IDC
Analyze the Future

Sponsored by:
Hewlett Packard Enterprise

Author:
Ashish Nadkarni

March 2017

Quantifying Datacenter Inefficiency: Making the Case for Composable Infrastructure

Nearly all datacenters in enterprise IT have areas where efficiency can be improved. IDC study shows:

Median efficiency in enterprise IT:

PEOPLE
50%

PROCESS
30%

TECHNOLOGY
50%

EXECUTIVE SUMMARY

Digital transformation (DX) — a technology-based business strategy — is an essential mandate for businesses to thrive in a digital economy. Businesses embark on DX initiatives to create value and extend their competitive advantage through new products and services, new business relationships, improved customer experience, and increased operational efficiencies. Next-generation applications (NGAs) and technologies (collectively referred to as IDC's 3rd Platform technologies) form the essential underpinning for DX. In the near term, these next-generation applications will coexist with traditional or current-generation applications (CGAs), as the latter support day-to-day revenue-generating business operations. That means in the immediate future, IT departments have to manage an infrastructure "duality" — the ability to deploy and manage two sets of applications, each with vastly diverse infrastructure requirements and service-level objectives.

IDC recently conducted a study on datacenter infrastructure and operations efficiency among medium-sized and large enterprises. The study found that nearly all datacenters in enterprise IT have areas where efficiency can be improved. People, process, and technology resources are poorly utilized, causing unnecessary churn and wastage. IDC found that by implementing next-generation apps on infrastructure designed for current-generation apps, such businesses risk putting their DX initiatives in jeopardy. An immediate effect of such an implementation is that it further exacerbates existing datacenter technology and operations inefficiencies. Furthermore, the operations paradigm demanded by NGAs puts undue pressure on people resources and can lead to an unsustainable CGA-based operating model. For example, the amount of time it takes to provision virtual machines (VMs), which is typically in hours, is generally unacceptable in DevOps environments.

IDC
Analyze the Future

Cohosting NGAs with CGAs requires businesses to embrace a new infrastructure model. Composable infrastructure solutions are designed to inherently handle the current- and next-generation infrastructure duality.

Cohosting NGAs with CGAs requires businesses to embrace a new infrastructure model. Composable infrastructure solutions are designed to inherently handle the current- and next-generation infrastructure duality. Solutions such as Hewlett Packard Enterprise (HPE) Synergy take key benefits of converged and hyperconverged infrastructure to the next level by operating on fluid (composable) compute, networking, and storage resource pools and are optimized for operations as well as applications. Such infrastructure solutions help businesses realize the vision of "infrastructure as code" — an operating paradigm that is specifically meant to support the "idea economy."

Methodology and Key Findings

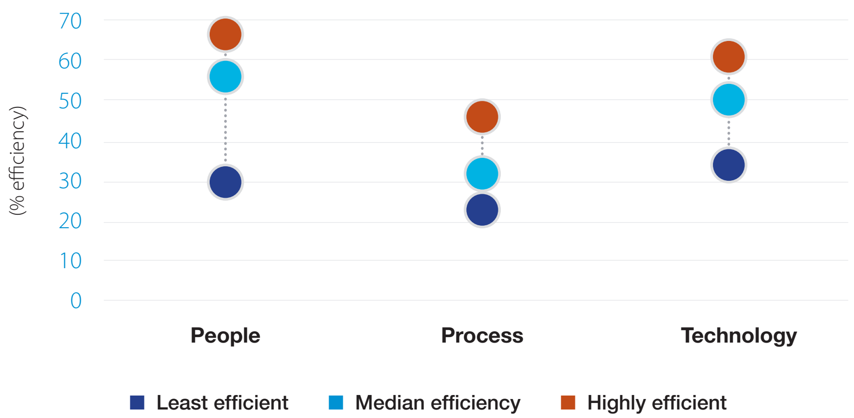
This white paper discusses the findings of a study commissioned by Hewlett Packard Enterprise, in which IDC surveyed 301 IT users from medium-sized and large enterprises (as defined in terms of the number of employees or number of servers). The survey sought to quantify insight into current infrastructure adoption and usage trends and ascertain potential capital expense (capex) and operational expense (opex) wastage because of common infrastructure provisioning and operations practices. To qualify, respondents had to be intimately familiar with their infrastructure, which had to be on-premises or a mix of on-premises and off-premises and be managed by internal staff.

The study found that (findings illustrated in Figure 1):

- **Median people efficiency in enterprise IT is 55%.** This can be attributed to how people spend their time on routine operations tasks. Repetitive provisioning tasks add to this inefficiency because many of them are performed manually (a sampling of routine tasks and potential wastage is illustrated in Figure 20).
- **Median process efficiency in enterprise IT is 30%.** This is mostly due to the number of steps needed to complete provisioning tasks and the lack of automation. Additional paperwork requirements add to this inefficiency.
- **Median technology efficiency in enterprise IT is 50%.** This is due to a combination of overprovisioning for redundancy and resiliency and ongoing underutilization from a capacity and performance perspective, leading to a lot of idle hours during which the infrastructure is not servicing any applications.

FIGURE 1

People, Process, and Technology Efficiency in Enterprises



n = 301

Notes: Figure is based on the analysis of survey data. Data is weighted by company size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

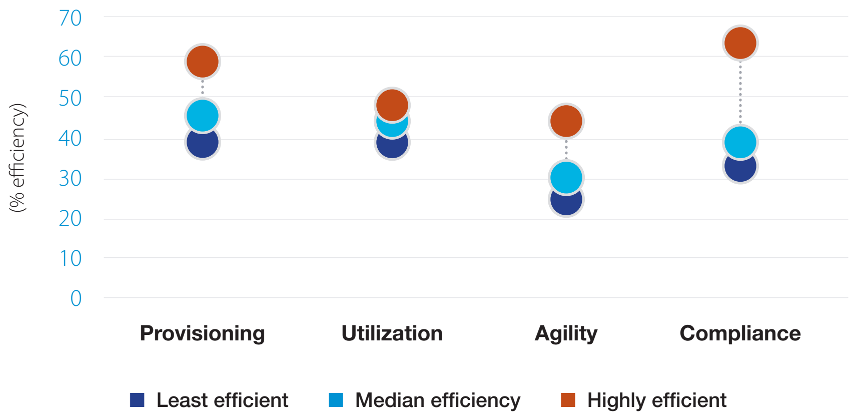
The study also found that in most enterprise datacenters, infrastructure is (findings illustrated in Figure 2):

- **45% provisioned:** This is measured in terms of the total capacity available in the infrastructure versus what is made available to the physical or virtual machines running applications.
- **45% utilized:** This is measured in terms of the amount of idle compute hours and unused storage capacity for provisioned components.
- **30% agile:** This is measured in terms of the perceived inefficacies in the time it takes to provision new compute instances and storage instances or make changes to existing instances. Making changes includes adding or removing capacity, the latter being especially time consuming.
- **40% compliant:** This is measured in terms of the perceived confidence in the ability to meet stated service-level agreements (SLAs) for infrastructure and its components, such as requisite redundancy and resiliency. To protect against human errors and/or component failures, businesses tend to overprovision infrastructure, leading to unnecessary wastage without eliminating the risk.

Additional survey data can be found in the Appendix.

FIGURE 2

Datacenter Efficiency in Enterprises



n = 301

Notes: Figure is based on the analysis of survey data. Data is weighted by company size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

Situation Overview

Thriving in the Digital Economy

Businesses have a clear mandate: Thrive in the digital economy — that is, embrace a disruptive technology-centric business strategy (aka digital transformation, as illustrated in Figure 3) — or perish. DX is thus a high-stakes game — DX initiatives are often net new and highly outcome dependent and thrust businesses into uncharted territory:

DX is people-process-technology transformation, not just technology transformation. Unlike typical incremental initiatives that are one of the three, DX requires all three to transform in parallel.

- DX requires that businesses initially complement and eventually replace old business models with new business models built from the ground up, without impacting bread-and-butter revenue.
- DX initiatives are often led by lines of business (LOBs) that have little to no operational experience with IT infrastructure.
- DX is people-process-technology transformation, not just technology transformation. Unlike typical incremental initiatives that are one of the three, DX requires all three to transform in parallel.

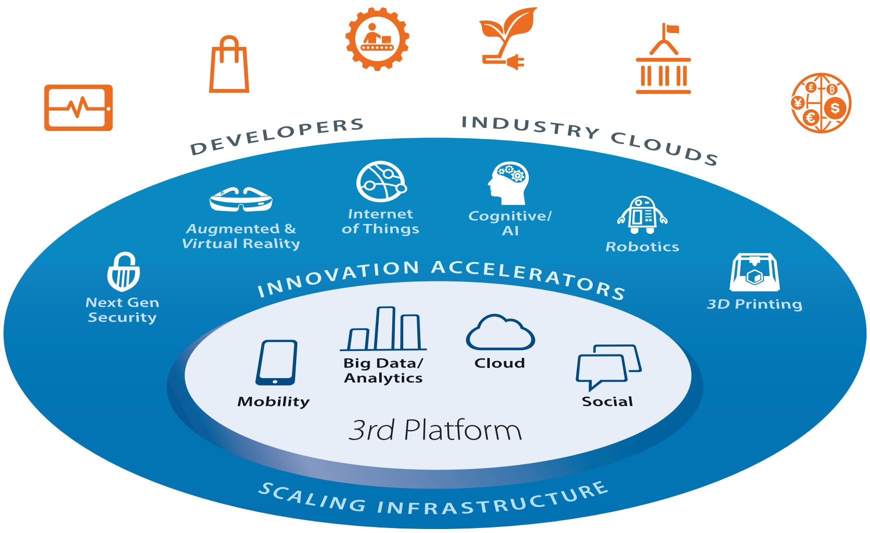
Next-generation applications and technologies form the essential underpinning for DX from a technology perspective.

Successful DX is rewarding: It enables businesses to create value and extend their competitive advantage through new products and services, new business relationships, improved customer experience, and increased operational efficiencies.

Next-generation applications and technologies form the essential underpinning for DX from a technology perspective. For businesses to effectively embrace 3rd Platform technologies, they need to ensure that people and processes are in lockstep with technology. That is where many DX initiatives hit hurdles, but the obstacles can be avoided with careful planning.

FIGURE 3

Digital Transformation



Source: IDC, 2017

The Infrastructure Duality: Advent of Next-Gen Apps

Taking on next-generation applications means getting accustomed to an entirely new people-process-technology paradigm.

In most businesses, IT departments are used to managing current-generation applications, which support day-to-day revenue-generating business operations and run on traditional IT infrastructure that they know well. Taking on next-generation applications means getting accustomed to an entirely new people-process-technology paradigm:

» CGAs assume infrastructure resiliency and require computing technologies such as virtualization and clustering that provide portability and transparently preserve application state. They also depend on external shared storage arrays that provide multiple layers of resiliency and redundancy. CGAs include enterprise resource planning (ERP), customer relationship management (CRM), and other enterprise applications that utilize relational databases.

- NGAs are often custom developed and deployed using newer methodologies such as continuous deployment/continuous integration (CD/CI) and DevOps. Unlike CGAs, NGAs require that IT operations and developers work as a team to continuously release updates to the application to accomplish a quicker release cadence. In contrast, most CGAs have sluggish upgrade cycles that are driven by vendor patches, requiring more planning ahead of time. Examples of continuously updated NGAs are mobile, online gaming, IaaS/SaaS, IoT applications, and big data.
- Unlike CGAs, NGAs are designed to be horizontally scalable, and with data layer processes that do not assume infrastructure resilience. Such applications are designed to run on industry-standard computing platforms and inside lightweight application containers and carry no bias toward certain silicon architectures. This means that traditional infrastructure that runs CGAs is not optimized for NGAs.

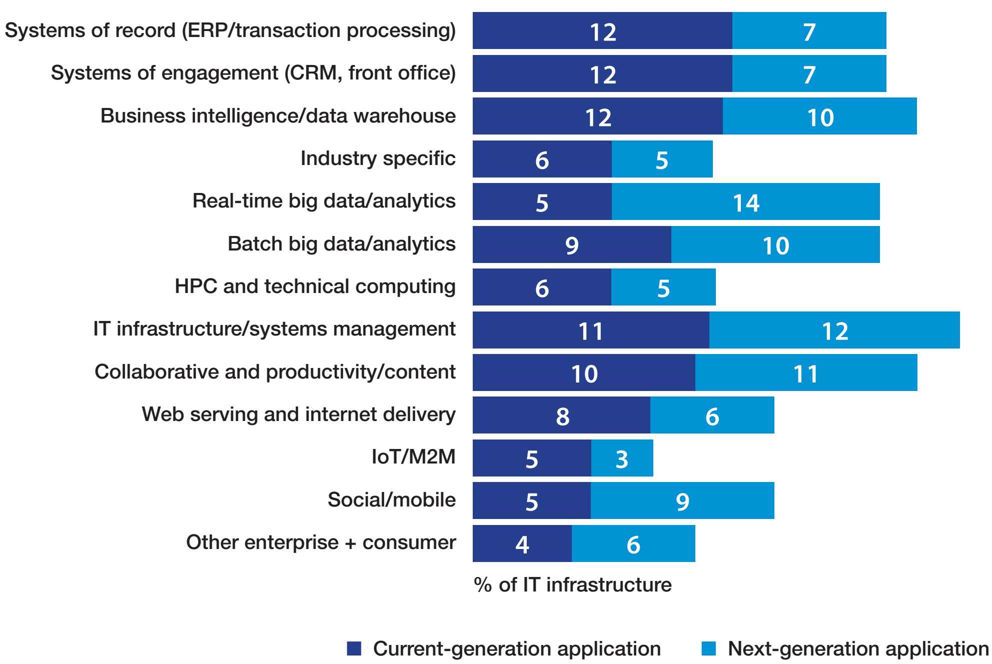
With DX, IT departments have to embrace an "infrastructure duality," at least in the short to medium term. In other words, they need to manage the infrastructure for NGAs alongside that for CGAs.

With DX, IT departments have to embrace an "infrastructure duality," at least in the short to medium term. In other words, they need to manage the infrastructure for NGAs alongside that for CGAs.

Figure 4 illustrates the general spread of CGAs and NGAs in different types of workloads.

FIGURE 4

Distribution of Current- and Next-Generation Applications in a Typical Enterprise



n = 619

Notes: Data is not weighted. Multiple responses were allowed. Use caution when interpreting small sample sizes.

Source: IDC's Next-Generation Applications Survey, *2016*

In most IT departments, capex and opex wastage is taken for granted — much of it is considered to be necessary to overdeliver on SLAs.

Examining the State of IT Infrastructure and Operations

In most IT departments, capex and opex wastage is taken for granted — much of it is considered to be necessary to overdeliver on SLAs. In the aforementioned study, IDC found that:

» Overprovisioning and idle infrastructure resources contribute to capex wastage. Infrastructure copies for disaster recovery (DR), test, and development exacerbate this problem further.

» Time spent on provisioning and maintenance activities contributes to opex wastage. This wastage is amplified by the number of applications.

» Many businesses have moved to converged infrastructure/hyperconverged infrastructure (CI/HCI) to streamline provisioning and operations. Such infrastructure is still largely optimized for CGAs.

Infrastructure

Table 1 and Figures 12–18 (see the Appendix) provide a snapshot of a typical IT infrastructure. Key highlights are:

» **Most environments are virtualized.** Table 1 illustrates that virtualization density (in terms of the total number of both servers and VMs per server) is still a bit underwhelming, resulting in heavy capex (spent on provisioning and managing physical servers). This disproves a common assumption of the pervasiveness of virtualization in the datacenter. In fact, it shows that most physical servers aren't virtualized at all, and this might be where composable infrastructure has the greatest opportunity to provide benefits — by allowing the operating environment for those bare-metal workloads to be manipulated more easily. Businesses have generally embraced multihypervisor environments, though the preference is still to standardize on a vSphere or Hyper-V for production.

» **Containers are slowly but surely gaining traction.** Docker is the preferred container format, though CoreOS is gaining popularity.

» **Much of the infrastructure is still on-premises and managed by internal IT staff.** This infrastructure is considerably traditional, but a relatively small number of organizations have a private cloud deployed. This is noteworthy because such environments will struggle with NGAs — which require much of the provisioning and operations processes to be automated.

- **There is generally an even and consistent distribution of compute and storage resources across the various types of workloads.** However, workloads like content/collaboration, IT management systems, and virtual client infrastructure consume a higher portion of the overprovisioning and idle resource time. These workloads also have the highest year-over-year (YoY) growth, leading to the conclusion that much of the overprovisioning is anticipatory in nature.

In summary, there are plenty of opportunities to reduce inefficiencies in the infrastructure and streamline the components to make it easier to handle the CGA-NGA duality.

In summary, there are plenty of opportunities to reduce inefficiencies in the infrastructure and streamline the components to make it easier to handle the CGA-NGA duality.

TABLE 1

Density of Compute Infrastructure

Q. When thinking about the server infrastructure supporting your organization (worldwide), please provide the following.

	Average (n = 301)	50 to 500 Servers Servers (n = 209)	More than 500 Servers (n = 92)
Total number of physical servers and blades	386	100	711
Percentage of physical servers hosting virtual servers (%)	23	21	25
Total number of virtual servers	1,061	261	1,971
Approximate virtual machine density (number of VMs per physical server)	13	13	12

n = 301

Base = all respondents

Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

People and Process

Figures 19–21 (see the Appendix) provide a snapshot of typical IT operations. Key highlights are:

- More people resources are allocated to carry out provisioning, configuration, and deployment functions. This is followed by resources allocated to operations and ongoing management.

- From an operations perspective, people in general spend more time provisioning or decommissioning (physical or virtual) server instances, performing troubleshooting, and recovering unresponsive application instances.

- From a process perspective, much of the time is spent handling new service requests, followed by software installation requests and patching and upgrading servers.

This is not surprising considering that much of IT infrastructure is run in a traditional manner. Furthermore, lack of automated processes in IT infrastructure means that routine tasks have to be carried out manually, creating inefficiencies in how people allocate their time.

Making the Case for Composable Infrastructure

Key Technical and Business Requirements for Infrastructure

Figures 22–29 (see the Appendix) provide a snapshot of what IT organizations look for when procuring new infrastructure solutions. Key highlights are:

- From a compute perspective, there is a preference for quad-core or higher-end processors, followed by shared compute (virtualized) and then single-core processors. In addition, businesses are increasingly adopting GPUs for accelerating certain workloads. From a storage perspective, there is a preference for internal storage drives — for both boot and local data — followed by internal flash and external hybrid storage arrays. Then there is a preference for external all-flash arrays to cater to low-latency, high IOPS workloads. These preferences are similar for mission- and non-mission-critical workloads, indicating that businesses seek to place applications on the appropriate service tier based on their performance requirements. This directly influences capex.
- With regard to infrastructure optimization, there is a strong preference to adopt infrastructure solutions that streamline and potentially reduce maintenance (such as patching and firmware upgrades), remediation (such as monitoring and troubleshooting), and service requests (such as server, storage, and network provisioning) tasks. This directly impacts opex.
- Businesses are sold on the benefits of private cloud and will invest in infrastructure that enables a move to deploying a private cloud. The top 3 requirements are the ability to offer multiprotocol (file, block, and object) access, dedicated image storage, and flexibility to switch between performance-optimized and capacity-optimized storage. These are followed by dedicated fabric, stateless computing, and unified API. This impacts both capex and opex.

Composable infrastructure (aka composable/disaggregated infrastructure) is an emerging category of datacenter infrastructure that seeks to (dis)aggregate compute, storage, and networking fabric resources into shared resource pools that can be available for on-demand allocation (i.e., "composable").

In other words, composable infrastructure makes it easier for businesses to embrace infrastructure duality.

- The top 3 criteria for selecting infrastructure components are total cost of operations, efficient ongoing management, and the ability to efficiently provision and deprovision resources. When these criteria are coupled with the top 2 business outcomes — which are mitigate risk to the business and reduce capex — it is clear that IT organizations are highly focused on servicing their internal clients with a cost-efficient and streamlined infrastructure.
- To service their internal clients, businesses are not shy about switching vendors. Vendor consolidation, infrastructure consolidation (by way of adopting converged or hyperconverged infrastructure), lower capex, and better service and support are cited as top reasons why businesses will consider such a switch.
- Businesses are examining alternatives to an on-premises infrastructure — alternatives such as public cloud (software as a service, infrastructure as a service, and platform as a service) are evaluated alongside traditional IT and converged/hyperconverged infrastructure. To a lesser extent, businesses also examine composable infrastructure, provided the solution meets their requirements.

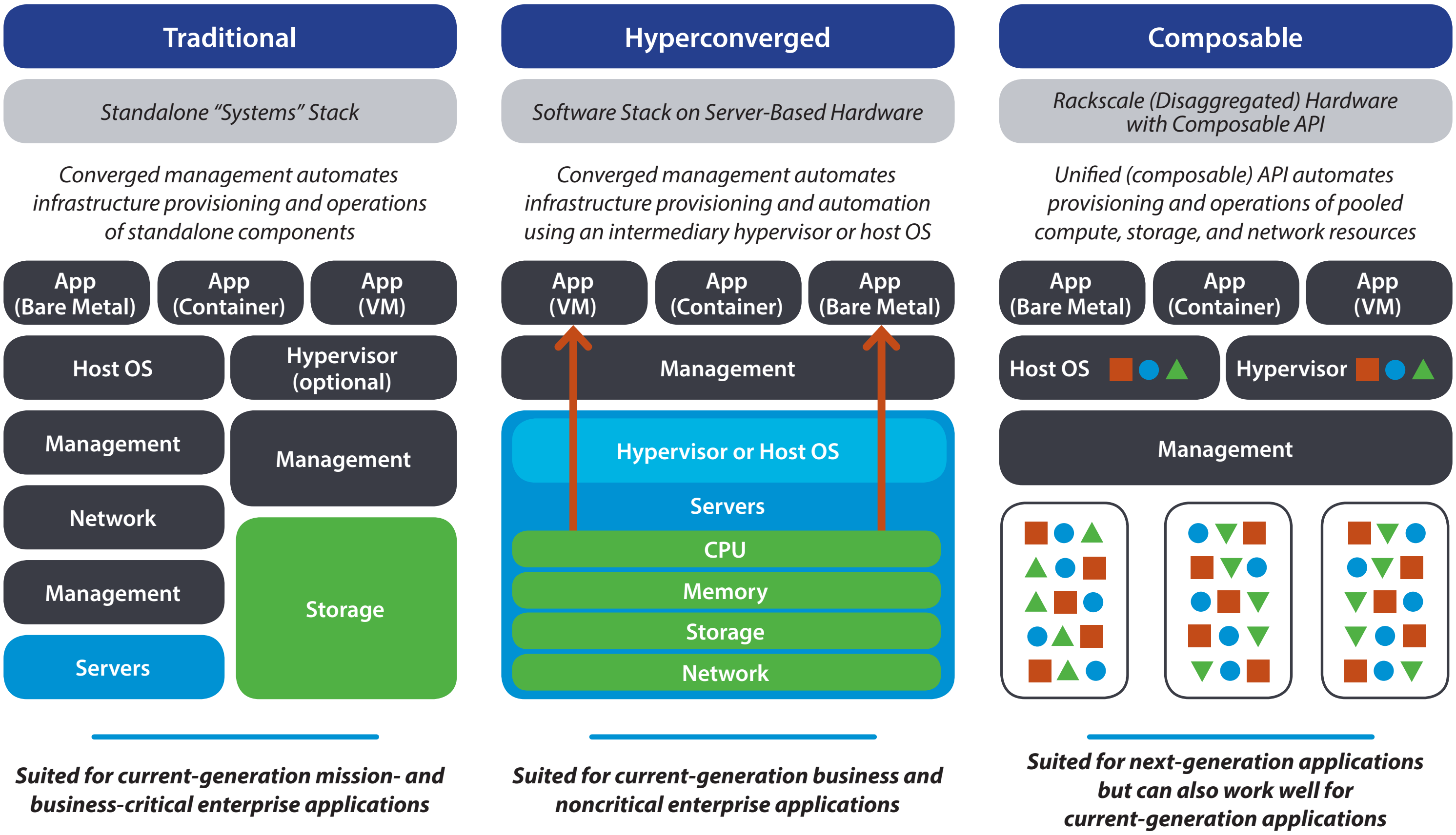
What Is Composable Infrastructure?

Composable infrastructure (aka composable/disaggregated infrastructure) is an emerging category of datacenter infrastructure that seeks to (dis)aggregate compute, storage, and networking fabric resources into shared resource pools that can be available for on-demand allocation (i.e., "composable"). Systems such as Hewlett Packard Enterprise Synergy are full-stack systems comprising the disaggregated hardware platform and a suite of management software that allows rapid provisioning and seamless ongoing management of resources. Such systems are targeted at IT organizations that require a common (read: flexible and agile) infrastructure to host NGAs and CGAs — whose design principles and therefore operational characteristics are fairly divergent. In other words, composable infrastructure makes it easier for businesses to embrace infrastructure duality. Highlights of composable infrastructure are:

- "Composability" is driven at the software (API) level, and "disaggregation" is primarily driven at the hardware level. To fully implement the design principles of composable infrastructure, the hardware it operates on has to support partial or full disaggregation (in which resources are pooled down to a component level).
- Composable systems are an evolution of converged and hyperconverged infrastructure (see Figure 5). While the new technology is a significant leap forward, the essence of this evolution is that the hardware side is moving toward disaggregation while the software side is moving toward composability — via a unified API-based provisioning, orchestration, and automation layer.

FIGURE 5

Composable Infrastructure as the Next Phase of Converged/Hyperconverged Infrastructure



Source: IDC, 2017

Composable Infrastructure Makes It Easier to Embrace Infrastructure Duality

Businesses in the midst of DX ought to treat infrastructure as a utility and, therefore, its ongoing deployment and management as an on-demand activity that can be carried out by those who use it. This means supporting users who develop CGAs as well as those who develop NGAs.

CGAs require workload-specific infrastructure environments. Applications are often run on bare metal or on virtual machines and are managed by dedicated IT resources. On-premises applications like Microsoft Exchange and ERP systems are two examples of applications that could be considered CGAs.

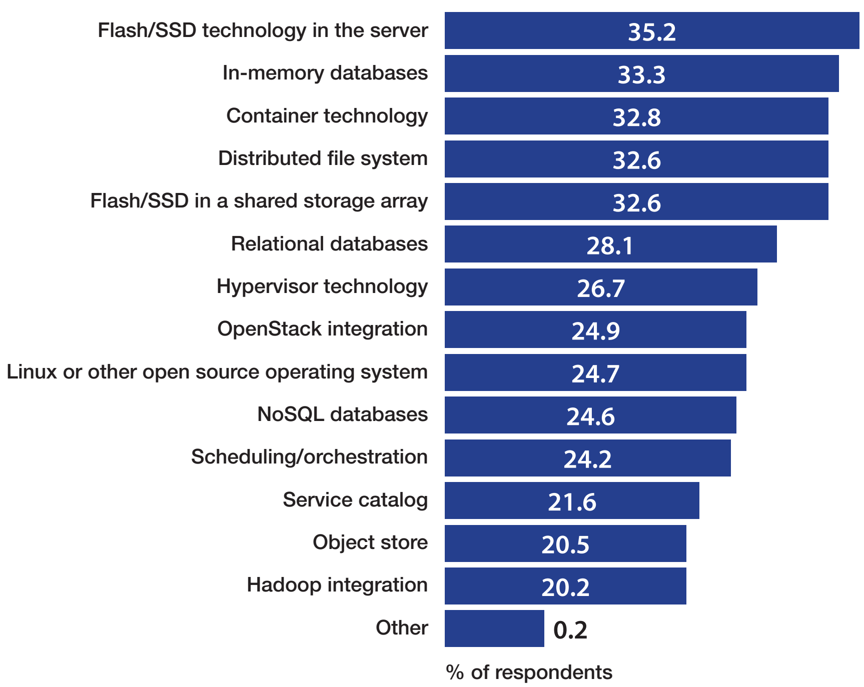
Figure 6 illustrates key infrastructure requirements for NGAs. The top requirements are support for containers, server-side flash/SSD, in-memory databases, and distributed file systems. NGAs are often developed via methodologies such as DevOps. Methodologies such as DevOps rest on collaboration as the cultural norm. No longer are development and operations separated by different chains of command. DevOps works well when the underlying infrastructure is "treated as code." As illustrated in Figure 7, infrastructure as code (also known as software-defined infrastructure) consists of four main principles: platforms, value, automation, and policy driven. Composable infrastructure enables the infrastructure-as-code paradigm and serves the crucial role in fulfilling all four principles stated previously.

Infrastructure as code (also known as software-defined infrastructure) consists of four main principles: platforms, value, automation, and policy driven. Composable infrastructure enables the infrastructure-as-code paradigm and serves the crucial role in fulfilling all four principles stated previously.

FIGURE 6

Mandatory Components of NGA Infrastructure

Q. Which, if any, of the following are mandatory infrastructure components for next-generation applications?



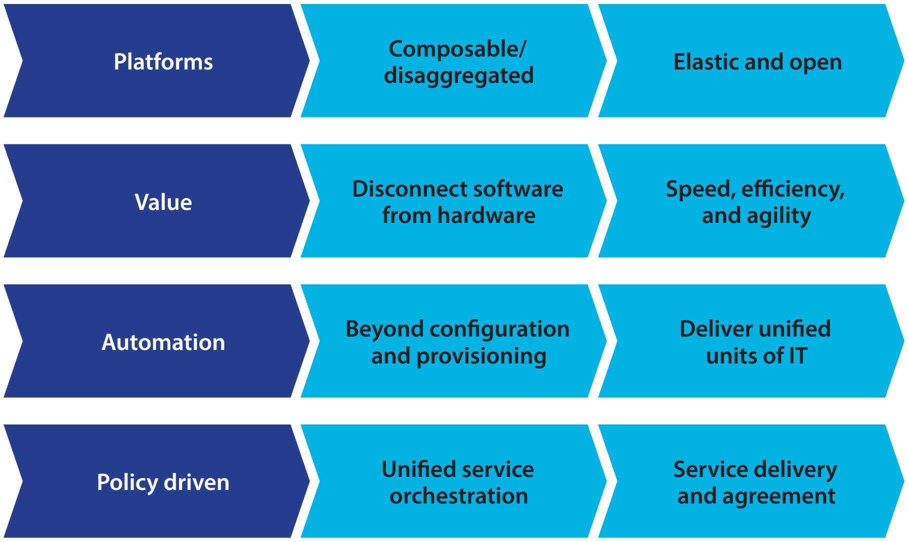
n = 619

Notes: Data is not weighted. Multiple responses were allowed. Use caution when interpreting small sample sizes.

Source: IDC's Next-Generation Applications Survey, *2016*

FIGURE 7

Infrastructure as Code



Source: IDC, 2017

Collectively, these requirements point to IT organizations seeking to leverage composable infrastructure as the vehicle to accelerate DX.

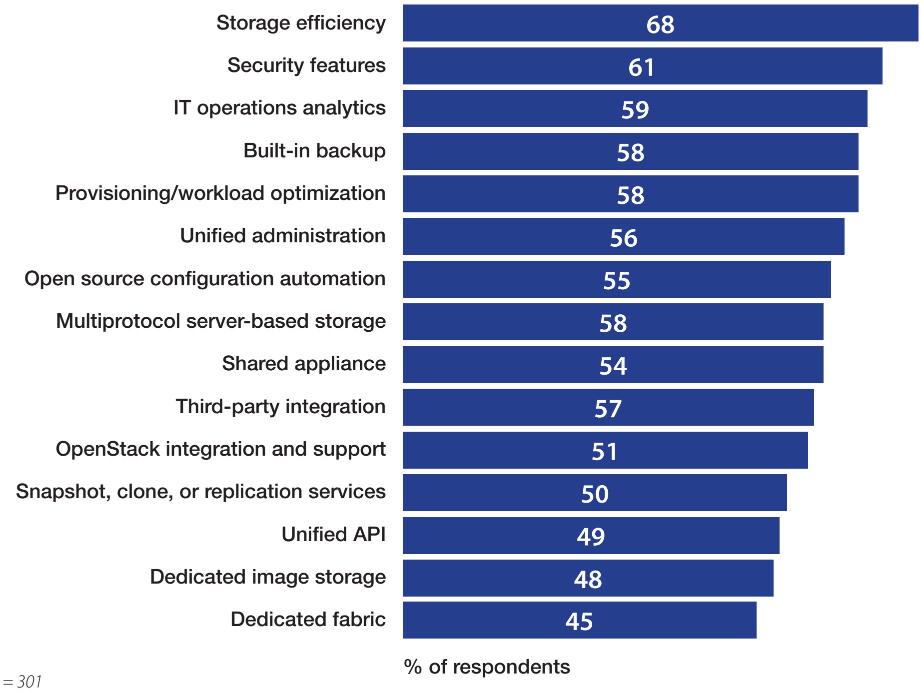
Figures 8 and 9 illustrate the key requirements and desired benefits of composable infrastructure.

- Top requirements are storage efficiency, security features, IT operations analytics, built-in backup, and provisioning/workload optimization. These are followed by unified administration, open source configuration automation (via tools like Puppet, Chef, and Ansible), and multiprotocol server-based storage.
- Top desired benefits of composable infrastructure are improved IT staff productivity, improved utilization of compute resources, and improved business agility. For businesses to accelerate DX, they have to embrace infrastructure duality, and that means a case for composable infrastructure, which addresses these requirements along with those for CGAs. Collectively, these requirements point to IT organizations seeking to leverage composable infrastructure as the vehicle to accelerate DX.

FIGURE 8

Key Requirements of Composable Infrastructure

Q. Which of the following management capabilities do you believe must be provided as part of the management software included with the composable infrastructure?



n = 301

Base = all respondents

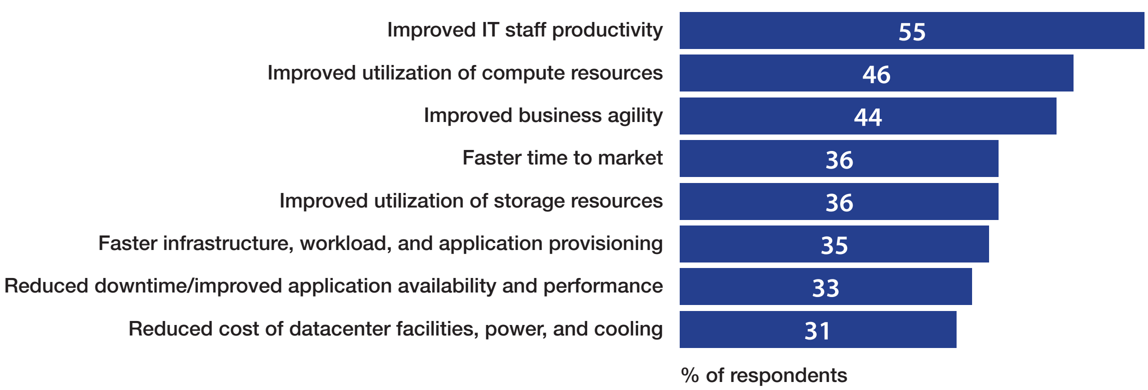
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 9

Top Desired Benefits of Composable Infrastructure

Q. Which of the following are your top priorities for benefits from the use of composable systems?



n = 301

Base = all respondents

Note: Data is weighted by employee size

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

Synergy is designed to serve both as a bridge for businesses that are in the thick of CGAs but are rapidly transitioning to supporting NGA environments and as the system of choice for businesses designing an infrastructure for NGAs from the ground up.

HPE Synergy: The First Instantiation in a New Class of Infrastructure

In late 2015, HPE announced a new class of infrastructure known as Synergy, which is built to further the company's vision of a fully composable datacenter infrastructure. Synergy is designed to serve both as a bridge for businesses that are in the thick of CGAs but are rapidly transitioning to supporting NGA environments and as the system of choice for businesses designing an infrastructure for NGAs from the ground up. From that side, it serves as the best of both worlds — for applications with infrastructure resiliency requirements and for applications that don't assume infrastructure resiliency. The three key foundational design elements of HPE Synergy that enable IT organizations to deploy NGAs at "cloud speed" or take a more traditional approach to deploy highly resilient infrastructure for CGAs are:

» **Unified APIs** allow organizations to implement infrastructure as code — a single line of code that abstracts all infrastructure. This enables developers to integrate infrastructure provisioning commands directly into the application development process, which in turn allows for a quicker deployment of applications.

» **Software-defined intelligence** enables organizations to take a template-driven approach to workload composition and management. This minimizes hiccups caused by operational activities such as operating system patching and firmware upgrades, thereby improving IT efficiency and reducing opex.

» **Fluid resource pools** provide the ability to compose pseudo-physical, virtual, and containerized computing units from compute (processor and memory) and data persistence (disk and flash) resource pools. This reduces the wastage and overhead caused by overprovisioning of resources and thus provides a means to capex.

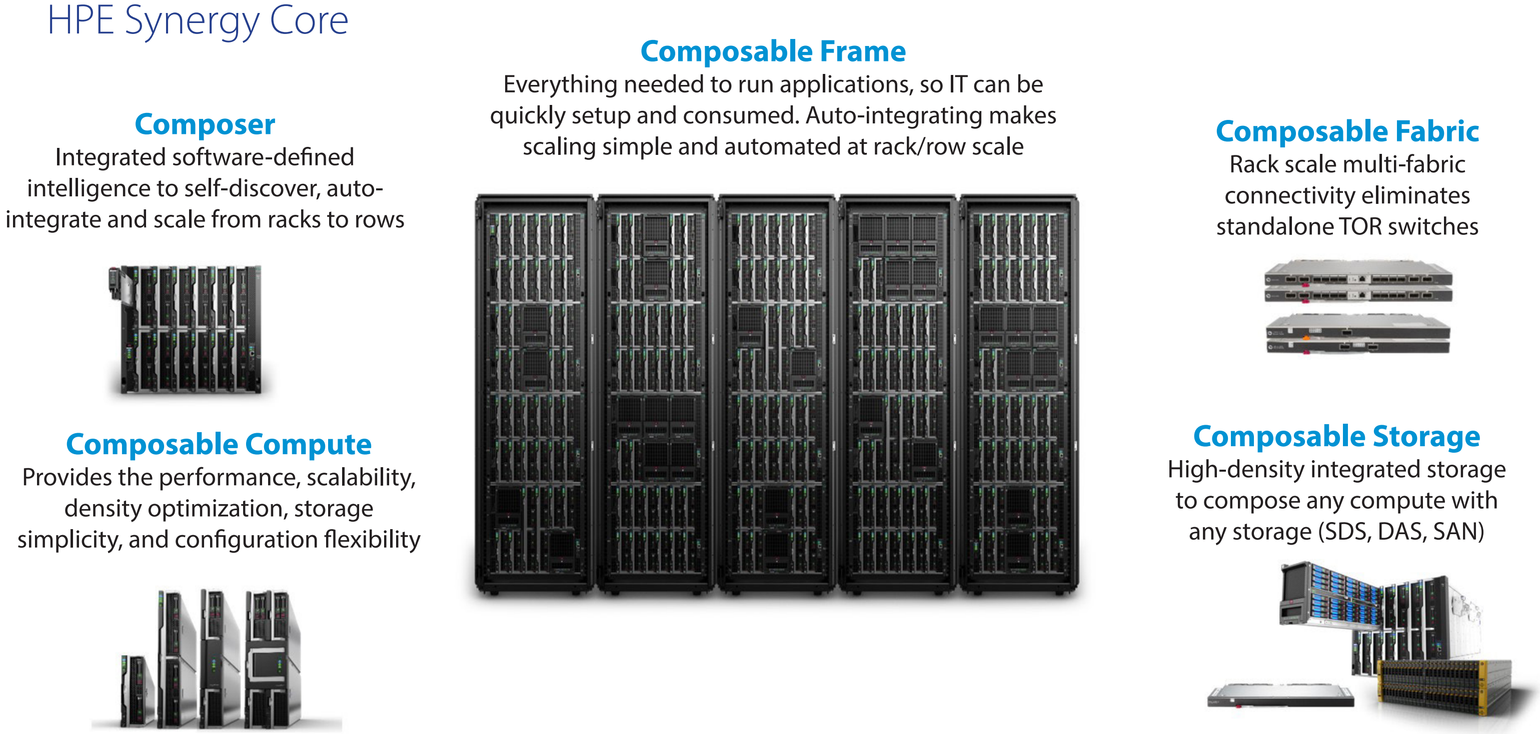
As Figure 10 illustrates, the main components of the HPE Synergy system are as follows:

» **Composable frame (includes compute, fabric, storage, and management modules):** The frame houses compute and storage modules. Compute modules provide CPU and memory capacity, while storage modules contain shareable DAS drives and flash. Each node contains up to 200 drives. Compute modules provide processor and memory capacity, while storage modules contain shareable DAS drives and flash. Each node contains up to 200 drives. Applications like Hadoop and email that require shared nothing capacity and compute dedicated for storage or software-defined storage (SDS) software like StoreVirtual VSA can utilize these drives for data persistence. Fabric modules allow frame expansion and external connectivity. The fabric modules can uplink to the datacenter directly — 10G or 40G or Fibre Channel. Internally, it runs on a shared PCIe fabric and can run across multiple frames.

- **Composable software suite:** Two key elements of this software stack are HPE Synergy Composer and HPE Synergy Image Streamer. **HPE Synergy Composer**, powered by HPE OneView, enables IT to deploy, monitor, update, and manage infrastructure through its life cycle from one interface. As the window into HPE Synergy, Composer provisions compute, storage, and fabric resources using a template process. It is the provisioning engine that allows resources from the frame to be carved into pseudo-physical compute and persistence units. This allows infrastructure to be deployed and consistently updated with the right configuration parameters and firmware versions — streamlining the delivery of IT services. **HPE Synergy Image Streamer** is the software/appliance (or module) that allows the operating platform to be provisioned onto these pseudo-physical units. Deploying a typical operating system or hypervisor can be time consuming because it requires customizing and/or copying each image for each compute module. HPE Synergy Image Streamer accelerates these processes through tight integration with HPE Synergy Composer templates. The image payload consists of the compute module's profile (from HPE Synergy Composer), the golden image (the operating environment with a bootable operating system and application and the I/O driver version), and the personality (the operating system and application configuration). Through the template, bootable images are deployed across stateless compute modules using a simple, consistent process.

FIGURE 10

HPE Synergy Core



Source: HPE, 2017

HPE Synergy also supports remote object, file, or SAN storage to be connected to the frame. In the case where the external storage supports the composable API (such as HPE 3PAR StoreServ arrays), the provisioning domain can be extended to include such storage (currently 3Par StoreServ arrays are supported). In theory, any hardware could be disaggregated, but reality is much different. This is where the HPE Synergy design is truly different. From a data perspective, it is stateless (i.e., the operating system images can be transferred anywhere in the frame). From a configuration state perspective, it is dynamic (i.e., such images can be altered on the fly).

HPE Synergy as the Building Block of a Future-Proof Infrastructure

A future-proof infrastructure is designed with components that can be upgraded continuously and without the infrastructure undergoing any major architectural change.

A future-proof infrastructure is designed with components that can be upgraded continuously and without the infrastructure undergoing any major architectural change. HPE Synergy helps organizations achieve important objectives related to infrastructure, objectives that are difficult — if not impossible — to achieve via traditional, converged, or even hyperconverged infrastructure solutions. For example, the ability to:

- Deploy CGAs or NGAs quickly with simple flexing, scaling, and updating
- Run workloads anywhere — on physical servers, on virtual servers, or in containers
- Operate any workload upon which the business depends, without worrying about infrastructure resources compatibility
- Ensure the infrastructure is able to provide the right service levels so the business can stay in business

As a next-generation datacenter architectural option, HPE Synergy embraces and extends key concepts and traits from the architectures that have come before it, including converged and hyperconverged systems.

As a next-generation datacenter architectural option, HPE Synergy embraces and extends key concepts and traits from the architectures that have come before it, including converged and hyperconverged systems. Simply trying to "bolt" Synergy onto one of these less inclusive architectures wouldn't have resulted in a complete, tightly integrated solution. With that in mind, HPE has developed the first platform architected for composability focused solely on helping businesses achieve the key outcomes required for a future-proof infrastructure.

Furthermore, by deploying Synergy, businesses can see a marked improvement in IT efficiency.

- People efficiency goes up because resources don't have to focus on routine operations tasks. Repetitive provisioning tasks are automated, freeing them to focus on strategic initiatives.
- Process efficiency improves as steps needed to complete provisioning tasks are automated and paperwork requirements are codified as templates. This process efficiency also includes audit capabilities, improving the overall confidence in the ability to meet stated service-level agreements for infrastructure and its components compliance.

- Technology efficiency goes up as the infrastructure is rightsized and inherent inefficiencies are removed. IT administrators can seek to turn off systems that are not utilized for long periods of time. They no longer need to overprovision resources for burst capacity — such capacity can be added when necessary.
- The infrastructure also gains agility because resource provisioning — regardless of whether it is for CGA or NGA environments — is now fast and on demand. Provisioning agility to create new compute instances and storage instances or to make changes to existing instances is now fully automated.

Future Outlook

As businesses move through their DX journey, the people-process-technology transformation will result in a shift in the role of IT. IT will become part of the fabric of products and services that are rapidly innovated at every company. IT will play a pivotal role in the quest for the business to remain competitive. CIOs and line-of-business executives will be at the forefront of a major initiatives aimed at leveraging the competitive advantages of the new hyperconnected enterprise. In other words, IT will gain the ability to influence new business opportunities by quickly delivering revenue-generating products, services, and experiences. Instead of just providing technology to automate internal business processes, IT can now directly impact business strategy and revenue by creating software-based services that:

- Energize growth
- Strengthen profitability
- Boost productivity
- Enhance innovation
- Increase organizational agility
- Improve the customer experience
- Provide a competitive advantage

Infrastructure as code — which at the moment is fairly nascent in terms of real-life implementations, especially in the enterprise — will truly be the overarching management paradigm. Composable infrastructure will form the core of any IT in that paradigm.

Infrastructure as code — which at the moment is fairly nascent in terms of real-life implementations, especially in the enterprise — will truly be the overarching management paradigm. Composable infrastructure will form the core of any IT in that paradigm.

Challenges/Opportunities

From a systems perspective, any composable/disaggregated architecture is made up of two parts: The first part is the ability to (dis)aggregate IT resources into compute, storage, and fabric pools, and the second part is the ability to compose consumable resources from such disaggregated pools via a unified API. Intelligent software therefore is needed to manage all the distinct assets and to compose the optimal configuration for a specific application. All the elements inside the installed infrastructure are pooled and require management, including:

» **Discovery and location of resources.** Monitoring and life-cycle management software are necessary to provide full awareness of the hardware assets and application workloads.

» **Self-discovery, provisioning workloads, orchestration, and healing.** It will be essential to have visibility into the utilization of the discrete resources and understand the load on the resource elements.

» **Orchestration layer.** It will be essential to enable a catalog of compute, storage, networking, and memory in an orchestration layer and also define resource requirements for specific applications and compose them in a set that is optimized for the workload.

Composable systems necessitate infrastructure analytics (i.e., data metrics are needed to understand how to optimally configure hardware for the applications). It is essential to have visibility into the utilization of the discrete resources and understand the load on the resource elements; for example, to answer questions like:

» How much is the application consuming?

» Which resources (such as processor and storage) are reaching thresholds or are underutilized?

These data metrics are also valuable feedback to enable orchestration and template provisioning. The dynamic nature of application capacity means manual processes, which consume time and are a source of potential errors, must be reduced as much as possible.

IDC sees these as necessary factors for composable systems to be successful in the enterprise — in accelerating provisioning times, improving IT utilization, and simplifying overall IT operations.

IDC sees these as necessary factors for composable systems to be successful in the enterprise — in accelerating provisioning times, improving IT utilization, and simplifying overall IT operations.

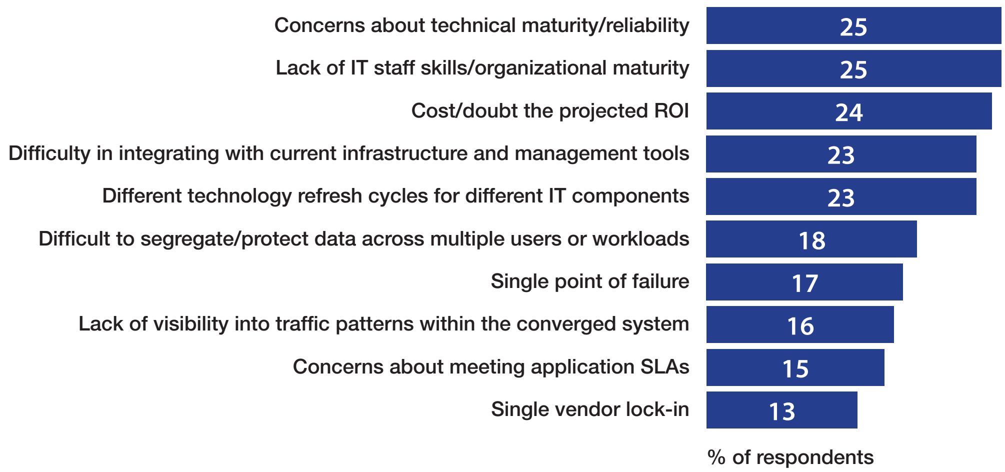
IDC Analyze the Future

Figure 11 illustrates the top challenges for businesses in embracing composable infrastructure. Concerns about technical maturity and reliability, the ability of staff to manage a new platform, doubts about promised ROI, and difficulty integrating with existing tools are at the top of mind for respondents. For system vendors like HPE to dominate this space, they have to address these concerns. Their credibility and their decades of collective experience in building enterprise infrastructure platforms are what will count.

FIGURE 11

Top Challenges in Adopting Composable Systems

Q. What do you see as the top 2 challenges in moving to composable systems?



n = 301
Base = all respondents
Note: Data is weighted by employee size.
Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

Concepts of disaggregating IT systems down to discrete resources have been discussed in the industry for some time, even though the enabling technologies have yet to mature. DX and NGAs driven by shifts in business strategy will fuel the demand for composable infrastructure.

Conclusion

Concepts of disaggregating IT systems down to discrete resources have been discussed in the industry for some time, even though the enabling technologies have yet to mature. DX and NGAs driven by shifts in business strategy will fuel the demand for composable infrastructure. This emerging category of systems seeks to create greater business agility, lower operational costs, and increased application performance. Businesses should stay open to embracing these systems provided concerns such as ease of management and promised ROI are addressed.

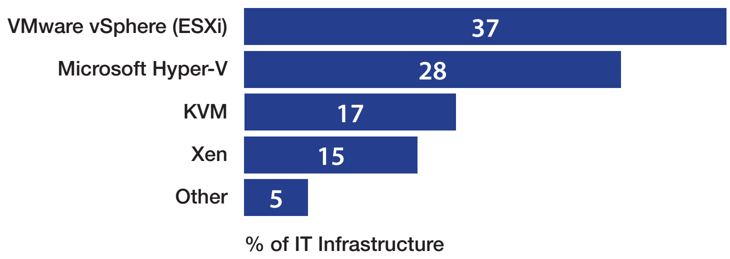
Appendix

Key Survey Findings

FIGURE 12

Multihypervisor Environments

Q. What percentage of your organization's IT infrastructure uses the following hypervisors?



n = 301

Base = all respondents

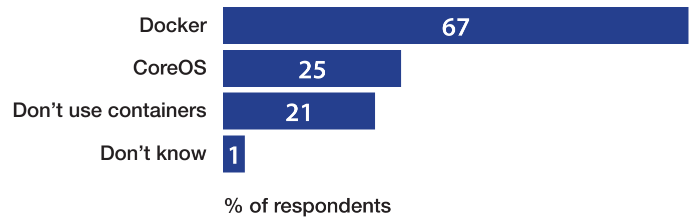
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 13

Multicontainer Environments

Q. Which container environments is your organization primarily using in production?



n = 301

Base = all respondents

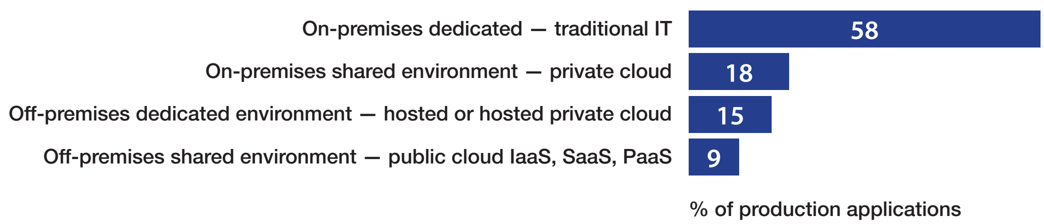
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 14

Location of Infrastructure

Q. What percentage of your organization's production applications (including both commercial and custom applications) are currently running in the following environments?



n = 301

Base = all respondents

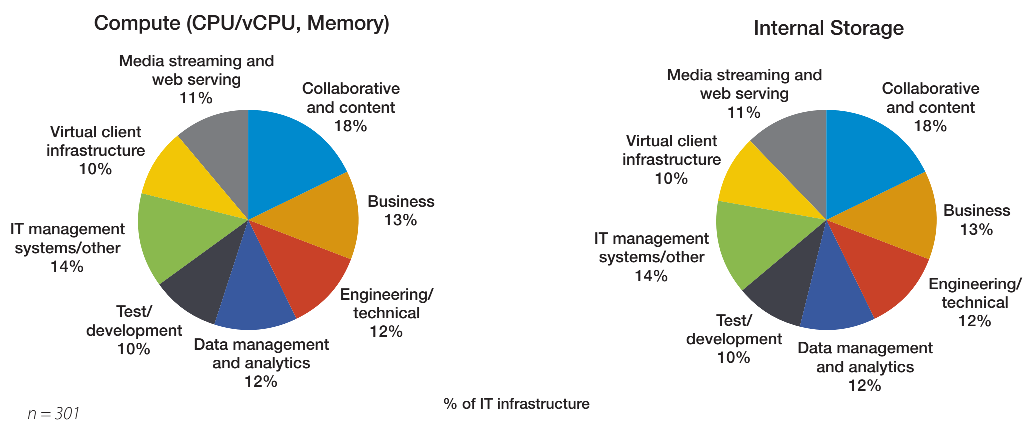
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 15

Workload Distribution: Compute

Q. How does the compute (CPU/vCPU) allocation break down among the following workloads?



n = 301

Base = all respondents

Notes: Data is weighted by employee size. IT management systems includes security, networking, and systems and storage management.

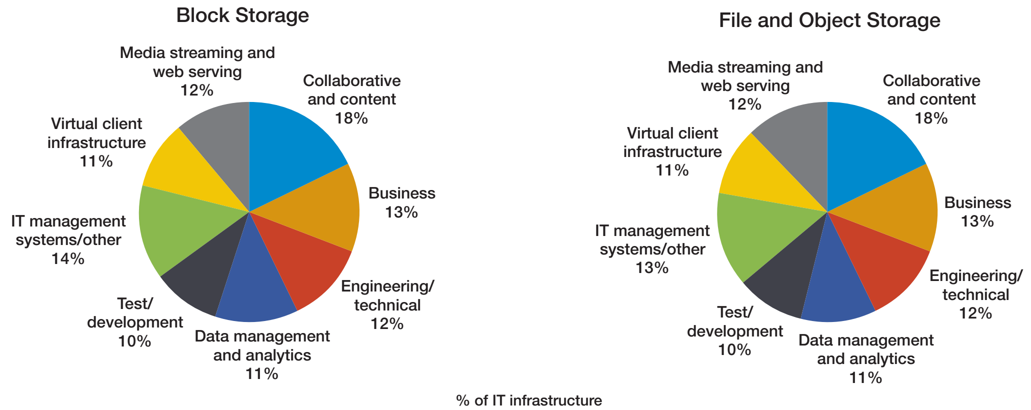
Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

IDC Analyze the Future

FIGURE 16

Workload Distribution: External Storage

Q. How does the storage allocation break down among the following workloads?



n = 301

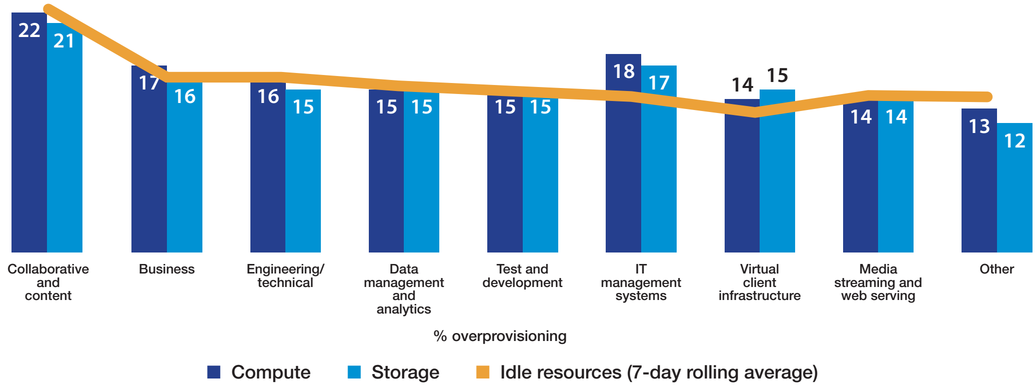
Base = all respondents

Notes: Data is weighted by employee size. IT management systems includes security, networking, and systems and storage management.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 17

Typical Compute and Storage Overprovisioning and Idle Resources



n = 301

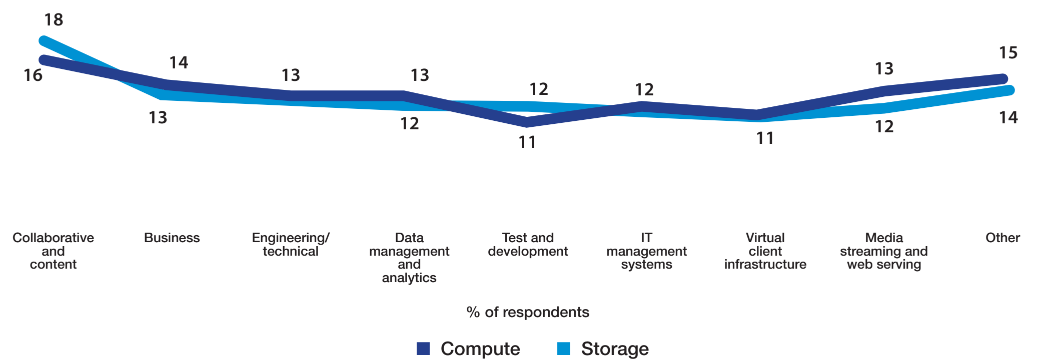
Base = all respondents

Notes: Data is weighted by employee size. IT management systems includes security, networking, and systems and storage management.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 18

YoY Compute and Storage Growth by Application



n = 301

Base = all respondents

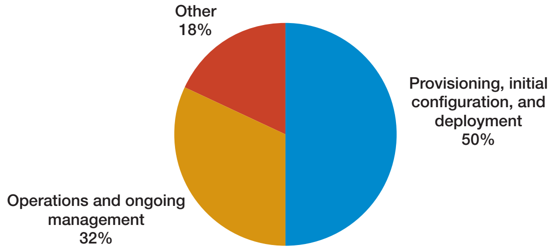
Notes: Data is weighted by employee size. IT management systems includes security, networking, and systems and storage management.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 19

Resource Allocation Between Provisioning and Operations

Q. Approximately how many full-time and contract personnel does your organization have on staff to provision and manage its IT infrastructure?



n = 301

Base = all respondents

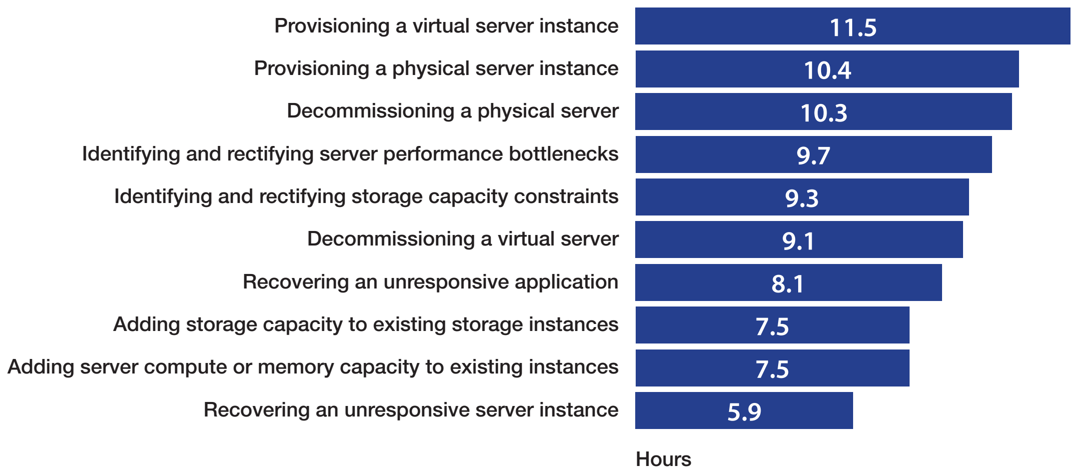
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 20

Time Spent on Provisioning Tasks

Q. How much time do you spend performing each task?



n = 301

Base = all respondents

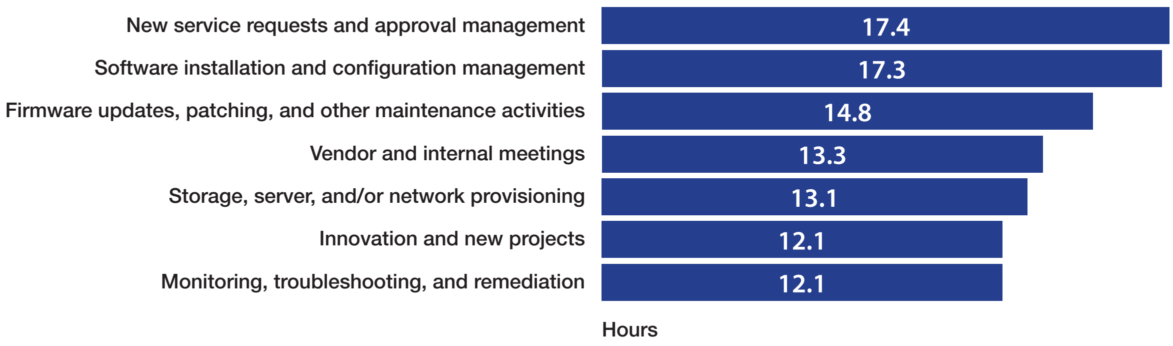
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 21

Time Spent on Maintenance Tasks

Q. How much time do you spend performing each task, and how much do you believe is wasted due to process or technology inefficiencies?



n = 301

Base = all respondents

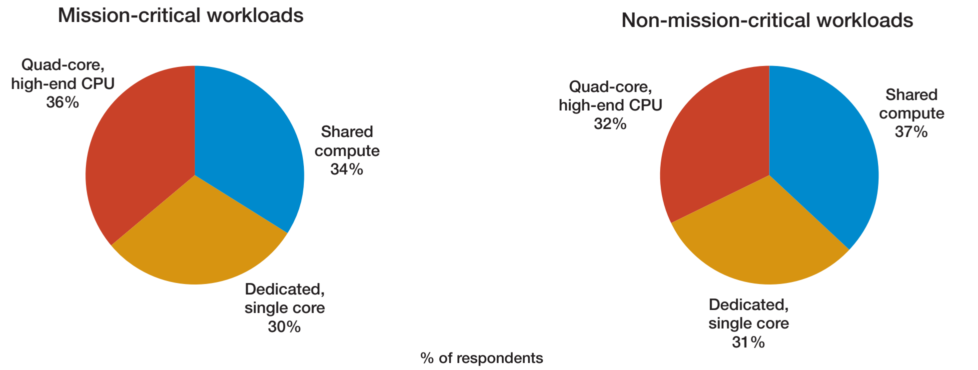
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 22

Compute Attributes for Workloads

Q. Which specific compute attributes are important for deploying workloads in your organization?



n = 301

Base = all respondents

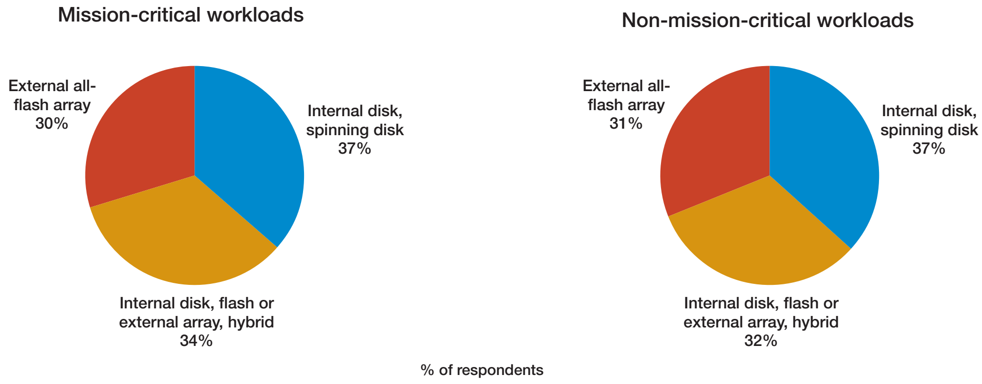
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 23

Storage Attributes for Workloads

Q. Which specific storage attributes are important for deploying workloads in your organization?



n = 301

Base = all respondents

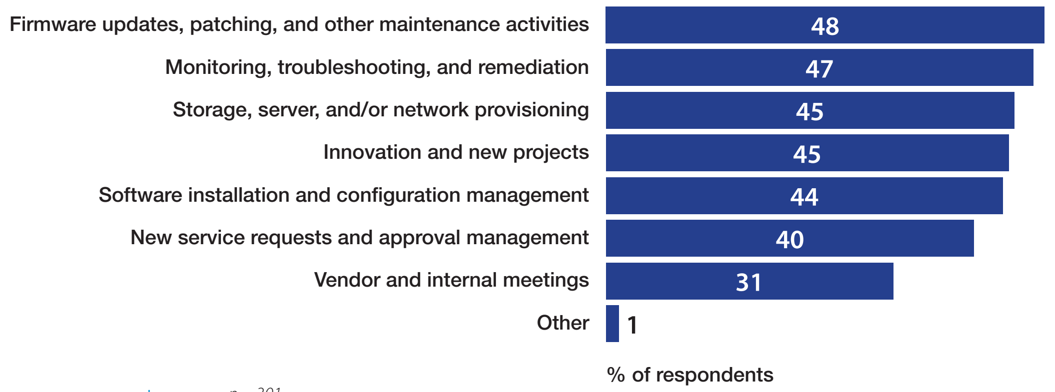
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 24

Top Priorities for Infrastructure Optimization

Q. Which do you believe are the top 3 priorities for optimization in how you manage your organization's infrastructure?



n = 301

Base = all respondents

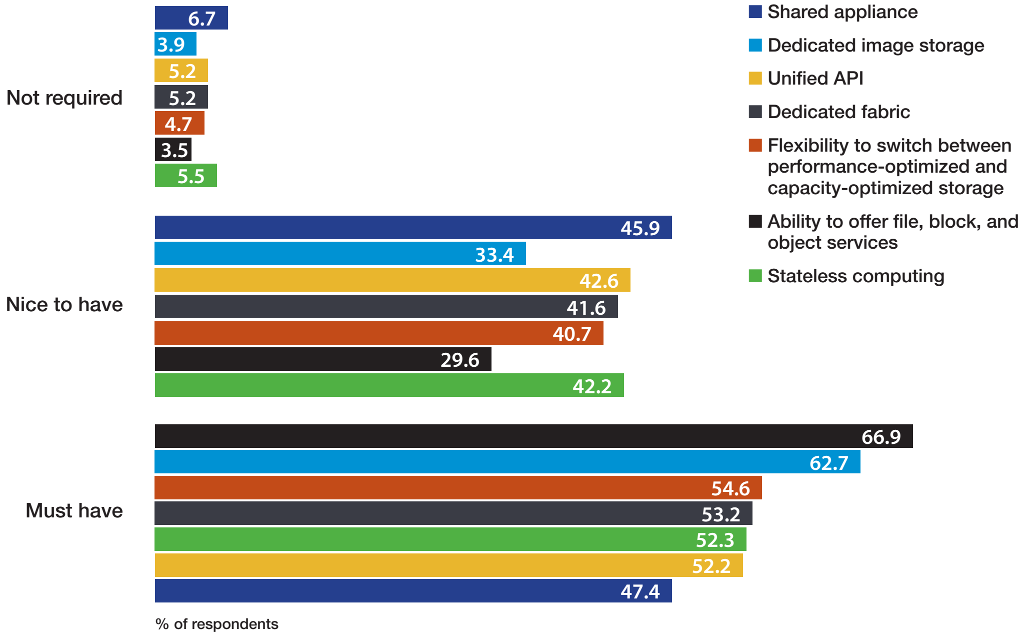
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 25

Important Attributes for Private/Hybrid Cloud

Q. What specific infrastructure attributes are crucial for your private cloud or on-premises portion of your hybrid cloud?



n = 301

Base = all respondents

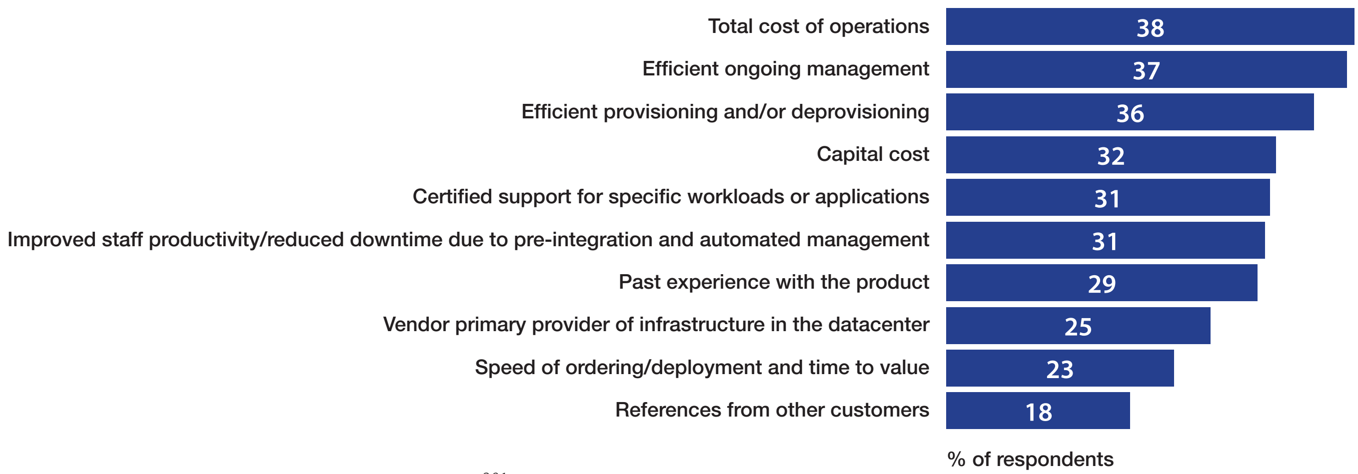
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 26

Selection Criteria for Infrastructure Components

Q. When it comes to evaluating the purchase of infrastructure equipment, what are the three most important selection criteria for your organization?



n = 301

Base = all respondents

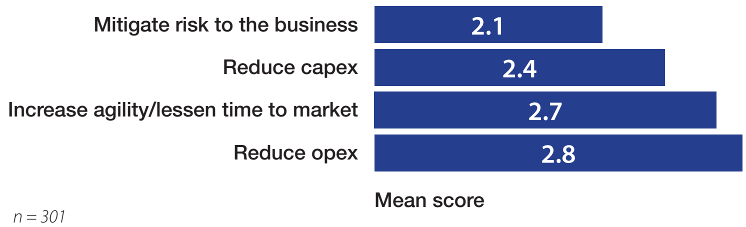
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 27

Business Outcomes

Q. For the following business outcomes that can be achieved through IT, rank them in the order of importance (1, 2, 3) to your organization.



n = 301

Base = all respondents

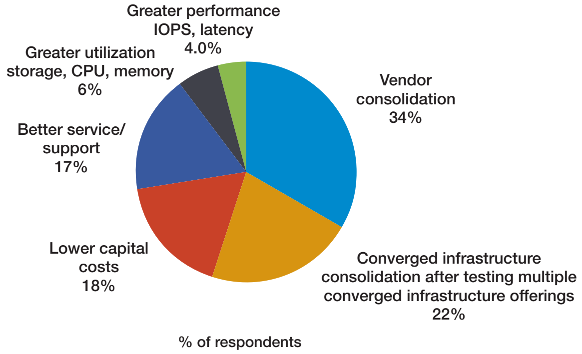
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 28

Reasons for Switching Suppliers

Q. What are your top reasons for evaluating other suppliers?



n = 301

Base = all respondents

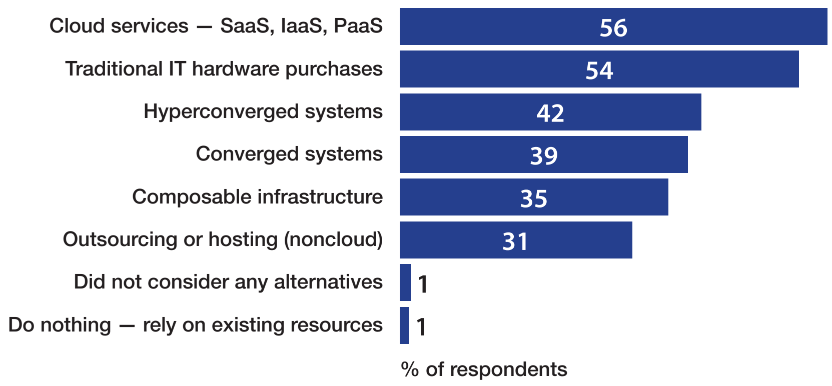
Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

FIGURE 29

Examining Alternatives

Q. In the process of evaluating new infrastructure systems, what other alternatives, if any, will you consider to meet your compute, storage, and network requirements?



n = 301

Base = all respondents

Note: Data is weighted by employee size.

Source: IDC's Infrastructure Usage and Overprovisioning Trends Survey, *November 2016*

IDC Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Analyze the Future