

Gartner.

This research note is restricted to the personal use of ().

Magic Quadrant for Cloud Infrastructure as a Service, Worldwide

Published 23 May 2018 - ID G00336148 - 84 min read

By Analysts Dennis Smith, Lydia Leong, Raj Bala

The global market for cloud IaaS has consolidated around hyperscale service providers. Infrastructure and operations leaders should adopt strategically, but consider scenario-specific providers as well.

Market Definition/Description

This document was revised on 25 May 2018. The document you are viewing is the corrected version. For more information, see the [Corrections \(https://www.gartner.com/technology/about/policies/current_corrections.jsp\)](https://www.gartner.com/technology/about/policies/current_corrections.jsp) page on gartner.com.

Cloud computing is a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service using internet technologies. Cloud infrastructure as a service (IaaS) is a type of cloud computing service; it parallels the infrastructure and data center initiatives of IT. Cloud compute IaaS constitutes the largest segment of this market (the broader IaaS market also includes cloud storage and cloud printing). Only cloud compute IaaS is evaluated in this Magic Quadrant. It does not cover service providers that exclusively offer cloud storage, platform as a service (PaaS), SaaS, cloud service brokerage (CSB) or any other type of cloud service, nor does it cover the hardware and software vendors that may be used to build cloud infrastructure. Furthermore, this Magic Quadrant is not an evaluation of the broad, generalized cloud computing strategies of the companies profiled.

In the context of this Magic Quadrant, cloud compute IaaS (hereafter referred to simply as "cloud IaaS" or "IaaS") is defined as a standardized, highly automated offering, where compute resources, complemented by storage and networking capabilities, are owned by a service provider and offered to the customer on demand. The resources are scalable and elastic in near real time, and metered by use. Self-service interfaces are exposed directly to the customer, including a web-based UI and an API. The resources may be single-tenant or multitenant, and hosted by the service provider or on-premises in the customer's data center. Thus, this Magic Quadrant covers both public and private cloud IaaS offerings. Further information about IaaS is available in "Technology Insight for Cloud Infrastructure as a Service."

This Magic Quadrant evaluation includes not only the cloud IaaS resources themselves, but also the automated management of those resources, management tools delivered as services and cloud software infrastructure services. The last category includes middleware and databases as a service, up to and including some capabilities that might be classifiable as PaaS. New computing abstractions, especially serverless computing, have increasingly blurred the line between IaaS and high-control PaaS capabilities (see "Evolution of Server Computing: VMs to Containers to Serverless – Which to Use When?"). We consider a capability to be an infrastructure capability regardless of whether hardware infrastructure or software infrastructure is being delivered as a cloud service.

IaaS and PaaS represent a continuum, as discussed in "Technology Insight for Integrated IaaS and PaaS," and for the last several years, customers have simply consumed cloud services across this continuum without concerning themselves with categorization. Fully integrated offerings are referred to in this Magic Quadrant as "integrated IaaS+PaaS." However, we do not evaluate the full spectrum of an integrated provider's offerings in this Magic Quadrant; for instance, we do not evaluate noninfrastructure machine learning (ML) services, or Internet of Things (IoT) device management services.

Even though some businesses may use an application PaaS (aPaaS) in a very IaaS-like manner, we have excluded PaaS providers from this Magic Quadrant, with the exception of those PaaS providers that also have a qualifying IaaS offering. PaaS offerings do not allow customers to obtain raw virtual machines (VMs) that can be loaded with arbitrary operating systems, middleware and applications, which is a requirement for being considered as IaaS. For PaaS providers, see "Magic Quadrant for Enterprise Application High-Productivity Platform as a Service" and "Magic Quadrant for Enterprise Integration Platform as a Service."

We draw a distinction between cloud infrastructure *as a service*, and cloud infrastructure *as an enabling technology*; we call the latter "cloud-enabled system infrastructure" (CESI). In cloud IaaS, the capabilities of a CESI are directly exposed to the customer through self-service. However, other services, including noncloud services, may be delivered on top of a CESI; these cloud-enabled services may include forms of managed hosting, data center outsourcing and other IT outsourcing services. In this Magic Quadrant, we evaluate only cloud IaaS offerings; we do not evaluate cloud-enabled services.

Gartner's clients are mainly enterprises, midmarket businesses and technology companies of all sizes, and the evaluation focuses on typical client requirements. This Magic Quadrant covers all the common use cases for cloud IaaS, including development and testing, production environments (including those supporting mission-critical workloads) for both internal and customer-facing applications, batch computing (including high-performance computing [HPC]), and disaster recovery. It includes not only the hosting of single-application workloads, but also the replacement of traditional enterprise data centers with cloud environments that can support a highly diverse range of workloads. It includes suitability for a wide range of application design patterns, including cloud-native applications, web-era applications and legacy enterprise applications.

Customers typically exhibit a bimodal IT sourcing pattern for cloud IaaS (see "Bimodal IT: How to Be Digitally Agile Without Making a Mess" and "Best Practices for Planning a Cloud Infrastructure-as-a-Service Strategy – Bimodal IT, Not Hybrid Infrastructure"). Most cloud IaaS is bought for Mode 2 agile IT, emphasizing developer productivity and business agility, but an increasing amount of cloud IaaS is being bought for Mode 1 traditional IT, with an emphasis on cost reduction, safety and security. Infrastructure and operations (I&O) leaders typically lead the sourcing for Mode 1 cloud needs. By contrast, sourcing for Mode 2 offerings is typically driven by enterprise architects, application development leaders and digital business leaders. This Magic Quadrant considers both sourcing patterns and their associated customer behaviors and requirements.

This Magic Quadrant strongly emphasizes self-service and automation in a standardized environment. It focuses on the needs of customers whose primary need is self-service cloud IaaS, although this may be supplemented by a small amount of colocation or traditional hosting. In self-service cloud IaaS, the customer retains most of the responsibility for IT operations (even if the customer subsequently chooses to outsource that responsibility via third-party managed services). Third-party managed service providers are covered in the "Magic Quadrant for Public Cloud Infrastructure Managed Service Providers, Worldwide."

Organizations that need significant customization or managed services for a single application, or that are seeking cloud IaaS as a supplement to a traditional hosting solution ("hybrid hosting"), should consult the Market Guide and Magic Quadrants for managed hosting instead ("Market Guide for Managed Hybrid Cloud Hosting, North America," "Magic Quadrant for Managed Hybrid Cloud Hosting, Europe" and "Magic Quadrant for Managed Hybrid Cloud Hosting, Asia/Pacific"). Organizations that want a fully custom-built solution, or managed services with an underlying CESI, should consult the Magic Quadrants for data center outsourcing and infrastructure utility services ("Magic Quadrant for Data Center Outsourcing and Infrastructure Utility Services, North America," "Magic Quadrant for Data Center Outsourcing and Infrastructure Utility Services, Europe" and "Magic Quadrant for Data Center Outsourcing and Infrastructure Utility Services, Asia/Pacific").

This Magic Quadrant evaluates all industrialized cloud IaaS solutions, whether public cloud (multitenant or mixed-tenancy), community cloud (multitenant, but limited to a particular customer community), or private cloud (fully single-tenant, hosted by the provider or on-premises). It is not merely a Magic Quadrant for public cloud IaaS. To be considered industrialized, a service must be standardized across the customer base; it is insufficient to use a common reference architecture. Although most of the providers in this Magic Quadrant do offer custom private cloud IaaS, we have not considered these nonindustrialized offerings in our evaluations. Organizations that are looking for custom-built, custom-managed private clouds should use our Magic Quadrants for data center outsourcing and infrastructure utility services instead (see above).

Understanding the Vendor Profiles, Strengths and Cautions

Cloud IaaS providers that target enterprise and midmarket customers generally offer a high-quality service, with excellent availability, good performance, high security and good customer support. Exceptions will be noted in this Magic Quadrant's evaluations of individual providers. Note that, when we say "all providers," we specifically mean "all the evaluated providers included in this Magic Quadrant," not all cloud IaaS providers in general. Keep the following in mind when reading the vendor profiles:

- All the providers have a public cloud IaaS offering. A few also have an industrialized private cloud offering, where every customer is on standardized infrastructure and cloud management tools, although this may or may not resemble the provider's public cloud service in either architecture or quality. A single architecture and feature set and cross-cloud management, for both public and private cloud IaaS, make it easier for customers to combine and migrate across service models as their needs dictate. They also enable the provider to use its engineering investments more effectively. Most of the providers also offer custom private clouds.
- All of the providers target midmarket businesses and enterprises, as well as other companies that use technology at scale. Some of the providers may also target small businesses and startups. Just because a provider targets a segment, however, does not necessarily mean that it is well-suited to that segment's needs. Furthermore, not all providers have the capacity to serve very large-scale customers, and some have capacity constraints in particular regions.
- Most of the providers are oriented toward the needs of Mode 2 agile IT. These providers typically emphasize capabilities for new applications and a DevOps orientation, but are also capable of running legacy applications and being managed in a traditional fashion.

Most providers are also capable of serving Mode 1 traditional IT, especially IT operations organizations, with an emphasis on control, governance and security. Keep in mind that a "rented virtualization" approach can be used for both new and legacy applications, but is unlikely to provide transformational benefits.

- All the providers offer basic cloud IaaS — compute, storage and networking resources as a service. They also offer additional value-added capabilities as well, notably cloud software infrastructure services — typically middleware and databases as a service — up to and including PaaS capabilities. These services, along with IT operations management (ITOM) capabilities as a service (especially DevOps-related services), are a vital differentiator in the market, especially for Mode 2 agile IT buyers. While all providers have both IaaS and PaaS offerings, only some offer integrated IaaS+PaaS; these providers are explicitly noted.
- We consider an offering to be public cloud IaaS if the storage and network elements are shared; the compute can be multitenant, single-tenant or both. Private cloud IaaS uses single-tenant compute and storage, but unless the solution is on the customer's premises, the network is usually still shared.
- All the providers claim to have high security standards. The extent of the security controls provided to customers varies significantly, though. All the providers evaluated can offer solutions that will meet common regulatory compliance needs, unless otherwise noted. All the providers have SSAE 16 audits for their data centers (see Note 1). Some may have security-specific third-party assessments such as ISO 27001 or SOC 2 for their cloud IaaS offerings (see Note 2). Both provide a relatively high level of assurance that the providers are adhering to generally accepted practices for the security of their systems, but do not address the extent of controls offered to customers. Security is a shared responsibility; customers need to correctly configure controls and may need to supply additional controls beyond what their provider offers. Furthermore, providers vary in their degree of transparency, although customers typically have access to third-party assessment reports under a nondisclosure agreement.
- In general, monthly compute availability SLAs of 99.95% and higher are the norm, and they are typically higher than availability SLAs for managed hosting. Service credits for outages in a given month are typically capped at 100% of the monthly bill, but some providers have caps as low as 25%. This availability percentage is typically non-negotiable, as it is based on an engineering estimate of the underlying infrastructure reliability. Maintenance windows are normally excluded from the SLA.
- Some providers have a compute availability SLA that requires the customer to use compute capabilities in at least two fault domains (sometimes known as "availability zones" or the like); an SLA violation requires both fault domains to fail. Providers with an SLA of this type are explicitly noted as having a multi-fault-domain SLA.
- Very few of the providers have an SLA for compute or storage performance. None of these providers oversubscribe compute or RAM resources in standard compute instances, but some may have special, less expensive instance types that do, such as "burstable" instances.
- Many providers have additional SLAs covering network availability and performance, customer service responsiveness and other service aspects.
- Infrastructure resources are not normally automatically replicated into multiple data centers, unless otherwise noted; customers are responsible for their own business continuity. Some providers offer optional disaster recovery solutions.
- All providers offer, at minimum, per-hour metering of VMs, and some can offer shorter metering increments, which can be more cost-effective for short-term batch jobs. Providers charge on a per-VM basis, unless otherwise noted.
- Some of the providers are able to offer bare-metal physical servers on a dynamic basis, priced by the hour. Providers with a bare-metal option are noted as such.
- All the providers offer an option for colocation, unless otherwise noted. Many customers have needs that require a small amount of supplemental colocation in conjunction with their cloud — most frequently for a large-scale database, but sometimes for specialized network equipment, software that cannot be licensed on virtualized servers, or legacy equipment. Colocation is specifically mentioned only when a service provider actively sells colocation as a stand-alone service; a significant number of midmarket customers plan to move into colocation and then gradually migrate into that provider's IaaS offering. If a provider does not offer colocation itself, but can meet such needs via a partner exchange, this is explicitly noted.
- Some providers offer a software marketplace where software vendors specially license and package their software to run on that provider's cloud IaaS offering. Marketplace software can be automatically installed with a click, and can be billed through the provider. Some marketplaces also contain other third-party solutions and services.

- All providers offer enterprise-class support with 24/7 customer service, via phone, email and chat, along with an account manager. Most providers include this with their offering. Some offer a lower level of support by default, but allow customers to pay extra for enterprise-class support.
- All the providers will sign contracts with customers, can invoice and can consolidate bills from multiple accounts. While some may also offer online sign-up and credit card billing, they recognize that enterprise buyers prefer contracts and invoices. Some will sign "zero dollar" contracts that do not commit a customer to a certain volume.
- Unless otherwise noted, all providers will sign the following contract addendums: a U.S. Health Insurance Portability and Accountability Act Business Associate Agreement (HIPAA BAA); an EU Data Protection Directive (95/46/EC) data processing agreement (DPA), which includes the model clauses; and an EU General Data Protection Regulation (GDPR) DPA.
- Some of the providers offer optional managed services on IaaS. Such providers may not offer the same type of managed services on IaaS as they do in their broader managed hosting or data center outsourcing services. Some may have managed service provider (MSP) or system integrator (SI) partners that provide managed and professional services.
- All the evaluated providers offer a portal, documentation, technical support, customer support and contracts in English. Some can provide one or more of these in languages other than English. Most providers can conduct business in local languages, even if all aspects of service are English-only.
- All the providers are part of very large corporations or otherwise have a well-established business. However, many of the providers are undergoing significant re-evaluation of their cloud IaaS businesses. Existing and prospective customers should be aware that such providers may make significant changes to the strategy and direction of their cloud IaaS business. These changes may include replacing their current offering with a new platform, or exiting this business entirely in favor of partnering with a more successful provider.

In previous years, this Magic Quadrant has provided significant technical detail on the offerings. These detailed evaluations are now published in "Critical Capabilities for Public Cloud Infrastructure as a Service, Worldwide" instead.

The service provider descriptions are accurate as of the time of publication. Our technical evaluation of service features took place between January 2018 and March 2018.

Format of the Vendor Descriptions

When describing each provider, we first summarize the nature of the company and then provide information about its industrialized cloud IaaS offerings in the following format:

Offerings: A list of the industrialized cloud IaaS offerings (both public and private) that are directly offered by the provider. Also included is commentary on the ways in which these offerings deviate from the standard capabilities detailed in the Understanding the Vendor Profiles, Strengths and Cautions section above. We also list related capabilities of interest, such as object storage, content delivery network (CDN), aPaaS and managed services, but this is not a comprehensive listing of the provider's offerings.

Locations: Cloud IaaS data center locations by country, languages that the company does business in and languages that technical support can be conducted in.

Provider maturity: Cloud IaaS providers vary dramatically in their level of risk – the degree to which a customer can trust them to be secure, reliable, stable businesses. We provide a three-tier maturity model in "Inform Your Cloud Service Choice With Provider Maturity," and for each provider, we list its tier in that maturity model. Tier 1 providers are global megavendors. Tier 2 providers are engaged in a struggle for sustainability, with the largest hoping to break into the top tier. Tier 2 is divided into two categories: The Tier 2A category is composed of established technology vendors, while the Tier 2B category consists of cloud-only (or cloud-primary) vendors that have grown enough to be significant. Tier 3 providers are emerging and risky; there are none on this Magic Quadrant. We recommend that customers focus risk assessment and mitigation efforts on Tier 2 providers, which may be undesirably immature, but are more likely to be willing to offer better contractual terms, SLAs and pricing in order to ease customer concerns. Tier 2 providers require attentive vendor management and a potential exit strategy.

Recommended mode: We note whether the vendor's offerings are likely to appeal to Mode 1 safety-and-efficiency-oriented IT, Mode 2 agility-oriented IT, or both. We also note whether the offerings are likely to be useful for organizations seeking IT transformation. This recommendation reflects the way that a provider goes to market, provides service and support, and designs its offerings. All such statements are specific to the provider's cloud IaaS offering, not the provider as a whole.

Recommended uses: These are the circumstances under which we recommend the provider. These are not the only circumstances in which it may be a useful provider, but these are the scenarios for which, in Gartner's opinion, the provider is best-suited. For all the vendors, the recommended uses are specific to self-managed cloud IaaS. However, some of the providers also have managed services, as well as other cloud and noncloud services that may be used in conjunction with cloud IaaS. These include hybrid hosting (customers sometimes blend solutions, such as an entirely self-managed front-end web tier on public cloud IaaS, with managed hosting for the application servers and database), as well as hybrid IaaS/PaaS solutions. Even though we do not evaluate managed services, nonintegrated PaaS and the like in this Magic Quadrant, they are part of a vendor's overall value proposition and we mention them in the context of providing more comprehensive solution recommendations.

In the list of offerings, we state the basis of each provider's virtualization technology and, if relevant, its cloud infrastructure framework (CIF) – the software used to create the cloud service, such as OpenStack or VMware vCloud Director. We also state what APIs it supports – the Amazon Web Services (AWS), OpenStack and vCloud APIs are the three that have broad adoption, but many providers also have their own unique APIs. Note that supporting one of the three common APIs does not provide assurance that a provider's service is compatible with a specific tool that purports to support that API; the completeness and accuracy of API implementations vary considerably. Furthermore, the use of the same underlying CIF or API compatibility does not indicate that two services are interoperable. Specifically, OpenStack-based clouds differ significantly from one another, limiting portability; the marketing hype of "no vendor lock-in" is, practically speaking, untrue.

For many customers, the underlying hypervisor will matter, particularly for those that intend to run commercial software on IaaS. Many independent software vendors (ISVs) support only VMware virtualization. Vendors that support Xen may support only Citrix XenServer, not open-source Xen (which is often customized by IaaS providers and is likely to be different from the current open-source version). Similarly, some ISVs may support the Kernel-based Virtual Machine (KVM) hypervisor in the form of Red Hat Enterprise Virtualization, whereas many IaaS providers use open-source KVM.

For a detailed technical description of public cloud IaaS offerings, along with a use-case-focused technical evaluation, see "Critical Capabilities for Public Cloud Infrastructure as a Service, Worldwide."

We also provide a detailed list of evaluation criteria in "Evaluation Criteria for Cloud Infrastructure as a Service." We have used those criteria to perform in-depth assessments of several providers: see "In-Depth Assessment of Amazon Web Services IaaS, March 2018," "In-Depth Assessment of Google Cloud Platform IaaS, March 2018" and "In-Depth Assessment of Microsoft Azure IaaS, March 2018."

Magic Quadrant

Figure 1. Magic Quadrant for Cloud Infrastructure as a Service, Worldwide



Source: Gartner (May 2018)

Customers that are comparing the 2017 and 2018 Magic Quadrants may notice that the scale of the Magic Quadrant graphic has changed very significantly. This year, we chose more stringent inclusion criteria, which had the effect of only including global vendors that currently have hyperscale integrated IaaS and PaaS offerings, or that are currently developing those offerings. Consequently, we dropped eight of the 14 vendors that were on the 2017 Magic Quadrant. Those were the vendors on the left-hand side of the graph, so effectively, the 2018 Completeness of Vision axis has expanded, such that its entirety now encompasses what was previously just the right-hand side of the graph. Consequently, individual vendor movement on Completeness of Vision represents both this rescaling effect and the broadening or narrowing of a vendor's vision between the previous evaluation and this year's evaluation.

These changes reflect Gartner's belief that customer evaluations are currently primarily focused on vendors for strategic adoption across a broad range of use cases. While customers still search for more focused, scenario-specific providers, these providers should be evaluated in the context of that specific workload, rather than compared in a broader market context.

Furthermore, the overall customer requirements continue to expand, and there are still many unmet needs in this market, resulting in greater room for improvement for all vendors in both Completeness of Vision and Ability to Execute (see the What Trends Are Currently Influencing Buyers? section for details).

Customers comparing the 2017 and 2018 Magic Quadrant should also keep in mind that the Magic Quadrant shows the comparative positioning of vendors within the market. Vendors may substantially improve their capabilities from year to year, yet not achieve significant movement in their position, because their position is relative to the overall market. Due to the Completeness of Vision scale change, this effect is particularly pronounced this year.

Vendor Strengths and Cautions

Alibaba Cloud

Alibaba Cloud, a subsidiary of Alibaba Group, is a cloud-focused service provider with headquarters in China. It was established in 2009, and initially provided services to Alibaba Group's e-commerce businesses. This Magic Quadrant evaluation is focused upon Alibaba Cloud's international business, which is headquartered in Singapore, and our technical assessment was performed using the international service.

Offerings: Alibaba Cloud is integrated IaaS+PaaS. It offers Xen and KVM-virtualized multitenant compute (Elastic Compute Service [ECS]) with compute-independent block storage (cloud disks), object storage (Object Storage Service [OSS]), a CDN service, a Docker-based container service (Alibaba Cloud Container Service) and a variety of PaaS-layer services, including a family of database services (ApsaraDB). More capabilities are offered in the China service than in the international service.

Locations: Alibaba Cloud operates multiple regions in China and additionally has a presence in the U.S. (East and West Coasts), Germany, Australia, Hong Kong, Japan, Singapore, India, Malaysia and the United Arab Emirates. It has local sales in the U.S., China, Germany, Australia, Japan, Hong Kong and Singapore. The China service portal, documentation and support are in Mandarin. The international portal, documentation and support are in English, Mandarin and Japanese.

Provider maturity: Tier 2B. Alibaba Cloud is a market leader in China, but is a relatively recent entrant to the global market.

Recommended mode: Alibaba Cloud appeals to Mode 2 buyers that seek infrastructure that supports agile workloads.

Recommended uses: Digital business workloads for customers that are based in China, or need to locate cloud infrastructure in China.

Strengths

- Alibaba Cloud has an extensive set of public cloud integrated IaaS+PaaS offerings, comparable to the service portfolios of other hyperscale providers, and continues to aggressively invest in research and development. It also offers an on-premises private cloud stack of software and services, Apsara Stack (though this is not yet available internationally), which offers hybrid cloud compatibility with Alibaba Cloud's public cloud offerings.
- Alibaba Group has the financial wherewithal to continue investing in global expansion via new regions, international-specific engineering efforts, and regional sales, support, and marketing for Alibaba Cloud. In China, Alibaba has built an impressive ecosystem consisting of MSPs and ISVs, and it has been successful at gradually building a global ecosystem for its international offering. It has the potential to become an alternative to the global hyperscale cloud providers in select regions over time.
- Alibaba Cloud's technical capabilities, built for the China market, can usually be readily leveraged for the international offering. Alibaba Cloud is the current market share leader for cloud IaaS in China, and performs particularly well with Chinese digital businesses and agencies within the Chinese government.

Cautions

- Alibaba Cloud needs to build both technical and go-to-market capabilities desired by traditional enterprise customers with non-cloud-native workloads, especially Microsoft Windows-based workloads. Alibaba has striven for feature parity against its global competitors, but as a result, Alibaba's international offering has very little in the way of unique differentiation compared to other global hyperscale providers. Additionally, Alibaba Cloud's vision seems inextricably tied to that of its global competitors; it takes liberal inspiration from competitors when developing service capabilities and branding.
- Alibaba Cloud's international offering was launched in mid-2016. It has a limited track record, and does not have the full capabilities or performance of the China offering. The international portal can be confusing to use, as the capabilities available in each region are not always clear.

- Alibaba Cloud has substantial challenges that it must overcome before it can translate its success in China to markets outside of its home territory. Alibaba Cloud has rapidly expanded its offering to markets outside of China since the launch of its international offering. However, the company does not have substantial mind share with buyers in most markets, as it is still building the required local talent, industry expertise and go-to-market capabilities. Prospective international customers may also perceive security and regulatory compliance concerns when using a Chinese company, even though Alibaba Cloud has undergone third-party audits.

Amazon Web Services

Amazon Web Services (AWS), a subsidiary of Amazon, is a cloud-focused service provider. It pioneered the cloud IaaS market in 2006.

Offerings: AWS is integrated IaaS+PaaS. Its Elastic Compute Cloud (EC2) offers metered-by-the-second multitenant and single-tenant VMs, as well as bare-metal servers. AWS's hypervisors are based on Xen and KVM. There is multitenant block and file storage, along with extensive additional IaaS and PaaS capabilities. These include object storage with an integrated CDN (Amazon Simple Storage Service [S3] and CloudFront), Docker container services (Amazon Elastic Container Service [ECS], ECS for Kubernetes [EKS], and Fargate container instances), a batch computing service (AWS Batch), event-driven "serverless computing" (Lambda) and an aPaaS-like developer experience (Elastic Beanstalk). It is willing to negotiate large-scale single-tenant and on-premises deals (such as the U.S. intelligence community cloud deal). The AWS Marketplace has an extensive selection of third-party software and services. VMware offers a VMware Cloud Foundation service within AWS data centers (VMware Cloud on AWS). Enterprise-grade support is extra. It has a multi-fault-domain SLA. Colocation needs are met via partner exchanges (AWS Direct Connect).

Locations: AWS groups its data centers into regions, each of which contains at least two availability zones (data centers). It has multiple regions across the U.S., as well as in Canada, France, Germany, Ireland, U.K., Australia, India, Japan, Singapore, South Korea and Brazil. It also has one region dedicated to the U.S. federal government. There are two China regions — Beijing (operated by Sinnet) and Ningxia (operated by Ningxia Western Cloud Data Technology [NWCD]) — which require a China-specific AWS account. It has a global sales presence. The portal and documentation are provided in English, Dutch, French, German, Italian, Japanese, Korean, Mandarin, Portuguese and Spanish. The primary languages for support are English, Japanese and Mandarin, but AWS will contractually commit to providing support in a large number of other languages.

Provider maturity: Tier 1. AWS has been the market pioneer and leader in cloud IaaS for over 10 years.

Recommended mode: AWS strongly appeals to Mode 2 buyers, but is also frequently chosen for Mode 1 needs. AWS is the provider most commonly chosen for strategic, organizationwide adoption. Transformation efforts are best undertaken in conjunction with an SI.

Recommended uses: All use cases that run well in a virtualized environment.

Strengths

- AWS has been the dominant market leader and an IT thought leader for more than 10 years, not only in IaaS, but also in integrated IaaS+PaaS, with an end-of-2017 revenue run rate of more than \$20 billion. It continues to aggressively expand into new IT markets via new services as well as acquisitions, adding to an already rich portfolio of services. It also continues to enhance existing services with new capabilities, with a particular emphasis on management and integration.
- AWS is the provider most commonly chosen for strategic adoption; many enterprise customers now spend over \$5 million annually, and some spend over \$100 million. While not the ideal fit for every need, it has become the "safe choice" in this market, appealing to customers that desire the broadest range of capabilities and long-term market leadership.
- AWS is the most mature, enterprise-ready provider, with the strongest track record of customer success and the most useful partner ecosystem. Thus, it is the provider not only chosen by customers that value innovation and are implementing digital business projects, but also preferred by customers that are migrating traditional data centers to cloud IaaS. It can readily support mission-critical production applications, as well as the implementation of highly secure and compliant solutions. Implementation, migration and management are significantly eased by AWS's ecosystem of more than 2,000 consulting partners that offer managed and professional services. AWS has the broadest cloud IaaS provider ecosystem of ISVs, which ensures that customers are able to obtain support and licenses for most commercial software, as well as obtain software and SaaS solutions that are preintegrated with AWS.

Cautions

- AWS's extensive portfolio of services requires expertise to implement. Customers should be aware that while it's easy to get started, optimal use — especially keeping up with new service innovations and best practices, and managing costs — may challenge even highly agile, expert IT organizations, including AWS partners. As new, less-experienced MSPs are added to AWS's Audited MSP Partner program, this designation is becoming less of an assurance of MSP quality. However, since it sets a high minimum bar for MSP

capabilities, it is still the best way to identify partners that use AWS best practices, especially when used in conjunction with the DevOps and Migration Partner competency designations.

- Throughout its history, AWS has near-invisibly updated its service implementations as its scale and customer requirements have changed. Customers may need to make the active choice to gain the full advantage of some modernizations that represent a generational shift in technology, such as the new Nitro offloaded-virtualization platform that is used for some new EC2 instance types. Competitors that copy AWS ideas sometimes have the opportunity to improve on the implementation or service construct.
- AWS is still adapting to the emergence of meaningful competitors — not only cloud providers, but also entrenched competitors in the new markets that it is entering with services that displace existing solutions. Also, as AWS introduces more open-source-compatible services, it has increasingly needed to work with open-source communities in a mutually beneficial fashion, and has begun to change its approach accordingly. Customers' future technology choices are likely to be influenced by AWS's ecosystem relationships.

Google

Google is an internet-centric provider of technology and services. Google has had an aPaaS offering since 2008, but did not enter the cloud IaaS market until Google Compute Engine was launched in June 2012 (with general availability in December 2013).

Offerings: Google Cloud Platform (GCP) is integrated IaaS+PaaS. It combines an IaaS offering (Compute Engine), an aPaaS offering (App Engine) and a range of complementary IaaS and PaaS capabilities, including object storage, a Docker container service (Google Kubernetes Engine) and event-driven "serverless computing" (Google Cloud Functions, in beta). Compute Engine VMs are KVM-virtualized and metered by the second. Enterprise-grade support is extra. It has a multi-fault-domain SLA. Colocation needs are met via partner exchanges (Google Cloud Interconnect).

Locations: Google groups its IaaS data centers into regions, each of which contains at least three zones (data centers). There are East Coast, West Coast and central U.S. regions, as well as regions in Belgium, Japan, Singapore, Germany, Holland, the U.K., India, Australia, Brazil, Canada and Taiwan. Google has a global sales presence. Support is available in English and Japanese. The portal is available in English, Dutch, French, German, Italian, Polish, Spanish, Turkish, Russian, Portuguese, Korean, Japanese, Mandarin, Cantonese and Thai. Documentation is available in English, German, Japanese and Brazilian Portuguese.

Provider maturity: Tier 1. GCP benefits, to some extent, from Google's massive investments in infrastructure for Google as a whole.

Recommended mode: GCP primarily appeals to Mode 2 buyers.

Recommended uses: Big data and other analytics applications, machine learning projects, cloud-native applications, or other applications optimized for cloud-native operations.

Strengths

- Google's strategy for GCP centers on commercializing the internal innovative technology capabilities that Google has developed to run its consumer business at scale, and making them available as services that other companies can purchase. Google's roadmap of capabilities increasingly targets customers with traditional workloads and IT processes, as well as with cloud-native applications. Google has positioned itself as an "open" provider, with a portability emphasis that is centered on open-source ecosystems. Like its competitors, though, Google delivers value through operations automation at scale, and it does not open-source these proprietary advantages.
- GCP has a well-implemented, reliable and performant core of fundamental IaaS and PaaS capabilities — including an increasing number of unique and innovative capabilities — even though its scope of services is not as broad as that of the other market leaders. Google has been most differentiated on the forward edge of IT, with deep investments in analytics and ML, and many customers who choose Google for strategic adoption have applications that are anchored by BigQuery.
- Google can potentially assist customers with the process of operations transformation via its Customer Reliability Engineering program (currently offered directly to a limited number of customers, as well as in conjunction with Pivotal and Rackspace). The program uses a shared-operations approach to teach customers to run operations the way that Google's site reliability engineers do.

Cautions

- GCP positions itself as the cost leader in the market, but its deepest negotiated discounts are usually limited to a single-year contract. Customers evaluating competitive costs should separate the standard discounts (committed and sustained use) from negotiated enterprise discounts, and be aware that GCP discounts are per-service rather than for the overall contract. Google is frequently rigid in

contract negotiations, save for its largest customers, who may receive exceptional flexibility. Its sales force, while technically adept, often has limited experience dealing with enterprise procurement organizations.

- While Google has made significant progress in its efforts to build an ecosystem around its IaaS capabilities, it still only has a small number of experienced MSP and infrastructure-centric professional services partners. GCP's own professional services are treated as profit centers rather than loss-leaders to drive customer implementations. Some prospective customers find that these ecosystem limitations significantly heighten the challenges of adopting GCP, and believe that this adds additional cost and risk.
- Customers sometimes cite ISV licensing and support challenges when adopting GCP, despite Google's 2016 acquisition of the Orbitera software marketplace. In particular, note that Oracle will not normally license or support its software on Google Compute Engine; customers should discuss this directly with Oracle. GCP is increasingly supported by an ecosystem of both commercial and open-source management tools, but the depth of this support varies.

IBM

IBM is a large, diversified technology company with a range of cloud-related products and services. In July 2013, it acquired SoftLayer, an independent web hoster with a focus on small or midsize businesses (SMBs), and in January 2014, it shut down its own SmartCloud Enterprise cloud IaaS offering after migrating its existing customers to SoftLayer. IBM began to absorb the operations of SoftLayer, an IBM company, during 2016, and that process is ongoing. It is phasing out the SoftLayer and Bluemix brands in favor of a general IBM Cloud brand.

At the time of this evaluation, IBM had two portals for cloud IaaS – the IBM Cloud portal (formerly the Bluemix portal) and the SoftLayer portal, each with a distinct set of services. Now, all infrastructure services are available through the IBM Cloud portal (formerly the Bluemix portal), and the SoftLayer portal is simply maintained for legacy customers. Bluemix was originally IBM's PaaS offering and has since expanded into a broader platform for IBM Cloud; in this context, the SoftLayer services were previously branded IBM Bluemix infrastructure, and now simply carry the IBM brand. This Magic Quadrant evaluation considers the customer experience through both portals. We use "SoftLayer infrastructure" to refer to all SoftLayer services (whether cloud or noncloud), regardless of which portal is used to provision and manage them.

Offerings: IBM offers both multitenant and single-tenant Citrix-XenServer-virtualized compute (Virtual Servers), as well as paid-by-the-hour nonvirtualized dedicated servers (Bare Metal Servers). It has OpenStack-based object storage and S3-compatible Cloud Object Storage based on Cleversafe technology. CDN integration is offered via an Akamai partnership. SoftLayer also has noncloud offerings, such as paid-by-the-month dedicated servers (a broader range of configurations than is available per hour) and hosted appliances, but IBM does not make a clear distinction between these offerings and its cloud IaaS capabilities. Bluemix has a Docker-based container service (IBM Cloud Container Service), event-driven "serverless computing" (IBM Cloud Functions, formerly OpenWhisk), a Cloud Foundry-based aPaaS, and other PaaS capabilities. Managed services are optional. Colocation needs are met via partner exchanges (IBM Direct Link).

Locations: SoftLayer infrastructure is located in multiple data centers in the U.S., along with data centers in Canada, Mexico, Brazil, France, Germany, Italy, the U.K., the Netherlands, Norway, Australia, Hong Kong, India, Japan, Korea and Singapore. The IBM Cloud Container Service is located in the U.S., Canada, France, Germany, the Netherlands, the U.K., Australia, Hong Kong, Japan, Korea, Singapore and Brazil.

IBM has a global sales presence. It offers support in the wide range of languages in which IBM does business. The portal and documentation are available in English, French, German, Italian, Portuguese, Spanish, Cantonese, Mandarin, Korean and Japanese.

Provider maturity: Tier 2A. IBM's cloud infrastructure strategy has shifted over time. It has made multiple forays into the cloud IaaS market, and is currently building a new cloud IaaS offering.

Recommended mode: Mode 1. However, IBM Cloud's IaaS capabilities may be used to supplement Mode 2 applications that are primarily implemented on PaaS.

Recommended uses: IBM outsourcing deals that use bare-metal servers as the hosting platform, where the customer has a need for supplemental basic cloud IaaS. SoftLayer infrastructure may also be used as a component of applications built using the Bluemix PaaS capabilities. It should also be considered in circumstances that require both API control over scalable infrastructure and bare-metal servers in order to meet requirements for performance, regulatory compliance or software licensing.

Strengths

- IBM is focused on helping customers with significant IT legacies, especially mainframe customers, gradually begin to take advantage of cloud services. IBM intends its service businesses to assist these customers through the cloud transformation journey.

- IBM has a strong brand and existing customer relationships across the globe, and can offer support in local languages, local contracts and billing in local currency. IBM's base of strategic outsourcing customers help drive cloud-enabled data center outsourcing business into SoftLayer data centers. Its developer ecosystem may help to drive adoption of IBM Cloud infrastructure.
- IBM is in the midst of a Next-Generation Infrastructure (NGI) engineering project that will eventually produce a new set of cloud IaaS offerings, including new compute infrastructure. The work done to enable this new offering has also resulted in addressing some vital capability gaps for SoftLayer infrastructure — such as the introduction of software-defined networking (SDN) capabilities. It has also led to a gradual integration between SoftLayer and the rest of IBM Cloud (formerly Bluemix). The NGI design adopts the principles of hyperscale infrastructure. It is likely to represent a significant step forward in IBM infrastructure capabilities, as well as in IBM's ability to serve the needs of future cloud-native applications, particularly in relation to IBM's broader ambitions in cognitive computing.

Cautions

- The current offering is SoftLayer infrastructure, incorporating some of the work that was necessary to prepare for the rollout of the new NGI compute platform. During 2017, SoftLayer customers experienced several significant outages, some of which were related to upgrades, as well as 2018 downtime related to Spectre/Meltdown patching. IBM is working on improving its ability to conduct nondisruptive maintenance. SoftLayer infrastructure remains SMB-centric and hosting-oriented, and is missing many cloud IaaS capabilities required by midmarket and enterprise customers.
- The IBM Cloud experience remains disjointed, although the integration continues to advance. SoftLayer infrastructure is available in 16 countries, and the container service in 12 countries, but the rest of IBM Cloud (Bluemix) is available in only four countries and in just two U.S. cities. Consequently, customer infrastructure placement options are very limited if they want to use other IBM Cloud services. Note that the four-country limitation applies to some infrastructure capabilities, such as IBM Cloud Functions. SoftLayer and other IBM Cloud services do not share a consistent API, although SoftLayer can now be controlled by the IBM Cloud command line interface (CLI) as well as the legacy SoftLayer CLI. IBM Cloud Private, a Cloud Foundry and Kubernetes service offering that runs on top of existing infrastructure, can facilitate some needs for on-premises hybrid cloud services, but is also a distinct experience.
- IBM is currently unable to publicly supply a beta or general-availability date for new NGI elements. The details of NGI services have not been announced. IBM has, throughout its history in the cloud IaaS business, repeatedly encountered engineering challenges that have negatively impacted its time to market. Customers must thus absorb the risk of an uncertain roadmap. This uncertainty also impacts partners, and therefore the potential ecosystem.

Microsoft

Microsoft is a large and diversified technology vendor that is increasingly focused on delivering its software capabilities via cloud services. Its Azure business was initially strictly PaaS, but Microsoft entered the cloud IaaS market with the launch of Azure Virtual Machines in June 2012 (with general availability in April 2013).

Offerings: Microsoft Azure is integrated IaaS+PaaS. It offers metered-by-the-second Hyper-V-virtualized multitenant compute (Virtual Machines), as well as specialized large instances (such as for SAP HANA). There is multitenant block and file storage, along with many additional IaaS and PaaS capabilities. These include object storage (Blob Storage), a CDN, a Docker-based container service (Azure Container Service), a batch computing service (Azure Batch) and event-driven "serverless computing" (Azure Functions). The Azure Marketplace offers third-party software and services. Cray offers supercomputing infrastructure in Azure data centers. Enterprise-grade support is extra. Colocation needs are met via partner exchanges (Azure ExpressRoute).

Locations: Microsoft groups its Azure data center locations into regions. It is beginning to introduce availability zones (multiple data centers within a region). There are multiple Azure regions in the U.S., Canada, the U.K., France, Germany, Australia, India, Japan and Korea, as well as regions in Ireland, the Netherlands, Hong Kong, Singapore, and Brazil. There are also six regions for the U.S. federal government; two are dedicated to the Department of Defense. (The two Azure China regions are part of a separate service operated by 21Vianet Group, and require a China-specific account.) Microsoft has global sales. Documentation is available in English, French, German, Italian, Spanish, Portuguese, Japanese, Korean, Russian and Mandarin. Support and the service portal are available in those languages, plus Czech, Dutch, Hungarian, Polish, Russian, Swedish and Turkish.

Provider maturity: Tier 1. Microsoft's strong commitment to cloud services has been rewarded with significant market success.

Recommended mode: Microsoft Azure appeals to both Mode 1 and Mode 2 customers, but for different reasons. Mode 1 customers tend to value the ability to use Azure to extend their infrastructure-oriented Microsoft relationship and investment in Microsoft technologies. Mode 2 customers tend to value Azure's ability to integrate with Microsoft's application development tools and technologies, or are interested in integrated specialized PaaS capabilities, such as the Azure Data Lake, Azure Machine Learning or the Azure IoT Suite.

Recommended uses: All use cases that run well in a virtualized environment, particularly for Microsoft-centric organizations.

Strengths

- Microsoft Azure's core strength is its Microsoft heritage – its integrations (both current and future) with other Microsoft products and services, its leverage of the existing Microsoft ISV ecosystem, and its overall strategic importance to Microsoft's future. Azure has a very broad range of services, and Microsoft has steadily executed on an ambitious roadmap. Customers that are strategically committed to Microsoft technology generally choose Azure as their primary cloud provider.
- Microsoft Azure's capabilities have become increasingly innovative and open, with improved support for Linux and open-source application stacks. Furthermore, many customers that are pursuing a multicloud strategy will use Azure for some of their workloads, and Microsoft's on-premises Azure Stack software may potentially attract customers seeking hybrid solutions.
- Microsoft has sustained a very high growth rate over multiple years, and Gartner estimates its end-of-2017 revenue run rate for integrated IaaS+PaaS at more than \$4 billion. Microsoft is leveraging its tremendous sales reach and ability to bundle Azure with other Microsoft products and services in order to drive adoption; Office 365 customers often decide it is most logical to adopt Azure. Microsoft is steadily growing the size of Azure customers; many are beginning to spend more than \$500,000 a year, and a growing number exceed \$5 million in annual spending.

Cautions

- Some Gartner clients with larger-scale implementations have reported significant challenges with Azure adoption; smaller customers may experience the same challenges, but with less severe impact. This is most often because Microsoft's sales, field solutions architects and professional service teams did not have an adequate technical understanding of Azure. Technical support personnel may also lack adequate expertise. The FastTrack program, which provides onboarding support from Azure engineering, results in successful pilots. Customers should use an MSP to execute a more successful implementation, but Microsoft has just begun the process of certifying MSPs; expert, experienced MSPs will be identified in partner directories starting in 3Q18. Many traditional Microsoft partners are trying to transition to supporting Azure, but many do not do so well, so customers should be wary of solicitations from inexperienced partners.
- Microsoft Azure's user experience is optimized to deliver ease of use to novices with simple projects, but this comes at the cost of sometimes making complex configurations difficult and frustrating to implement. Customers shouldn't underestimate the expertise required to properly implement Azure in a performant, reliable and secure fashion. Furthermore, multiple generations of solutions, coupled with unclear guidance on when to use each, create significant complexity in determining the right implementation. Most Azure customers use the portal or CLI for manual management, rather than taking a more automated or DevOps approach. DevOps-oriented customers may encounter frustrations with a lack of strong Azure support in some open-source and other third-party tools and software. Visual Studio Team Services supports an agile development life cycle for Azure and should be a preferred solution.
- Azure does not have best-in-class infrastructure reliability, although reliability has been steadily improving, and Microsoft continues to invest in reducing disruptive maintenance. Customers indicate most reliability issues are related to virtual networks. Most such issues affect individual customers, not a region as a whole. Customer impact can be reduced with meticulous monitoring and use of a high-availability architecture.

Oracle

Oracle is a large, diversified technology company with a range of cloud-related products and services. In late 2015, it launched its first public cloud IaaS offering, the Oracle Cloud Infrastructure Classic (formerly Oracle Compute Cloud Service, the "Gen 1 Cloud"). In November 2016, it launched its next-generation offering, Oracle Cloud Infrastructure (OCI, formerly Oracle Bare Metal Cloud Services, the "Gen 2 Cloud"). In 2016, Oracle purchased Ravello (a cloud service that runs as an overlay on top of third-party clouds as well as Oracle's IaaS), and in 2017, Dyn (a managed DNS provider). Neither is in scope for this Magic Quadrant, but are closely related businesses.

Offerings: The Gen 2 service offers both paid-by-the-hour, KVM-virtualized VMs and bare-metal servers (including a one-click installation and configuration of Oracle Database, RAC and Exadata), a Docker-based container service (Oracle Container Cloud Service), and block, file, and object storage. The Gen 1 service offers paid-by-the-hour, Xen-virtualized VMs, as well as object storage. Oracle Cloud at Customer provides a Gen 1-compatible, on-premises private cloud IaaS offering.

Locations: The Gen 2 data centers are grouped into regions, each of which contains at least three availability domains (data centers); there is a Western U.S. region, a U.S. East Coast region, a U.K. region and a Germany region. The Gen 1 data centers are located in the Central and Eastern U.S., the U.K., Brazil, Australia, Japan and the Netherlands. There is also a Gen 1 government offering. Oracle has global sales.

The Gen 2 service is available only in English. The Gen 1 service is supported and documented only in English, but the service portal is also available in French, German, Italian, Spanish, Russian, Portuguese, Japanese, Korean and Mandarin.

Provider maturity: Tier 2A. Oracle's cloud IaaS strategy has evolved over time. It has made several previous forays into the market, and the Gen 2 offering is still in an early stage of maturity.

Recommended mode: The Gen 2 service will appeal to both Mode 1 and Mode 2 customers, especially those with performance needs that are well-suited to bare-metal servers, and those that do not need more than very basic cloud IaaS capabilities. The Gen 1 service will appeal to Mode 1 customers.

Recommended uses for Gen 2: Oracle application hosting, applications that use Oracle Databases on Exadata, and use cases that require bare-metal servers to be provisioned within minutes.

Recommended uses for Gen 1: General business applications or development environments that require basic cloud IaaS capabilities, for customers that are strategically committed to using Oracle solutions.

Strengths

- Oracle envisions customers running its technologies as an integrated stack. Its cloud strategy is anchored by its applications, database and other middleware, and spans IaaS, PaaS and SaaS. Oracle's cloud IaaS is primarily an infrastructure foundation for its other businesses. Oracle is in the process of migrating its PaaS and SaaS offerings to the new Gen 2 infrastructure, and Gen 2 is also an important component in its future Autonomous Database offerings. Oracle is primarily targeting customers that want to run Oracle software on cloud IaaS, particularly those that prefer to run on bare-metal servers. However, Oracle is also targeting cost-driven startups and small software companies that do not use Oracle technology.
- Oracle's Gen 2 offering has a well-designed hyperscale cloud architecture, a thoughtful selection of current and future features, an emphasis on capabilities related to Oracle workloads, solid reliability and performance, and a competitive price point. Its emerging Container-Native Application Development Platform integrates with a continuous integration/continuous delivery (CI/CD) pipeline derived from its acquisition of Wercker, and has the potential to be technically attractive to cloud-native developers.
- Oracle has a realistic perspective on its late entry into the market, and has a sensible engineering roadmap focused on building the capabilities necessary for targeted use cases. Its engineering and product leadership have been recruited primarily from hyperscale cloud providers. It has been relatively reliable at meeting its forecast engineering timelines.

Cautions

- The Gen 2 offering remains a bare-bones "minimum viable product," and it is arguably too minimal to be viable for a broad range of common cloud IaaS use cases. It has limited enterprise customer traction. Oracle has just begun to build a partner ecosystem, and Gartner clients report that MSPs are sometimes reluctant to support OCI. Customers need to have a high tolerance for risk, along with strong technical acumen.
- The Gen 1 offering is a basic cloud IaaS offering with little in the way of differentiation, and is primarily purchased as a base for Oracle's PaaS offerings. However, it is consistent with Oracle's Cloud Machine private cloud IaaS offering, and thus may be attractive to customers that are interested in a hybrid cloud solution — but such customers also will be aligned to the legacy Gen 1. Oracle is sustaining Gen 1 for the purposes of this solution and other cloud services that rely on Gen 1. Oracle's PaaS and SaaS offerings are otherwise being migrated to Gen 2 infrastructure. The Gen 2 offering will be Oracle's primary cloud IaaS offering going forward, and Gen 1 customers should factor this into their future planning.
- Oracle sometimes uses high-pressure sales tactics to sell its cloud IaaS offerings, including software audits or threatening to dramatically raise the cost of database licenses if the customer chooses another cloud provider. Customers should be cautious of these tactics, thoroughly understand their options and evaluate the risks of adopting OCI at this stage of its maturity. Gartner strongly encourages prospective customers to speak with references.

Vendors Added and Dropped

We review and adjust our inclusion criteria for Magic Quadrants as markets change. As a result of these adjustments, the mix of vendors in any Magic Quadrant may change over time. A vendor's appearance in a Magic Quadrant one year and not the next does not necessarily indicate that we have changed our opinion of that vendor. It may be a reflection of a change in the market and, therefore, changed evaluation criteria, or of a change of focus by that vendor.

Added

No vendors were added.

Dropped

We dropped the following vendors for not fully meeting the 2018 inclusion criteria:

- CenturyLink
- Fujitsu
- Interoute
- Joyent
- Rackspace
- NTT Communications
- Skytap
- Virtustream

Inclusion and Exclusion Criteria

To be included in this 2018 Magic Quadrant, vendors had to demonstrate the following, as of January 2018:

- **Market participation.** They must sell public cloud IaaS as a stand-alone service, without the requirement to use any managed services (including guest OS management), or to bundle it with managed hosting, application development, application maintenance or other forms of outsourcing. They may, optionally, also sell a private version of this offering that uses the same architecture but is single-tenant.
- **Market traction and momentum.** They must be among the top global providers for the relevant segments (public and industrialized private cloud IaaS, excluding small deployments of two or fewer VMs). They must have ISO 27001-audited (or equivalent) data centers on at least three continents. They must have at least one public cloud IaaS offering that meets the following criteria:
 - *If the offering has been generally available for more than three years:* A minimum of \$250 million in 2017 revenue, excluding all managed and professional services; or more than 1,000 customers with at least 100 VMs.
 - *If the offering has been generally available for less than three years:* A minimum of \$10 million in 2017 revenue, excluding all managed and professional services, as well as a growth rate of at least 50% exiting 2017.
- **Business capabilities relevant to Gartner clients.** They must offer the public cloud IaaS service globally (it must be purchasable outside their home region), be able to invoice, offer consolidated billing, and be willing to negotiate customized contracts. They must have 24/7 customer support (including phone support). They must offer the contract, service portal, documentation and support in English (either as the service's default language, or as an optional localization).
- **Technical capabilities relevant to Gartner clients.** They must have a public cloud IaaS service that is suitable for supporting mission-critical, large-scale production workloads, whether enterprise or cloud-native. Specific generally available service features must include:
 - Software-defined compute, storage and networking, with access to a web services API for these capabilities.
 - Cloud software infrastructure services facilitating automated management, including, at minimum, monitoring, autoscaling services and database services.
 - A distributed, continuously available control plane supporting a hyperscale architecture.
 - Real-time provisioning for compute instances (small Linux VM in five minutes, 1,000 Linux VMs in one hour) and a container service that can provision Docker containers in seconds.

- An allowable VM size of at least 16 vCPUs and 128GB of RAM.
- An SLA for compute, with a minimum of 99.9% availability.
- The ability to securely extend the customer's data center network into the cloud environment.
- The ability to support multiple users and API keys, with role-based access control.

The 2018 inclusion criteria were chosen to reflect the key traits that Gartner clients are seeking for strategic cloud IaaS providers, and thus reflect minimum requirements across a range of bimodal use cases.

Evaluation Criteria

Ability to Execute

We evaluated vendors' Ability to Execute in this market by using the following criteria:

- *Product/Service*: Service providers were evaluated on the capabilities of their cloud IaaS offering to support all use cases being evaluated. We evaluated the breadth and depth of the feature set, self-service capabilities, automated system management and suitability to run a broad range of workload types. This criterion is important to buyers that want to purchase the most capable, feature-rich service.
- *Overall Viability (Business Unit, Financial, Strategy, Organization)*: Providers were evaluated on the success of their cloud IaaS business, as demonstrated by current revenue and revenue growth since the launch of their service; their financial wherewithal to continue investing in the business and to execute successfully on their roadmaps; commitment to their current offerings, with no plans to execute disruptive platform transitions or migrations in the next two years; and their organizational commitment to this business, and its importance to the company's overall strategy. This criterion is important to buyers that prefer to purchase services from large vendors with ample financial resources, or from vendors that have a position of market leadership and are continuing to invest aggressively in the business. It is also important to buyers that are concerned about their long-term strategic investment in a particular vendor, or who want to avoid potentially disruptive service changes.
- *Sales Execution/Pricing*: Providers were evaluated on their ability to address the range of buyers for IaaS, including the different audiences in each mode of bimodal IT; adapt to "frictionless selling" with online sales, immediate trials and proofs of concept; provide consultative sales and solutions engineering; be highly responsive to prospective customers; and offer value for money. This criterion is important to buyers that value a smooth sales experience, the right solution proposals and competitive prices.
- *Market Responsiveness and Track Record*: This market is evolving extremely quickly and the rate of technological innovation is very high. Providers were evaluated on how well they have historically been able to respond to changing buyer needs and technology developments, rapidly iterate their service offerings, and deliver promised enhancements and services by the expected time. This criterion is important to buyers that value rapid delivery of cutting-edge capabilities.
- *Marketing Execution*: Providers were evaluated on their mind share and brand awareness in the market; their ability to convey marketing messages based on their ability to deliver real business value, not empty hype or misleading "cloudwashing" (the practice of rebranding or remarketing an existing offering under a cloud label without offering all the attributes of a cloud service); and the clarity and accuracy of their marketing messages, compared with their actual service offering. This criterion is important to buyers that prefer to buy from well-known vendors.
- *Customer Experience*: Providers were evaluated on the quality and responsiveness of their account management and technical support; the ease of use of their self-service functionality; the capabilities of their customer portal (additional functionality such as monitoring, reporting and trouble ticketing); the usefulness of their documentation and customer communications; the quality of their SLAs; the ease of doing business with them; and overall customer satisfaction. This criterion is important to buyers that value the aspects of the vendor relationship and capabilities beyond the IaaS platform itself.
- *Operations*: Providers were evaluated on their ability to meet their goals and commitments, including their track record of service delivery; the quality of their response to outages; their approach to emergency and scheduled maintenance; and their ability to meet timelines that are communicated to customers and to the market. This criterion is important to buyers that want a reliable, predictable service experience.

Our evaluation of a service provider's Ability to Execute remains similar to that of the 2017 Magic Quadrant, with increased expectations across all criteria.

Table 1: Ability to Execute Evaluation Criteria

Evaluation Criteria ↓	Weighting ↓
Product or Service	High
Overall Viability	High
Sales Execution/Pricing	Medium
Market Responsiveness/Record	High
Marketing Execution	Medium
Customer Experience	Medium
Operations	Medium

Source: Gartner (May 2018)

Completeness of Vision

We assessed vendors' Completeness of Vision in this market by using the following criteria:

- **Market Understanding:** Providers were evaluated on their understanding of the wants and needs of three different buying constituencies in this market — enterprises, midmarket businesses and digital businesses (whether technology companies or digital business units embedded in nontechnology businesses) — both currently and in the longer term as the use of IaaS matures. This criterion is important to buyers that value a provider's understanding of the market's evolution and broader business trends, which impact a provider's ability to plan a successful long-term strategy.
- **Marketing Strategy:** Providers were evaluated on their ability to articulate their position in the market and their competitive differentiation, and to communicate these messages clearly and consistently, both internally and externally. This criterion is important to buyers that believe that providers should have a clear focus and direction.
- **Sales Strategy:** Providers were evaluated on their understanding of the buying centers for the market, and the way that these different buying centers want to engage with sales, as well as their strategy for adapting their sales force, online channel and partner channels to the IaaS market. This criterion is important to buyers that value a provider's ability to grow its business over the long term.
- **Offering (Product) Strategy:** Providers were evaluated on the breadth, depth, quality and differentiation of their service roadmaps, as relevant to the use cases under evaluation, with an emphasis on self-service, management capabilities (both traditional and DevOps-oriented), and overall feature set, including cloud software infrastructure services. This criterion is important to buyers that want a provider that will lead the market in service capabilities.
- **Business Model:** Providers were evaluated on their overall value proposition and their strategy for providing solutions for the use cases under consideration, not just raw infrastructure elements. This included evaluating how IaaS fits into their broader product portfolio and product strategy. This criterion is important to buyers that view IaaS as part of an integrated set of solutions from a particular provider.
- **Vertical/Industry Strategy:** Providers were evaluated on their ability to offer targeted services for particular vertical markets, such as government, biotechnology, media and entertainment, and retail. This includes sales and marketing to such verticals, their ability to meet specialized compliance needs, and vertical-specific solutions. This criterion is not directly important to most buyers, except to the extent that a provider has a vertical-specific offering that is relevant to them or meets their specific regulatory compliance requirements.
- **Innovation:** Providers were evaluated on the level of investment in the future of their business, and the quality of those investments, whether financial or human capital; this includes aspects such as the deployment of engineering resources, investments in new technology, mergers and acquisitions, and partnerships and alliances. This criterion is important to buyers that care about leading-edge capabilities, and the strength of a provider's ecosystem.
- **Geographic Strategy:** Providers were evaluated on their ability to expand their offering beyond their home region, serving the needs of multinational businesses, as well as adapting their offerings to other geographies. In particular, this included their strategy for

international sales and support, as well as their data center footprints and internationalization efforts. This criterion is important to buyers that want to use a global vendor.

Our evaluation of Completeness of Vision remains similar to that of the 2017 Magic Quadrant. However, we have continued to increase our expectations for the breadth and depth of a provider's vision, encompassing both technical capabilities and business alliances that create an ecosystem of supporting partners, with a focus on the traits that make a provider suitable for strategic adoption.

Table 2: Completeness of Vision Evaluation Criteria

Evaluation Criteria ↓	Weighting ↓
Market Understanding	High
Marketing Strategy	Medium
Sales Strategy	Medium
Offering (Product) Strategy	High
Business Model	Medium
Vertical/Industry Strategy	Low
Innovation	High
Geographic Strategy	Low

Source: Gartner (May 2018)

Quadrant Descriptions

Leaders

Leaders distinguish themselves by offering a service suitable for strategic adoption and having an ambitious roadmap. They can serve a broad range of use cases, although they do not excel in all areas, may not necessarily be the best providers for a specific need, and may not serve some use cases at all. They have a track record of successful delivery, significant market share and many referenceable customers.

Challengers

Challengers are well-positioned to serve some current market needs. They deliver a good service that is targeted at a particular set of use cases, and they have a track record of successful delivery. However, they are not adapting to market challenges sufficiently quickly, or do not have a broad scope of ambition.

Visionaries

Visionaries have an ambitious vision of the future, and are making significant investments in the development of unique technologies. Their services are still emerging, and they have many capabilities in development that are not yet generally available. While they may have many customers, they might not yet serve a broad range of use cases well.

Niche Players

Some Niche Players may be excellent providers for the use cases in which they specialize, but do not serve a broad range of use cases well or have a broadly ambitious roadmap. Some may have solid leadership positions in markets adjacent to this market, but have only developed limited capabilities in cloud IaaS. Providers that specialize in managed services on top of a "good enough" IaaS platform may be in this category. Finally, some Niche Players have weak offerings, or have cloud IaaS businesses with uncertain futures, and should only be chosen with careful attention to managing vendor-related risks.

Context

When people think about "cloud computing," cloud IaaS is often one of the first things that comes to mind. It's the "computing" in cloud computing – on-demand compute, storage and network resources, delivered on-demand, in near-real time, as a service. The market is maturing rapidly; IaaS is on the Slope of Enlightenment on Gartner's "Hype Cycle for Cloud Computing, 2017." However, because the market has consolidated around the hyperscale market leaders – particularly Amazon Web Services and Microsoft Azure – many of the other competitors now face significant business challenges, and the customers of those competitors now face significant supplier-related risks.

The stakes involved in this market are increasing because the relevant total addressable market size is increasing. Cloud IaaS and PaaS increasingly represent a continuum of integrated services delivered by a single provider, and the leading cloud IaaS providers also have strong PaaS capabilities. IaaS and PaaS represent a spectrum of offerings that balance greater control and customization against greater ease of management and developer productivity. Most customers that adopt the infrastructure resources within a cloud IaaS offering will also adopt associated management services, such as monitoring, and are highly likely to adopt PaaS-level capabilities, such as database as a service, over time.

Consequently, the value proposition of cloud IaaS is no longer simply compute and storage capabilities delivered on-demand, but rather a complete infrastructure platform that delivers both efficiency and agility, combined with unprecedented scalability and global presence. This market direction favors the incumbent market leaders, and significantly raises the barriers for other vendors trying to gain traction in the market.

Cloud IaaS has broad, mainstream adoption across a wide variety of use cases. While most businesses initially adopted cloud IaaS for Mode 2, agile IT projects, an increasing number of organizations are now migrating Mode 1, safety-and-efficiency-oriented applications – and even entire data centers – to cloud IaaS. Cloud IaaS is increasingly critical not only to digital business, but also to IT modernization and transformation initiatives. Cloud IaaS can now be used for nearly all use cases that can be reasonably hosted on virtualized x86-based servers; the question is no longer, "Is cloud IaaS a *viable* solution for my application?" but rather, "Is cloud IaaS the *best possible* solution for my application?" Furthermore, cloud IaaS is now a viable alternative to running an internal data center, but it is not the right decision for everyone (see "15 Reasons Not to Migrate Your Data Center to Public Cloud Infrastructure as a Service").

Bimodal IT impacts cloud IaaS sourcing decisions. Mode 2, agile IT organizations typically value cloud IaaS providers that invest deeply in engineering in order to provide a rich suite of features and extensive automation for self-service enablement. Mode 2 adoption is often business-led – driven by business managers who hold the budget, need greater agility and have shorter time frames than I&O organizations are able to accommodate. They therefore turn to application developers and enterprise architects for a solution. I&O organizations typically have a Mode 1 mindset. They may initially look for service providers that provide a basic set of IaaS features within a familiar environment that is similar to their existing virtualized infrastructure. However, they are likely to rethink this approach if their ultimate goal is IT transformation. Cloud IaaS providers vary in their ability to target these different buying centers. Furthermore, most providers focus on either a Mode 1 or Mode 2 audience, and their feature set and style of service are oriented accordingly, although leading providers offer capabilities attractive to both audiences.

Most organizations now choose one or two long-term strategic partners for cloud IaaS, although they may still use other cloud IaaS providers in a tactical fashion. Most organizations make the choice of which of these cloud IaaS providers to use on a per-project basis. Typically, though, one of the providers is the primary strategic partner, and other providers are only used when they are a significantly better fit for the project in question.

Market Overview

The cloud IaaS market is consolidating rapidly. Customer expectations continue to escalate. Customers now expect cloud IaaS providers to offer a very broad array of services that encompass hardware infrastructure, software infrastructure, management and governance capabilities, and preintegrated value-added solutions. Moreover, they expect that providers will have robust ecosystems for managed and professional services, as well as a software and tools ecosystem. Providers must either achieve both breadth and depth of capabilities with a global platform deployed at scale, or be forced into a niche.

The 2018 Magic Quadrant contains only six cloud IaaS providers. Some of these providers currently provide a broad array of cloud infrastructure capabilities, while others simply have the ambition to eventually do so. This is a reduction from 2017's 15 vendors, reflecting the consolidating market and heightened buyer expectations that led to more stringent inclusion criteria.

What Does the Cloud IaaS Market Include?

Cloud IaaS provides on-demand, near-real-time, self-service access to abstracted, programmatically accessible and highly automated infrastructure resources (at minimum, compute resources, along with associated storage and network resources), on-demand and in near real time. In IaaS, the provider manages the data center facilities, hardware and virtualization. Everything above the hypervisor layer, however – the operating system, middleware and application – is managed by the customer, or is an add-on managed service from the provider or another third party.

This market is wholly separate and distinct from cloud SaaS, but is increasingly entangled with the PaaS market. Cloud IaaS providers are increasingly offering middleware and other software infrastructure capabilities as a service, as well as services that provision and orchestrate application containers (particularly Docker containers). Customers want to develop, deploy and manage applications efficiently, and will choose the combination of IaaS and PaaS capabilities that best suits their needs — and often, neither customers nor providers will make a definitional distinction between IaaS and PaaS. To make it easy for applications to span this spectrum of capabilities, an integrated IaaS+PaaS provider needs a single self-service portal and catalog, common identity and access management, an integrated low-latency network context, and an integrated security context.

Cloud IaaS is owned, built and operated by a service provider, but it may be delivered on-premises within a customer's data center or hosted in the provider's data center. It may be "public" (multitenant) or "private" (single-tenant), although, in practice, there is no consistency in the application of these labels to varying degrees of resource isolation, and most hosted offerings use some degree of shared resources in services labeled "private."

Cloud IaaS is not a commoditized service, and even providers with very similar offerings and underlying technologies often have sufficiently different implementations that there is a material difference in availability, performance, security and service features. As a result, risks related to vendor lock-in or application portability need to be thoughtfully managed. (See "Addressing Lock-In Concerns With Public Cloud IaaS" for details.)

What Types of Workload Are Being Placed on Cloud IaaS?

There are four broad categories of customer need in cloud IaaS:

- Digital business enablement
- Mode 2, agile IT projects
- Mode 1, traditional IT data center substitution
- Batch computing

Digital business needs account for the majority of workloads in cloud IaaS. Digital business, however, is not limited to technology companies. Almost every business is being impacted by digital disruption, and an increasing number of businesses have "internal startups" or digital business units. (See "Building and Expanding a Digital Business Primer for 2018.") Digital business use cases are very broad, and include digital marketing, e-commerce, e-CRM, SaaS, data services, big data analytics and IoT applications. These are generally production applications, although cloud IaaS is typically used for the whole application life cycle. Many of these customers have mission-critical needs.

In addition to digital business projects, many organizations have a wide variety of IT projects that they are executing in an agile fashion. Rapid application development, prototyping, experiments and other IT projects that require agility, flexibility and the ability to meet urgent infrastructure needs are frequently executed on cloud IaaS. Although most such Mode 2, agile IT projects are not core to the organization's overall IT portfolio, they may have high visibility and high business impact.

While these Mode 2 needs are still the primary driver of customer adoption, Mode 1 use cases are now common. In many organizations, cloud IaaS is gradually replacing or supplementing traditional data center infrastructure. It is typically used very similarly to the organization's internal virtualization environment. Organizations typically begin with development environments or less-mission-critical production applications, but gradually expand to also host mission-critical applications on cloud IaaS. Mode 1, traditional IT organizations typically look to cloud IaaS to deliver cost reductions, but may also be interested in long-term IT transformation. (See "Three Journeys Define Migrating a Data Center to Cloud Infrastructure as a Service" for details.)

The least common need, but one that nevertheless generates significant revenue and accounts for significant cloud infrastructure consumption, is batch computing, which may be either Mode 1- or Mode 2-oriented. For these customers, IaaS serves as a substitute for traditional HPC or grid computing. Customer needs include rendering, video encoding, genetic sequencing, modeling and simulation, and numerical analysis. These customers need to access large amounts of commodity compute at the lowest possible price, with little concern for infrastructure reliability. Some HPC use cases benefit from specialized hardware such as graphics processing units (GPUs) and high-speed interconnects.

Cloud IaaS can now be used to run most workloads, although not every provider can run every type of workload well. Service providers are moving toward infrastructure platforms that can offer physical (nonvirtualized) and virtual resources, priced according to the level of availability, performance, security and isolation that the customer selects. This allows customers to run "cloud native" applications — those architected with cloud-native principles and design patterns in mind (see "How to Architect and Design Cloud-Native Applications") — and

migration of existing business applications from their own virtualized servers in internal data centers into the cloud, without changes. Cloud IaaS is best used to enable new IT capabilities, but it has become a reasonable alternative to an internal data center.

What Trends Are Currently Influencing Buyers?

Customer expectations continued to increase significantly over the course of 2017. Customers now have high expectations from their cloud IaaS providers. They demand market-leading technical capabilities — depth and breadth of features, along with high availability, performance and security. They expect not only "hardware" infrastructure features, but also management features, developer services and cloud software infrastructure services, including fully integrated PaaS capabilities. Customers also expect that a service provider has the ability to engage successfully with enterprises that have mission-critical production workloads, large-scale migrations to the cloud, or digital transformation initiatives. Further, they expect that the provider also has an ISV, consulting partner and managed service provider (MSP) ecosystem that is capable of supporting such needs. Cloud IaaS providers that do not meet these expectations are typically adopted only for narrow, specialized use cases.

Integrated management capabilities and developer services are key generators of value. One of the primary generators of both greater agility and greater efficiency is automation — specifically the ability to reduce the burden of operations management and to empower developers to self-serve. The more management capabilities are integrated across the cloud provider's portfolio, the more easily system and application infrastructure can both be managed as a single entity, regardless of whether a traditional or DevOps style of operations is used. Developer services and other cloud software infrastructure services hide most infrastructure management concerns from developers, and significantly reduce operations-related labor. When the provider preintegrates all the components across its platform, the customer receives more value at less cost than if the customer were to choose multivendor components and integrate those components themselves.

Most customers have a multicloud strategy. Most customers choose a primary strategic cloud IaaS provider, and some will choose a secondary strategic provider as well. They may also use other providers on a tactical basis for narrow use cases. While it is relatively straightforward to move VM images from one cloud to another, cloud IaaS is not a commodity. Customers choose to adopt multiple providers in order to have a broader array of solutions to choose from. Relatively few customers use multicloud architectures (where a single application or workload runs on multiple cloud providers), as these architectures are complex and difficult to implement. (See "Assessing the Strengths and Weaknesses of High-Value IaaS and PaaS Multicloud Use Cases" for details.)

Managing multiple cloud IaaS providers is challenging. Many organizations are facing the challenge of creating standardized policies and procedures, repeatable processes, governance, and cost optimization across multiple cloud providers. "Single pane of glass" management, seamless movement across infrastructure platforms and "cloudbursting" are unlikely to become reality, even between providers using the same underlying CIF or with use of portable application container technology. Note that the claim that an ecosystem is "open" has nothing to do with actual portability. Due to the high degree of differentiation between providers, the organizations that use cloud IaaS most effectively will embrace cloud-native management, rather than allow the legacy enterprise environment to dictate their choices.

Customers frequently use third-party management tools for governance, especially multicloud governance. Most customers that make substantive use of cloud IaaS supplement the native management capabilities of the providers with third-party management tools. These tools can be very helpful for governance functions, and may be designed for single-cloud or multicloud use. If multicloud, the tool should support integrated cost management, identity and access management, security and compliance reporting, and networking. Management tools cover a wide range of possible functions. Examples include:

- Cloud management platforms (CMPs), such as Cisco CloudCenter (formerly CliQr), RightScale and Scalr (see "Market Guide for Cloud Management Platforms")
- Cloud service expense management (CSEM) tools, such as CloudHealth and Cloudability (see "How to Identify Solutions for Managing Costs in Public Cloud IaaS")
- Continuous configuration automation tools, such as HashiCorp's Terraform (see "Market Guide for Agile and DevOps Services")
- Cloud workload protection platforms, such as Dome9 (see "Market Guide for Cloud Workload Protection Platforms")
- Cloud infrastructure security posture assessment tools, such as Evident.io (see "Cloud Security Primer for 2018")

Application platform strategy, and the relevant vendor relationships, are important to many customers. Very few customers prioritize their relationships with their existing IT infrastructure vendors when choosing a cloud IaaS provider. However, most customers consider their application platform strategy, and the related vendors, when choosing a cloud IaaS provider. For instance, customers that use Microsoft development tools and Microsoft middleware are much more likely to choose Microsoft Azure for those applications, and their overall

application strategy may be more aligned to Microsoft's strategic direction for Azure. Conversely, a customer may align their future application strategy to their chosen cloud provider's capabilities. This is especially true if the customer is migrating from commercial application infrastructure to an open-source application stack, whether obtained as software or as a cloud service.

Ecosystems are vital. Customers need their cloud IaaS providers to have strong ecosystems that can successfully support a broad array of customer use cases. ISVs need to license and support their software on the cloud IaaS offering. Open-source software needs to be readily available. Tools, especially ITOM and DevOps tools, must be integrated with the cloud provider's API. Professional services must be available for migration, application integration and application development on the platform. Managed services need to be delivered in an expert fashion, which requires experience as well as supporting automation tools. While some customers will use global partners, many customers need local or regional partners, especially for managed and professional services.

Managed and professional services greatly increase the likelihood of a successful cloud IaaS implementation. Such services should be sourced in a bimodal fashion. Some MSPs specialize in cloud-native operations, usually with significant use of DevOps, and can help customers through the transformation process. This may be attractive to both Mode 1 and Mode 2 customers, as well as digital businesses (see "Use Managed and Professional Services to Improve Cloud Operations for Digital Business"). Mode 1 data center migrations also benefit strongly from managed and professional services, even if the approach is not cloud-native (see "Three Journeys Define Migrating a Data Center to Cloud Infrastructure as a Service"). See "How to Choose a Managed Service Provider for a Hyperscale Cloud Provider" for MSP selection guidance, and the "Magic Quadrant for Public Cloud Infrastructure Managed Service Providers" for a market evaluation.

"Lift and shift" migrations rarely achieve the desired business outcomes. Most customers who simply treat cloud IaaS like "rented virtualization" do not achieve significant cost savings, increased operational efficiency or greater agility. It is possible to achieve these outcomes with a "lift and optimize" approach — cloud-enabled virtual automation — in which the applications do not change, but the IT operations management approach changes to be more automated and cloud-optimized. Customers who execute a lift-and-shift migration often recognize, after a year, that optimization is needed. Gartner believes it is more efficient to optimize during the migration rather than afterward, and that customers typically achieve the best outcomes by adopting the full range of relevant capabilities from a hyperscale integrated IaaS+PaaS provider.

Customers do not always save money by using cloud IaaS. Although many customers first investigate using IaaS to achieve cost savings, most customers buy IaaS to achieve greater business agility or to access infrastructure capabilities that they do not have within their own data center. IaaS can drive significant cost savings when customers have short-term, seasonal, disaster recovery or batch-computing needs. It can also be a boon to companies with limited access to capital and to small companies — especially startups — that cannot afford to invest in infrastructure. For larger businesses with existing internal data centers, well-managed virtualized infrastructure, efficient IT operations teams and a high degree of automation, IaaS for steady-state workloads is often no less expensive, and may be more expensive, than an internal private cloud. The less efficient your organization, the more likely you are to save money by using a cloud provider, especially if you take advantage of this opportunity to streamline and automate your operations. The largest-scale providers are continually lowering their prices, and automated managed services will substantially drive down the cost of infrastructure management over time, so cost advantages will continue to accrue to the providers. (See "Can You Save Money Migrating to Cloud IaaS?" for guidance.)

Single-tenant options in public cloud IaaS are preferred over hosted private cloud IaaS. Increasingly, public cloud IaaS providers have introduced options for single-tenant compute instances and single-tenant storage. Many customers are able to meet regulatory compliance or software licensing requirements that require single tenancy in this fashion. Because public cloud IaaS offerings typically have a much broader, deeper feature set than hosted private cloud IaaS, customers have primarily turned to single-tenant options in public cloud IaaS, or supplemented public cloud IaaS with colocated equipment. As a result, few providers have found significant success in hosted private cloud IaaS.

On-premises private cloud IaaS is rarely successful. Most IT organizations that have tried to implement on-premises private cloud IaaS have met with limited success. Most are turning to alternatives — enhancing their virtualized infrastructure with additional automation, building containerized infrastructure, deploying PaaS frameworks on-premises or purchasing private PaaS, or using public cloud IaaS. Notably, the providers evaluated on this Magic Quadrant that have private cloud IaaS offerings also deliver PaaS as part of those offerings, and customers frequently choose these offerings specifically to obtain the PaaS capabilities.

Serverless computing is most easily adopted via public cloud services. The serverless model of computing abstracts the underlying infrastructure and hides many management considerations from the application developer. Integrated IaaS+PaaS providers are offering an increasing array of serverless capabilities — including but not limited to function PaaS (fPaaS) offerings, such as AWS Lambda and Azure Functions. Few enterprises have the wherewithal to implement such capabilities themselves. These capabilities empower developers and reduce operational effort, and are vital to optimizing for digital business agility. As such, they will continue to help drive agility-focused customers toward public cloud services, but cost-efficiency-driven customers may also be able to achieve some cost savings by migrating middleware components to serverless cloud software infrastructure services. (See "Predicts 2018: Compute Infrastructure" for details.)

Public cloud IaaS providers are beginning to deliver comprehensive solutions for microservices infrastructure. Microservices infrastructure is a composite of the various types of application infrastructure technologies used to build, deploy, run and manage microservices and miniservices. It includes microservices runtime infrastructure, development frameworks and life cycle automation technologies. Public cloud providers are increasingly delivering integrated solutions across all three of these aspects. Managed container systems – which include container IaaS, high-control aPaaS and fPaaS – are a key component. Most microservices infrastructure technologies are immature, and are among the most innovative capabilities delivered by cloud providers. (See "Innovation Insight for Microservice Infrastructure" for details.)

Customers frequently acknowledge that public cloud IaaS providers offer superior security capabilities when compared to their own data centers. Although many security controls are the responsibility of the customer, not the provider, most major cloud IaaS providers offer a high degree of security on the underlying platform. Transparent encryption of LAN, WAN and storage will become increasingly commonplace as a bundled element of cloud IaaS offerings, as providers react to defend themselves against intrusion from sophisticated attackers. (See "Take a Risk-Based Approach to Public Cloud IaaS" for guidance.) Gartner strongly recommends the adoption of a DevSecOps approach to managing security in cloud IaaS environments. (See "10 Things to Get Right for Successful DevSecOps" and "Integrating Security into the DevSecOps Toolchain" for implementation advice.)

What Key Market Aspects Should Buyers Be Aware Of?

The global market remains consolidated around two clear leaders. The market consolidated dramatically over the course of 2015. Since 2016, just two providers – AWS and Microsoft Azure – have accounted for the overwhelming majority of the IaaS-related infrastructure consumption in the market, and their dominance is even more thorough if their PaaS-related infrastructure consumption is included as well. Furthermore, AWS is many times the size of Microsoft Azure, further skewing the market structure. Most customers will choose one of these leaders as their strategic cloud IaaS provider.

Chinese cloud providers have gone global, but still have limited success outside of the domestic Chinese market. The sheer potential size of the market in mainland China has motivated multiple Chinese cloud providers to build a broad range of capabilities; such providers are often trying to imitate the global leaders feature-for-feature. While this is a major technological accomplishment, these providers are primarily succeeding in their domestic market, rather than becoming global leaders. Their customers are currently China-based companies, international companies that are doing business in China and some Asia/Pacific entities that are strongly influenced by China.

The remainder of the market is highly fragmented. Despite the thorough dominance of two market leaders, there are still thousands of service providers that offer cloud IaaS. Some of these are managed hosting providers or local managed service providers, for whom cloud IaaS is simply an infrastructure platform and a means to an end. Many such providers are also pivoting to offer their managed services on third-party cloud IaaS offerings. There are also many virtual private server (VPS) hosting providers that serve small businesses and have successful cloud VPS offerings; many such providers serve local markets or a single country. However, such providers typically have highly limited capabilities, and most have no supporting ecosystem.

Local sourcing matters to some customers, but such offerings are limited. Customers normally prefer to keep data in-region for reasons of network latency. However, regulatory concerns that require keeping data in-country, as well as revelations about foreign intelligence agencies obtaining access to private data, have heightened the desire of non-U.S.-based customers to purchase cloud IaaS from local providers. (See "The Snowden Effect: Data Location Matters" and "Transfer Personal Data Worldwide.") Unfortunately, local providers typically lack the scale and capabilities of the global providers, and may focus primarily on small businesses, not enterprises. Furthermore, keeping data local is no guarantee of freedom from either domestic or foreign surveillance. Customers outside the U.S. that cannot use a foreign provider (even when that provider has local presence) are likely to have access to only basic, commodity capabilities. The lack of a high-quality local offering may significantly slow cloud IaaS adoption in a country.

Beware of new offerings. Several cloud providers introduced new or significantly altered cloud IaaS platforms in 2016 or 2017, or are in the process of doing so during 2018. These new service offerings usually have a minimalistic feature set, may have poor operational reliability and lack a supporting ecosystem; even offerings that are more than two years old may not have established a track record of enterprise customer success. Nevertheless, many such providers are aggressively pursuing new customers for these platforms, especially if they have existing customer relationships, and they may be willing to offer generous discounts in order to win customers. Prospective customers of these providers need to carefully manage vendor risks. They should ensure that they speak with reference customers that are similar to them in organization type, IT management style, software development life cycle, workload type, implementation size and tolerance for risk.

There are significant barriers to becoming a successful hyperscale integrated IaaS+PaaS provider. The cloud providers evaluated in this Magic Quadrant either have hyperscale integrated IaaS+PaaS offerings or are currently developing such offerings. Entering this market requires an enormous initial investment, and a provider has to commit to further deep investments in software development as well as infrastructure deployment for years to come. It also requires deep engineering expertise – including the ability not only to deliver

technology that has never existed before, but also to do so on time and within budget. Finally, entering the hyperscale integrated IaaS+PaaS market requires the ability to build an ecosystem. Potential partners prefer to devote the bulk of their resources to working with market leaders, which makes it harder for later entrants to gain traction with partners. Newer entrants may still be relegated to niches.

Specialized providers can be the right choice for some use cases. Cloud IaaS providers that have built offerings to serve highly targeted use cases, especially application-specific use cases, can be highly successful within their niche. Specialized providers typically only have a basic cloud IaaS feature set, without extensive management or cloud software infrastructure services. However, the capabilities they have are tailored to the use cases they serve, and they may have additional use-case-specific capabilities. A deeply differentiated feature set aimed at a particular use case can set a provider apart from the rest of the market, enabling it to win deals even when a customer chooses one or more strategic providers for general-purpose workloads.

APIs anchor a partner ecosystem. Programmatic (API) access to infrastructure is crucial, as it enables customers, as well as third parties, to build management tools for their platforms, and to enable applications to take maximum advantage of the infrastructure environment. Providers need to foster rich ecosystems of capabilities. The leading providers are likely to build a substantial number of capabilities themselves. Meanwhile, partners will extend the range of their capabilities, provide overlays for complex heterogeneous multivendor environments, and add "stickiness" to these platforms by offering tight integrations between applications, middleware and infrastructure. Furthermore, cloud IaaS providers that are launching new platforms and hope to "catch up" to the market leaders will be highly dependent upon partners that can supply missing capabilities. Yet the trend is toward proprietary APIs, rather than "open" APIs, such as OpenStack.

Cloud IaaS is not a commodity. Providers vary significantly in their features, performance, cost and business terms. Although in theory, cloud IaaS has very little lock-in — a VM is just a VM, in the end — in truth, cloud IaaS is not merely a matter of hardware rental, but an entire data center ecosystem as a service. This encompasses the entirety of the ITOM stack, including traditional IT service management capabilities and automation that reduce the burden of operational chores, such as patching and backups; DevOps-oriented capabilities; and new forms of automation, analytics and insight (including "smart" infrastructure capabilities) that take advantage of the unique perspective offered by the delivery of integrated compute, storage and networking resources. The more you use those capabilities, the more value you will receive from the offering, but the more you will be tied to that particular service offering. The dynamics of this market resemble a software market, not a traditional IT services market. Providers are in a race to deliver features, and the "winners" are likely to be those that are highly innovative and that have the most resources to invest in the breadth and depth of capabilities development.

Evidence

- Gartner client inquiries in 2017 and 2018 (currently more than 1,000 cloud IaaS-related inquiries per quarter)
- Service provider interviews and product demonstrations in 2017 and 2018
- Surveys of more than 75 cloud IaaS providers in 2017 and 2018
- Customer references from the service providers in 2017 and 2018, as well as Gartner Peer Insights data
- Hands-on trials of service offerings in 2017 and 2018
- Performance benchmarks and monitoring of service offerings via Gartner's Cloud Decisions service, throughout 2017 and 2018
- Public information from sources such as U.S. Securities and Exchange Commission filings, press releases, vendor websites and community support forums

Note 1 SSAE 16

Statement on Standards for Attestation Engagements (SSAE) 16 — that is, Service Organization Control (SOC) 1. See "Market Guide for Organization Security Certification Services."

Note 2 ISO 27001

International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) 27001. See "Security Research Roundup for ISO 27001 Compliance."

Evaluation Criteria Definitions

Ability to Execute

Product/Service: Core goods and services offered by the vendor for the defined market. This includes current product/service capabilities, quality, feature sets, skills and so on, whether offered natively or through OEM agreements/partnerships as defined in the market definition and detailed in the subcriteria.

Overall Viability: Viability includes an assessment of the overall organization's financial health, the financial and practical success of the business unit, and the likelihood that the individual business unit will continue investing in the product, will continue offering the product and will advance the state of the art within the organization's portfolio of products.

Sales Execution/Pricing: The vendor's capabilities in all presales activities and the structure that supports them. This includes deal management, pricing and negotiation, presales support, and the overall effectiveness of the sales channel.

Market Responsiveness/Record: Ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve and market dynamics change. This criterion also considers the vendor's history of responsiveness.

Marketing Execution: The clarity, quality, creativity and efficacy of programs designed to deliver the organization's message to influence the market, promote the brand and business, increase awareness of the products, and establish a positive identification with the product/brand and organization in the minds of buyers. This "mind share" can be driven by a combination of publicity, promotional initiatives, thought leadership, word of mouth and sales activities.

Customer Experience: Relationships, products and services/programs that enable clients to be successful with the products evaluated. Specifically, this includes the ways customers receive technical support or account support. This can also include ancillary tools, customer support programs (and the quality thereof), availability of user groups, service-level agreements and so on.

Operations: The ability of the organization to meet its goals and commitments. Factors include the quality of the organizational structure, including skills, experiences, programs, systems and other vehicles that enable the organization to operate effectively and efficiently on an ongoing basis.

Completeness of Vision

Market Understanding: Ability of the vendor to understand buyers' wants and needs and to translate those into products and services. Vendors that show the highest degree of vision listen to and understand buyers' wants and needs, and can shape or enhance those with their added vision.

Marketing Strategy: A clear, differentiated set of messages consistently communicated throughout the organization and externalized through the website, advertising, customer programs and positioning statements.

Sales Strategy: The strategy for selling products that uses the appropriate network of direct and indirect sales, marketing, service, and communication affiliates that extend the scope and depth of market reach, skills, expertise, technologies, services and the customer base.

Offering (Product) Strategy: The vendor's approach to product development and delivery that emphasizes differentiation, functionality, methodology and feature sets as they map to current and future requirements.

Business Model: The soundness and logic of the vendor's underlying business proposition.

Vertical/Industry Strategy: The vendor's strategy to direct resources, skills and offerings to meet the specific needs of individual market segments, including vertical markets.

Innovation: Direct, related, complementary and synergistic layouts of resources, expertise or capital for investment, consolidation, defensive or pre-emptive purposes.

Geographic Strategy: The vendor's strategy to direct resources, skills and offerings to meet the specific needs of geographies outside the "home" or native geography, either directly or through partners, channels and subsidiaries as appropriate for that geography and market.

be reproduced or distributed in any form without Gartner's prior written permission. If you are authorized to access this publication, your use of it is subject to the [Gartner Usage Policy](#) posted on gartner.com. The information contained in this publication has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information and shall have no liability for errors, omissions or inadequacies in such information. This publication consists of the opinions of Gartner's research organization and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice. Although Gartner research may include a discussion of related legal issues, Gartner does not provide legal advice or services and its research should not be construed or used as such. Gartner is a public company, and its shareholders may include firms and funds that have financial interests in entities covered in Gartner research. Gartner's Board of Directors may include senior managers of these firms or funds. Gartner research is produced independently by its research organization without input or influence from these firms, funds or their managers. For further information on the independence and integrity of Gartner research, see "[Guiding Principles on Independence and Objectivity](#)."

[About](#) [Careers](#) [Newsroom](#) [Policies](#) [Site Index](#) [IT Glossary](#) [Gartner Blog Network](#) [Contact](#) [Send Feedback](#) 

© 2018 Gartner, Inc. and/or its Affiliates. All Rights Reserved.