumption-driven offerings in the next three years, all three platforms can be consumed as a service through HPE GreenLake.

Artificial Intelligence

As AI increasingly takes hold in the healthcare industry, HPE is well positioned for the shift. HPE has been at the forefront of supporting swarm learning solutions from datacenter to edge to cloud, while Intel has a broad portfolio of processors with built-in AI acceleration using Intel Deep Learning Boost (Intel DL Boost), giving flexibility to run complex AI workloads on the same hardware as an organization's existing workloads.

Challenges

IDC data shows that the move to cloud continues unabated with more than 60% of organizations using public cloud for one or two small applications. However, not all workloads are suitable for the public cloud. The challenge for HPE is to make a case to organizations in the healthcare industry that workloads, such as those that use Epic and EHRs, require a mission-critical scale-up platform.

IDC believes that scale-up x86-based platforms such as the HPE Superdome Flex family and the new HPE Compute Scale-up Server 3200 will continue to command traction among healthcare organizations that require:

- Scale-up multisocket design for high-performance scaling
- Security, availability, and reliability for mission-critical deployments
- Flexible and modular design for opex-friendly deployments
- Optimizations for in-memory databases and real-time analytics applications
- Support for an open standards–based and cloud-ready design for deploying hybrid IT
- As-a-service consumption while maintaining on-premises control

Conclusion

In the healthcare industry, high availability and security are paramount to supporting the clinical and operational requirements of provider organizations and privacy of patients. Data continues to grow exponentially, and healthcare is frequently targeted by rogue actors due to the richness and value of its data. Healthcare organizations must be able to process and analyze vast amounts of data while infusing resiliency into business, clinical, and operational functions. And equipping for growth is important given the evolution toward more integrated, dynamic, and proximate care models.

Mission-critical real-time online transaction processing and online analytical processing platforms are designed for such environments, hosting core relational database management systems and modern data management systems. These platforms combine unique characteristics around transactional and analytical performance, reliability, availability, and security.

Healthcare organizations require quicker and more convenient access to clinical, claims, population health, social determinants of health, and other sources of data.

IDC does not expect growth in public cloud adoption to slow, but a significant portion of organizations, including in the healthcare industry, will leverage a private cloud to modernize their large installed base of noncloud applications. Many of these noncloud applications will be their mission-critical applications, including core databases, business processing, or apps that execute Big Data and analytics and AI.

HPE Compute Scale-up Server 3200 and the HPE Superdome Flex family provide such a platform for healthcare providers, comprehensively addressing the many unique challenges that exist in this industry, including through the platforms' modular design that helps organizations equip for growth and with HPE's as-a-service models that enable organizations to run the HPE mission-critical scale-up servers as a service, essentially as if it were a cloud. IDC recommends that providers consider these platforms as part of their modernization strategy.

About the Analysts

Peter Rutten, *Research Vice President, Performance-Intensive Computing*

Peter Rutten is a research vice president within IDC's Worldwide Infrastructure practice, covering research on computing platforms. Mr. Rutten is IDC's global research lead on performance-intensive computing solutions and use cases.

Mutaz Shegewi, *Research Director, Provider IT Transformation Strategies*

Mutaz Shegewi leads the provider research practice at IDC Health Insights covering topics of most relevance to healthcare provider organizations looking to digitally transform and become more digitally native than their competition.

MESSAGE FROM THE SPONSOR

More About HPE

HPE serves thousands of healthcare organizations worldwide and has delivered solutions to healthcare and life sciences for decades, providing the right technologies and expertise to help deliver transformative outcomes. HPE aims to improve health and community well-being through technology and health information that is accessible whenever and wherever it matters most, including enabling secure and reliable access to patient records, protecting and mobilizing medical imaging data, and accelerating medical insight and scientific discovery. With an unmatched combination of expertise, consumption-based models and mission critical infrastructure across servers, storage, and networking, HPE can help healthcare organizations react with digital initiatives and solutions that positively impact patients, providers, and caregivers across the entire continuum of care.

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

IDC Research, Inc.
140 Kendrick Street
Building B
Needham, MA 02494, USA
T 508.872.8200
F 508.935.4015
Twitter @IDC
idc-insights-community.com
www.idc.com