

The logo for 451 Research, featuring the number '451' in white inside a blue circle, followed by the word 'Research' in a dark blue sans-serif font.

451

Research®

PATHFINDER REPORT

The background of the page features a light blue and white geometric design with diagonal lines and circles. Three circles are connected by lines, each containing a blurred image of a circuit board. The main title is in a large, orange, sans-serif font.

Delivering the Cloud Experience on Hyperconverged Infrastructure

**HCI EVOLUTION REQUIRES HARDWARE AND
SOFTWARE INNOVATION TO ACCELERATE
NEXT-GENERATION WORKLOADS**

COMMISSIONED BY



FEBRUARY 2019

©COPYRIGHT 2019 451 RESEARCH. ALL RIGHTS RESERVED.

About this paper

A Pathfinder paper navigates decision-makers through the issues surrounding a specific technology or business case, explores the business value of adoption, and recommends the range of considerations and concrete next steps in the decision-making process.

ABOUT THE AUTHOR



HENRY BALTAZAR

RESEARCH VICE PRESIDENT

Henry Baltazar is a Research Vice President for the Storage Channel at 451 Research. Henry returned to 451 Research after spending nearly three years at Forrester Research as a senior analyst serving Infrastructure & Operations

Professionals and advising Forrester clients on datacenter infrastructure technologies. Henry has evaluated and tested storage hardware and software offerings for more than 15 years as an industry analyst and as a journalist.



LIAM ROGERS

ASSOCIATE ANALYST

As an Associate Analyst in 451 Research's Storage Channel, Liam Rogers covers technology and business-model innovation across the enterprise storage landscape – spanning primary storage systems and software, backup

and recovery, archiving, cloud storage services, cloud-enabling storage technologies and storage management. This includes markets such as data management, hyperconverged infrastructure and software-defined storage.

Executive Summary

Hyperconverged infrastructure (HCI) has been a driving force in simplifying IT for the past few years, enabling it to rapidly expand its role in enterprise, midsize and service provider environments. Its simplicity allows HCI platforms to be implemented and managed by generalist IT staff without consuming valuable time from storage and network specialists. HCI has been driven largely by software innovation up to this point, but changing workload dynamics and increasing performance demands will require HCI vendors to go beyond software features alone to expand overall solution capabilities. Given that cloud in public, private and hybrid formats will figure largely in future IT environments, on-premises infrastructures must not only integrate well with public cloud service providers but must also become more cloud-like in their delivery of IT resources.

While the simplicity introduced by HCI software has been nothing short of a phenomenon in the growth of the product, to become an even larger force in the infrastructure market, HCI vendors will need to do more than incrementally boost performance through software updates and wait around for the latest X86 processors to be released on commodity servers. Customer expectations continue to rise as more business-critical workloads migrate to HCI and applications become larger-scale with the emergence of IoT, artificial intelligence and other next-generation use cases. To drive adoption, HCI vendors will need to introduce new systems that incorporate hardware-acceleration innovations optimized through platform-level engineering to increase performance, scalability and efficiency.

In this paper we discuss:

- The current infrastructure challenges limiting cloud-like delivery experiences.
- The evolving role of HCI and its potential to impact and become a catalyst for IT hybrid cloud transformation.
- The opportunity to integrate emerging hardware technologies with HCI software innovations to meet requirements for hybrid cloud and next-generation workloads.

Infrastructure Challenges

Infrastructure professionals are under pressure to improve their ability to scale to deliver increasingly distributed IT resources to their applications and stakeholders. High capex costs and the manual process of meeting capacity growth using traditional three-tier infrastructure have been key issues for decades and continue to be major problems for organizations. Infrastructure professionals need to reduce the time they spend on core datacenter operations while delivering higher performance and data management. At the same time, there is a growing need to expand IT resources to the edge to bring computing closer to the source of massive data generation and free staff to implement emerging technologies such as AI/ML. All of this is forcing organizations to rethink how they store, deliver and protect their data.

Beyond these core issues, infrastructure professionals are under pressure to match the rapid delivery and elasticity of public cloud services – to essentially make their infrastructure more ‘cloud-like.’ For this reason, key areas such as resource provisioning time are under intense scrutiny and will require improvements such as automation to meet the expectations of an increasingly impatient customer base.

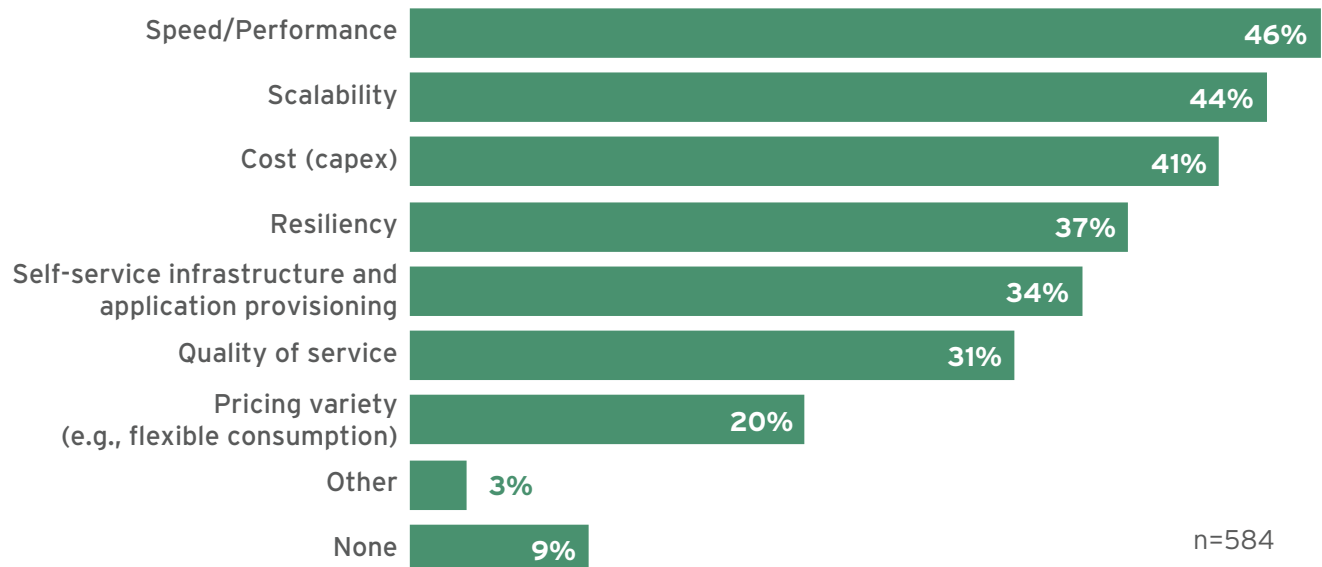
Performance, Scalability and Cost are Holding Back On-Premises Infrastructures

In our *VotE: Storage, Workloads and Key Projects 2018* survey, we asked respondents about where their on-premises infrastructure needs improvement, and speed/performance was cited most often – by 46% of respondents. Scalability was a close second with 44% of respondents citing it as a weakness.

Figure 1: On-premises infrastructure attributes to be improved to meet the needs of customers

Source: 451 Research's *Voice of the Enterprise: Storage, Workloads and Key Projects 2018*

Q. Which of the following attributes of your on-premises infrastructure are you looking to improve to meet the current and future needs of your customers? Please select all that apply.



PATHFINDER | DELIVERING THE CLOUD EXPERIENCE ON
HYPERCONVERGED INFRASTRUCTURE

Several key challenges are forcing organizations to rethink their infrastructure architectures, including:

- **Speed/performance:** Rising customer expectations and the need to create business value more quickly require improved infrastructure performance. In response, organizations will need to improve the processing, storage and networking performance of their infrastructures. They can accomplish this using hardware acceleration and innovative system design. In addition to centralized datacenters, vendors must improve performance in edge environments to deal with emerging IoT and AI/ML use cases.
- **Scalability:** Data is incredibly valuable, but for many organizations, the volume of unstructured data is growing at a rate that is disproportionate to the relatively modest increases in storage budgets, which presents a capacity and management challenge for many. Data growth has been an issue for several years, but it is further complicated by the changes happening in the compliance and regulatory landscape. With data privacy regulations such as the EU's GDPR, organizations must ensure their infrastructure can keep pace with data growth while ensuring that the stored data is in compliance.
- **Cost:** The majority of organizations are in unsustainable situations: they are facing rapid workload growth while most of their (relatively modest) infrastructure budgets are tied up in maintaining current operations. Going forward, technology buyers need to look at TCO savings not just from a capex perspective, but also from the perspective of opex savings delivered from their investments to enable reinvestment of both time and budget into strategic initiatives.

Growing Role of HCI in Datacenters and Beyond

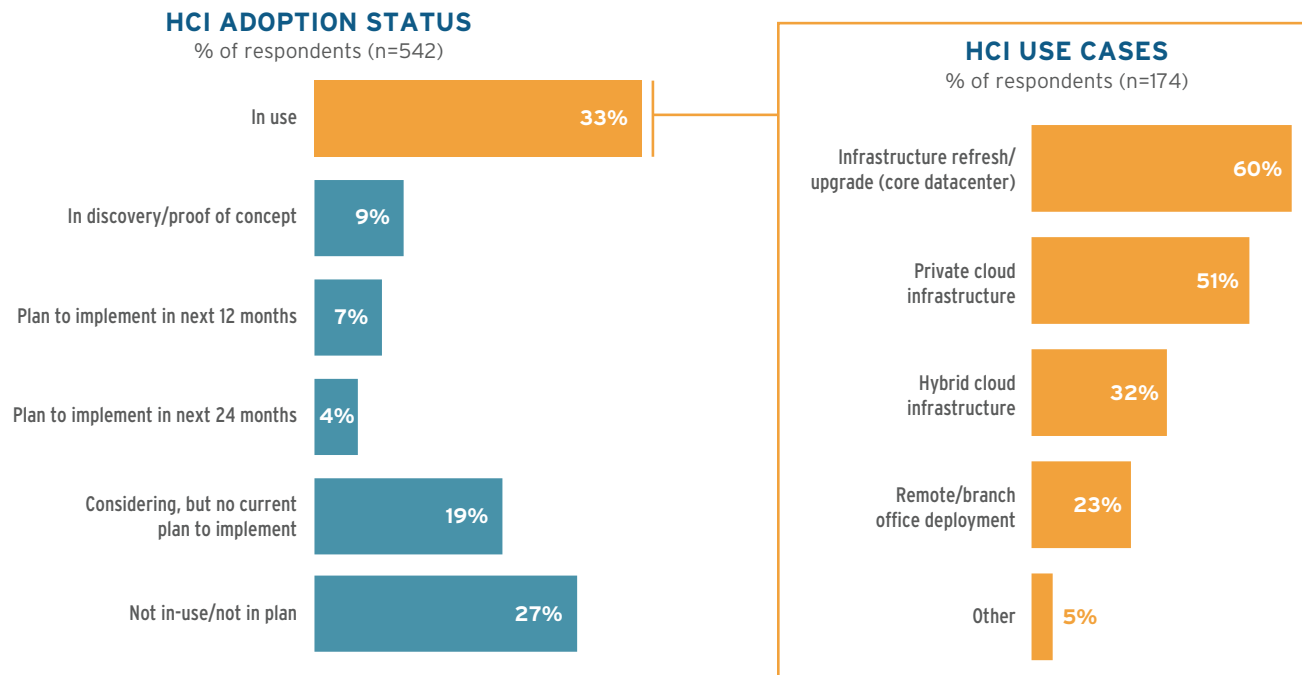
HCI is one route organizations can take to help modernize their IT environments and simplify both the deployment of infrastructure and its ongoing management. In the VotE: Storage Workloads and Key Projects survey, 53% of respondents indicated that they have HCI either in use or in plan within the next two years, and another 19% are considering it, though it's not in their current plan.

Figure 2: HCI plays key role in cloud strategy

Source: 451 Research's Voice of the Enterprise: Storage, Workloads and Key Projects 2018

Q: Please indicate your organization's adoption status for hyperconverged infrastructure (e.g., Nutanix, VMware VSAN).

Q: What role does hyperconverged infrastructure serve in your organization? Please select all that apply.



Based on our survey findings, the majority of current HCI adopters are using the technology as part of their infrastructure modernization initiative and/or for their private cloud infrastructure (60% and 51%, respectively) as they look for a solution that will meet application and governance requirements and deliver a sustainable cost model. HCI is also a prominent component of many hybrid cloud strategies because it can help bridge the gap between the functionality of on-premises infrastructure with the desirable aspects of public cloud such as easier provisioning, simplified management and opex pricing models.

HCI Requirements to Deliver a Cloud-Like Experience

Cloud vendors have elevated end-user expectations about the infrastructure experience through delivery of their simplified offerings that provision resources and scale on demand, provide integrated management of their clients' application ecosystem, and utilize single-source support for any issues their clients experience. This experience is driving organizations to plan and execute cloud strategies that use HCI solutions to deliver their private cloud and act as the foundation of their hybrid cloud architecture. Created from software innovations, HCI solutions eliminate the silos between compute, storage and networking with the promise of a cloud-like experience on-premises. The relationship between the hardware and the software, however, creates new silos that limit the promised cloud-like experience if there is not tight integration between them.

A common misconception among HCI buyers is that software-only models will give them the ability to run HCI software on any X86 platform and allow them to execute their cloud strategy at a lower cost, and without risking hardware lock-in, but there are some key factors to consider. Organizations choosing to invest in software-only HCI models still need to standardize on an approved hardware platform, configure that platform on a component level to the hardware compatibility lists published by the software vendor and then either do the integration in-house or pay for a third party to do the integration. These hidden costs can quickly add up in both capital and operational costs. There are also limitations on mixing server vendors in a cluster, so once a platform is selected, it's there to stay. Once the platform is configured and deployed, users will then have to manage multiple vendors across the software, server and component stack. This can add complexity at every stage – deployment, managing updates, and support. In the end, the teams that benefit from this model are procurement teams that can shop for the least expensive options to save on capex costs and don't have a stake in long-term TCO or user experience.

HCI solutions can deliver the cloud-like experience that most organizations envision as they put together their cloud strategies, but doing so requires tight integration between all components with single-source support for the solution stack. Because HCI solutions are on-premises infrastructure, they will never be truly invisible, but IT teams can minimize their time spent in areas such as deployment, lifecycle management (including hardware and software updates), and support through the deployment of integrated appliances. The appliance model is not a new concept; in fact, most HCI vendors in the market today have gone this route at one point. Some have moved away from it to broaden their distribution reach.

From an infrastructure hardware perspective, we are in the middle of a handful of major technology transitions that will reshape what our environments and business processes look like in the near future, so organizations looking to invest in an HCI solution should carefully consider the level of integration a vendor provides in order to yield the desired outcomes of their investment.

Leveraging Hardware Innovations to Deliver a Cloud-Like Experience

As anyone who has been involved in infrastructure knows, performance directly impacts application speeds, scalability and operating costs. Given that innovation takes place at different rates and at different times among the compute, networking and storage stacks, each of these sectors at one point has been the bottleneck that forces organizations to rethink and re-architect their on-premises infrastructure components. HCI solutions are not immune to this, and vendors that focus only on the software stack leave the end user with the challenge of keeping up with the rapid pace of innovation being delivered from component manufacturers. HCI has emerged as a powerful new infrastructure platform, and many organizations are now looking to leverage HCI to simplify the management and delivery of IT resources. All-flash systems have become the norm for application storage, and this transition from disk to flash will be driven even further by the advent of storage-class memory (SCM), such as 3D XPoint, which is even faster than NAND flash.

The growing popularity of graphics processing units (GPUs), application-specific integrated circuits (ASICs), and field-programmable gate arrays (FPGAs) is a major factor that cannot be ignored since these accelerators have already proven to be faster and more efficient than X86 in some specific situations. This is an important factor in a variety of use cases such as artificial intelligence and deep learning where overcoming the slowdown of Moore's Law is important. How organizations leverage these powerful new technologies will be a factor in how well they are able to compete with more aggressive competitors that will also be looking to harness performance and efficiency advantages. HCI solutions delivered as fully integrated appliances remove uncertainty in the adoption of these technologies. They are optimized out of the box with all components validated at the factory so that they can be deployed faster, and users can extract value from their investment quickly.

Although many of these technologies have existed for years (with the exception of SCM), mainstream adoption in infrastructure hardware platforms has been elusive. That being said, the future of infrastructure will be determined not only by the continued innovation in commodity X86 processors, but also by several key technologies that have been generating growing interest, such as:

- **Network:** The network is the backbone of a clustered architecture, but it is a commonly overlooked aspect of the overall solution with HCI vendors leaving it as a DIY project for end users. These technologies, however, will make network integration of an HCI solution play an even more vital role in the overall performance delivered. The benefits of accelerator components are quickly diminished if the network connecting the nodes in an HCI cluster is not optimized and cannot provide the high east-west bandwidth needed between nodes to support the cluster. While 10Gbps connectivity has become the standard for HCI clusters, faster connectivity is on the rise to support higher performance in a cluster, so the optimal and reliable configuration of the network and continued innovation in network ASIC technology to accelerate the network will continue to be important factors in HCI deployments.

- **GPUs:** Graphics processing units have existed for years and been used to boost performance for VDI deployments, but there is renewed interest in them for planned workloads such as artificial intelligence and machine learning. GPUs have emerged as a key technology for accelerating compute-intensive AI/ML workloads. These processors can complete certain tasks tens of times faster than conventional X86 processors because they use hundreds of processing cores to run thousands of concurrent threads.
GPUs in HCI: AI/ML applications are mainly distributed applications such as research, training and inferencing run between clouds and on-premises. Hybrid clouds on HCI need to support the GPU performance density and scalability for these AI/ML use cases, as well as the AI/ML development frameworks.
- **FPGAs and ASICs:** FPGAs and ASICs are chips that also function as coprocessors. ASICs are application-specific and thereby provide added performance, but at the cost of reduced flexibility and versatility. Both FPGAs and ASICs have been used within storage arrays in the past to accelerate and offload compute-heavy tasks such as compression and encryption – and we expect to see these processors returning to the scene with next-generation system architectures.
FPGAs and ASICs in HCI: Every HCI software stack consumes core processing and memory resources to perform storage controller functions such as deduplication and compression, which can be more resource-intensive. Offloading software functions into dedicated FPGA and ASIC hardware allows for much higher compression and deduplication ratios, increasing the storage density delivered per node. By also minimizing the storage resource consumption, the processing and memory resources can be reallocated to application workloads, increasing virtual machine density on the HCI platform and, therefore, driving down TCO.
- **Next-Generation Interfaces and Storage-Class Memory:** New interface technologies (e.g., NVMe) and storage-class memory (e.g., NAND-based 3D XPoint) have emerged as new types of media that can offer speeds faster than typical flash while delivering the persistence that is lacking with DRAM. NVMe storage such as Intel and Micron's 3D XPoint non-volatile memory technology is a good example of this, and you can expect to see this technology being used to power a range of use cases, including in-memory databases and real-time analytics. While this emerging technology is starting to see traction in the server market, HCI solutions have been slower to adopt it due to its vastly different storage architecture. We expect to see greater volume of this technology in the next 12-18 months, led by HCI vendors with tight integration between hardware and software layers that can overcome the reliability, availability and serviceability challenges required of a production-ready solution.
NVMe and 3D XPoint in HCI: Time is money when it comes to application performance and the revelation of new business insights, so these technologies will accelerate workloads to new levels. SCM will allow HCI platforms to deliver high performance with low read and write latency, which will reduce the number of nodes required for a workload to create cost savings in facilities (rack space and power consumption), as well as in software fees (hypervisor, HCI and application licenses). High-density flash drives with NVMe connectivity will also improve the datacenter efficiency of HCI platforms by allowing them to store hundreds of terabytes of data in a small form factor.

Developing HCI Software for Hardware Innovations

HCI solutions generally reduce the issue of performance bottlenecks that arise from separate silos of compute, storage and networking, but they are not immune to bottlenecks caused by performance imbalances between the hardware and software layers. We expect to see more widespread adoption of hardware acceleration and other innovations within HCI systems to meet growing workload demands, but the increasing disaggregation between the hardware and software layers of HCI solutions means that actual performance gains can vary greatly among vendor options. This is vital to the TCO of HCI because less performance per node drives increased node counts, which in turn increases the expense from hardware, HCI software licensing, and operations associated with deploying and managing additional nodes.

Incorporating hardware innovations can drive HCI to new levels of usability as a hybrid cloud platform, but it's important to note that the use of these technologies in HCI goes well beyond simple hardware qualification. While vendors selling HCI as a software offering can simply qualify new hardware technologies on their software platform, the product delivered to the end user might not yield the expected returns if the HCI stack of hardware and software is not optimized to function as a whole. New storage interfaces such as NVMe, for example, require dramatically different designs from the server forming the hardware foundation of an HCI node, and key features like reliability, availability and serviceability can only be addressed through tight integration between the hardware and software layers. Dedupe and compression also stand out. While these key storage features can benefit greatly from the use of FPGAs and ASICs in an HCI solution, the software must be updated to yield the high compression and dedupe ratios that the hardware can deliver. Vendors that choose to sell preintegrated appliances have the potential to offer far more optimization between the software and hardware layers to get the most out of new hardware-based innovations.

Pulling the Pieces Together

As we have detailed extensively in this section, there is no shortage of technology innovation for enhancing and transforming infrastructures. While component-level innovations such as GPUs, SCMs, and ASICs and FPGAs can provide game-changing performance and efficiency gains in solutions such as HCI, the potential benefits of these technologies could be limited if they are not integrated tightly with the rest of the infrastructure.

At the system architecture level, we expect to see HCI vendors providing tighter hardware integration because their architectures will need to evolve substantially to keep up with the new hardware innovations. For example, in edge environments where traditional datacenters and IT expertise is lacking, system resiliency and performance will be even more crucial because organizations will be reluctant to send experts to distant locations to remediate issues. As a result, we expect to see more all-flash and SCM-enhanced systems with ruggedized designs emerging for edge environments since these provide a reliability advantage over hard-drive-based systems, which are more prone to failure. In core datacenters, superior system design will also be valuable to increasing compute and storage density, which has been a key advantage for hyperscalers that have been designing their own platforms for years now.

This need for integration will also be significant on management software, where some vendors are racing to add and optimize the AI/ML capabilities of their platforms to provide proactive support and optimization for their customers' systems. As hardware acceleration and other new technologies emerge, they must be monitored and integrated to ensure that when changes are made to various parts of the compute, networking and storage stack, operations will continue with minimal disruption.

Recommendations

Modernization Could Save You Money

Getting more out of your infrastructure by using acceleration technology is one strategy for minimizing the number of servers needed to accomplish the tasks or a given use case and will, therefore, drive down TCO, but modernization also drives simplicity, meaningful automation and, therefore, speed.

In our VotE: Servers and Converged Infrastructure, Budgets & Outlook 2018 survey, the top reason for deploying HCI was to simplify infrastructure management/maintenance (chosen by 79% of respondents), which is not surprising given that nearly all organizations are looking to reduce opex, and improving infrastructure management is strong area for HCI because it allows IT generalists such as virtualization administrators to manage resources without the help of expensive specialists such as storage administrators. The second-highest-rated reason for deploying HCI was to accelerate provisioning and/or optimization, which was chosen by 60% of respondents. The automation and advanced management tools that HCI providers are deploying will ensure stakeholders not only get resources in a timely fashion, but also that when issues arise, the optimization capabilities will make sure performance-sensitive applications are not starved of resources when they need them most.

As we discussed earlier, organizations also want to make their on-premises infrastructure more cloud-like, which is why 51% of respondents are deploying to support their hybrid cloud strategy.

Prepare for Challenges Ahead: IoT, Edge Computing, Pervasive Analytics

As IoT, edge computing and analytics become more pervasive, there is a greater need for compute and storage resources to be deployed farther from central datacenters. Our VotE: Internet of Things, Organizational Dynamics 2017 survey found that half of the data collected at the edge is being discarded. This indicates that there is a need for greater processing power at the edge to mitigate waste and capitalize on the value of the data. In addition to edge computing, emerging technologies such as AI/ML, big data and in-memory computing reinforce the need for IT modernization and the closing of any performance gaps. Our Digital Pulse: Budgets and Outlook 2017 survey found that 32% of respondents have ML/AI in use or are in the discovery/proof-of-concept phase; this number was 38% for big data and 16% for blockchain.

In all of the aforementioned use cases, the inclusion of hardware acceleration offers the potential to meet the growing performance needs while simultaneously achieving higher levels of efficiency. Moving forward, customers will be presented with several options when it comes to hardware accelerators that are best suited for a variety of these different processing needs. Although infrastructure professionals have been notoriously hesitant about moving away from the traditional architectures that they have built their careers on, the ever-increasing pressure to reduce expenses while accelerating service delivery will continue to force naysayers to evaluate and eventually leverage next-generation HCI systems for their datacenters.

About 451 Research

451 Research is a preeminent information technology research and advisory company. With a core focus on technology innovation and market disruption, we provide essential insight for leaders of the digital economy. More than 100 analysts and consultants deliver that insight via syndicated research, advisory services and live events to over 1,000 client organizations in North America, Europe and around the world. Founded in 2000 and headquartered in New York, 451 Research is a division of The 451 Group.

© 2019 451 Research, LLC and/or its Affiliates. All Rights Reserved. Reproduction and distribution of this publication, in whole or in part, in any form without prior written permission is forbidden. The terms of use regarding distribution, both internally and externally, shall be governed by the terms laid out in your Service Agreement with 451 Research and/or its Affiliates. The information contained herein has been obtained from sources believed to be reliable. 451 Research disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although 451 Research may discuss legal issues related to the information technology business, 451 Research does not provide legal advice or services and their research should not be construed or used as such.

451 Research shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The reader assumes sole responsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice.



NEW YORK

1411 Broadway
New York, NY 10018
+1 212 505 3030



SAN FRANCISCO

140 Geary Street
San Francisco, CA 94108
+1 415 989 1555



LONDON

Paxton House
30, Artillery Lane
London, E1 7LS, UK
+44 (0) 203 929 5700



BOSTON

75-101 Federal Street
Boston, MA 02110
+1 617 598 7200