(http://www.gartner.com/home)

LICENSED FOR DISTRIBUTION

Critical Capabilities for Public Cloud Infrastructure as a Service, Worldwide

26 October 2015 | ID:G00270178

Analyst(s): Lydia Leong

Summary

Public cloud IaaS offerings are not commodities. Buyers must choose an offering that matches their use case and specific needs. We compare 15 public cloud IaaS services against eight critical capabilities in four use cases.

Overview

Key Findings

Most public cloud infrastructure as a service (IaaS) offerings can capably deliver virtual machines (VMs), along with the basic storage and networking capabilities associated with those compute resources. However, the depth, breadth, quality and manageability of these capabilities vary significantly. The best providers have an extensive array of additional services that extend across the spectrum from IaaS to platform as a service (PaaS).

The most capable public cloud IaaS offerings can be successfully used for both new and existing applications, and they can serve the needs of both developers and IT operations organizations; they enhance both developer productivity and operator efficiency, and they may enable IT transformation. The value of less feature-rich offerings is likely to be primarily in efficient provisioning and a shift to operational expenditures from capital expenditures.

Some service providers in this evaluation are in a state of upheaval. They have not truly achieved scale, have met with limited commercial success, or have not made the deep investments needed to drive automation, improve their capabilities and compete successfully with the market leaders. Some will exit the market, pivot, or build or acquire a new platform.

Recommendations

Adopt a bimodal IT sourcing strategy for cloud IaaS. Ensure that you meet the needs of developers and other technical end users who consume cloud IaaS, not just the needs of the infrastructure and operations organization.

Try several offerings before committing to any one service; this is the only way to get a feel for the nuances of depth, breadth, quality, manageability, user experience and cost. Many public cloud IaaS offerings can be bought by the hour, with no contractual commitment.

Assume that cloud IaaS offerings are not interchangeable, and that where you place a workload will be where it stays. Although there can be relatively little "lock-in" for public cloud IaaS, moving between providers is similar to doing a data center move in that it can be time-consuming, expensive and risky. Consider the strategic future of a provider before migrating a significant percentage of your applications into its cloud.

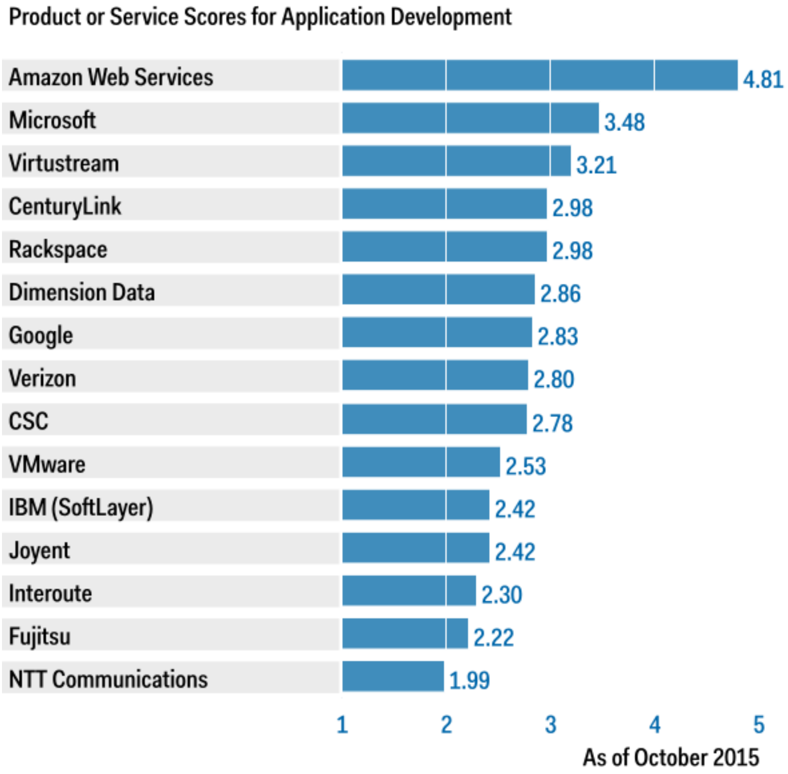
What You Need to Know

Public cloud IaaS can be used for most workloads that can run in a virtualized x86-based server environment. IT leaders should ensure that they select offerings that meet the needs of developers and other technical end users of cloud IaaS solutions, not just the needs of the IT operations organization. IT leaders should also keep in mind that the most effective way to adopt cloud IaaS usually requires embracing DevOps, taking advantage of the provider's proprietary capabilities, and an agile and transformative approach to IT; using cloud IaaS as "rented virtualization" may be of limited benefit. This Critical Capabilities report compares providers in the context of their ability to deliver value for four common use cases for public cloud IaaS — application development, batch computing, cloud-native applications and general business applications. It should help you draw up a shortlist of appropriate providers for the public cloud IaaS use cases in your organization. While it is still appropriate to choose providers in a use-case-specific way, most organizations will choose one or two providers for strategic adoption across multiple use cases.

Analysis

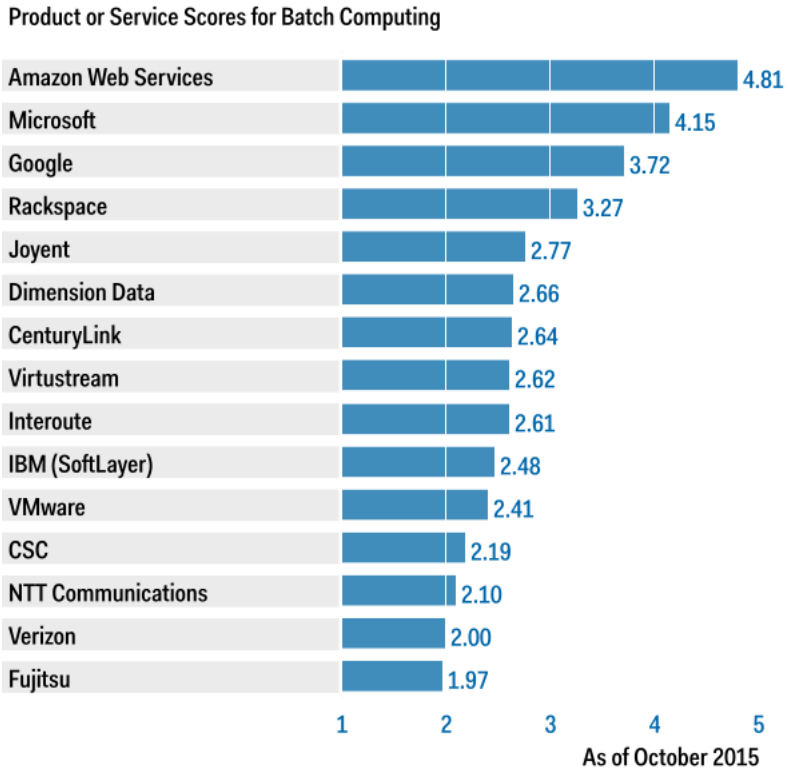
Critical Capabilities Use-Case Graphics

Figure 1. Vendors' Product Scores for Application Development Use Case



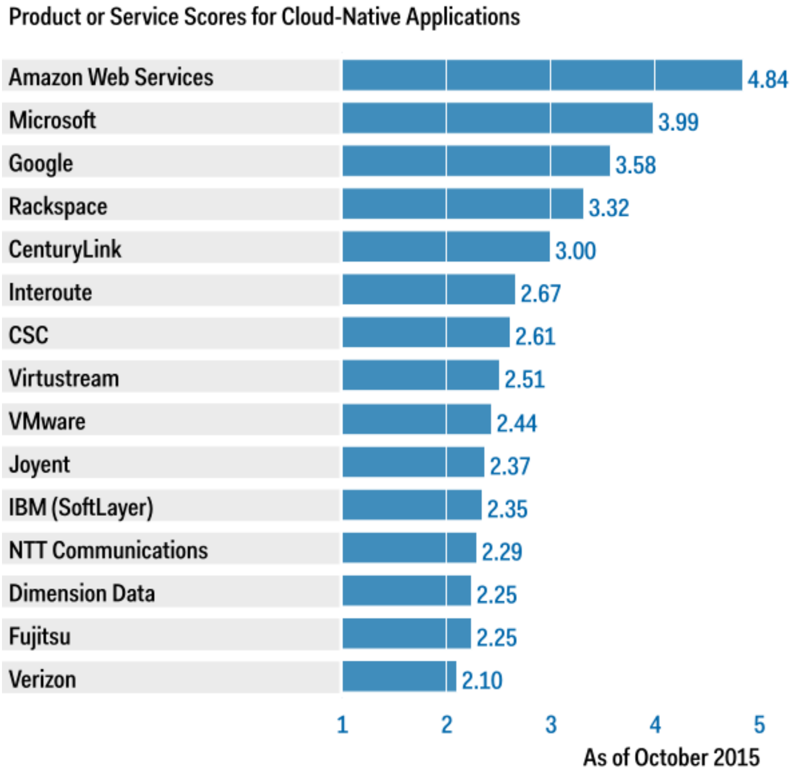
Source: Gartner (October 2015)

Figure 2. Vendors' Product Scores for Batch Computing Use Case



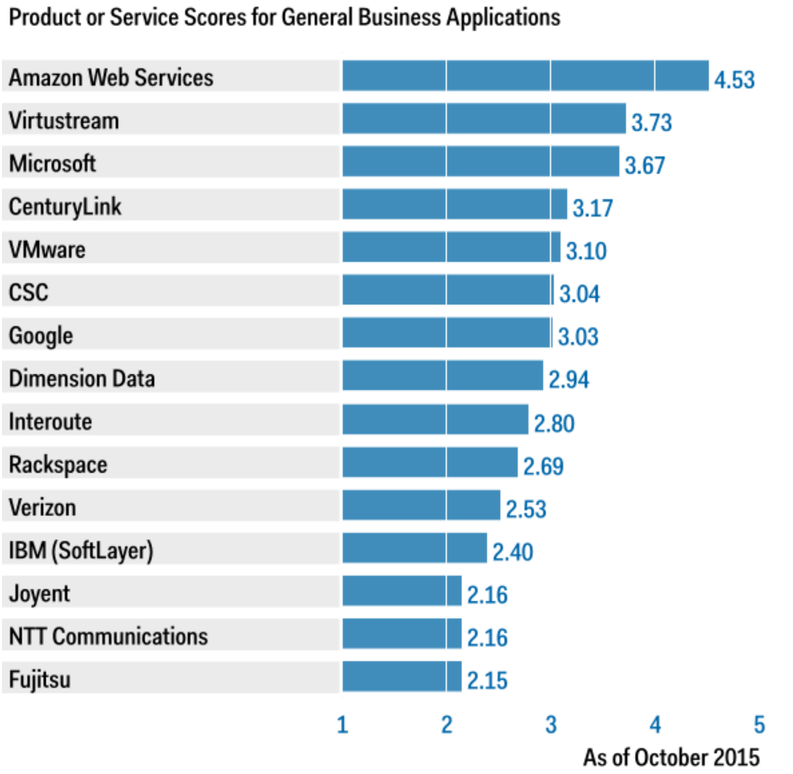
Source: Gartner (October 2015)

Figure 3. Vendors' Product Scores for Cloud-Native Applications Use Case



Source: Gartner (October 2015)

Figure 4. Vendors' Product Scores for General Business Applications Use Case



Source: Gartner (October 2015)

Vendors

All the providers evaluated in this Critical Capabilities research serve enterprise and midmarket customers, and they generally offer a high-quality service. Note that when we say "all providers," we specifically mean "all the evaluated providers included in this Critical Capabilities report," not all cloud IaaS providers in general. Keep the following in mind when reading the vendor profiles:

Most of the providers have resilient infrastructure, achieved through redundant infrastructure in conjunction with VM clustering, or the ability to rapidly detect VM failure and immediately restart it on different hardware. They are thus able to offer very high SLAs for infrastructure availability — sometimes as high as 99.999% (sometimes expressed as a 100% SLA with a 10-minute exclusion). Offerings without VM clustering or fast VM restart — which provide higher levels of infrastructure availability than can be expected from a single physical server — are noted as lacking autorestart.

Most of the providers have maintenance windows that result in downtime of the control plane (including the GUI and API), and they may require infrastructure downtime. Some offerings utilize a technology that allows VM-preserving host maintenance, but they may still have downtime for other types of maintenance. Some utilize live migration of VMs, largely eliminating the need for downtime to perform host or data center maintenance, but this does not eliminate maintenance windows in general.

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 the 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 providers offer either a shared resource pool (SRP) pricing model or are flexible about how they price the service. In the SRP model, customers contract for a certain amount of capacity (in terms of CPU and RAM) but can allocate that capacity to VMs in an arbitrary way, including being able to oversubscribe that capacity voluntarily; additional capacity can usually be purchased on demand by the hour.

Throughout the offering descriptions, VM sizes are stated as "vCPUxRAM"; for instance, "8x32" refers to eight virtual CPUs (vCPUs) and 32 gigabytes (GB) of RAM. In general, a vCPU maps to a physical core on a CPU, but not always. Implementations vary between providers — and may actually vary within a particular provider's infrastructure — since many providers have multiple generations of hardware in their cloud. CPU performance varies widely, so if it is very important to you, you should carry out your own benchmarks. Maximum VM sizes are provided as a guideline for understanding the scope of what a provider offers.

Some of the providers allow customers to choose arbitrary-size VMs — any combination of vCPUs, RAM and VM storage, subject to some limits. Providers that do not allow this are explicitly noted as offering fixed-size VMs. Some providers with fixed-size VMs have a very limited range of VM sizes, while others have a wide variety of sizes and suit a broad range of use cases. Some providers that offer arbitrary-size VMs may enforce a maximum ratio between vCPUs and RAM.

Most of the providers can resize an existing VM without needing to reprovision it; those that cannot are explicitly noted as offering nonresizable VMs. Some of the providers can resize an existing VM without needing to reboot it (if the customer's OS can also support it).

Most of the providers can provision a basic Linux VM within five minutes (although this will increase with large OS images, and Windows VMs typically take somewhat longer). Those that cannot are noted as having slow provisioning. Most providers can do simultaneous provisioning of multiple VMs; for example, provisioning 20 VMs will finish about as quickly as one VM. Those that cannot are noted as such, and the degradation can be significant (some providers take over an hour to provision 20 VMs). Nonsimultaneous provisioning has a major negative impact in disaster recovery, instant high-scalability and batch-computing scenarios.

Some of the providers are able to offer an option for single-tenant VMs within a public cloud IaaS offering, on a fully dynamic basis, where a customer can choose to place a VM on a host that is temporarily physically dedicated to just that customer, without the customer needing to buy a VM that is so large that it consumes the whole physical host. These VMs are typically more expensive than VMs on shared hosts. Providers that have this option are noted as such.

Some of the providers are able to offer "bare metal" physical servers on a dynamic basis. Due to the longer provisioning times involved for physical equipment (two hours is common), the minimum billing increment for such servers is usually daily, rather than hourly. 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.

Typically, the storage associated with an individual VM is persistent. However, some providers have ephemeral storage, where the storage exists only during the life of the VM; if the VM goes away unexpectedly (for instance, due to hardware failure), all data in that storage is lost. Ephemeral storage is always noted explicitly.

All the providers offer VM-independent block storage, unless otherwise noted. A few providers allow storage volumes to be mounted on multiple VMs simultaneously, although customers must correctly architect their solutions to ensure data integrity (just as they would with a traditional storage array).

Storage performance varies considerably between providers. Most providers can offer solid-state drives (SSDs). Providers that cannot do so, or that offer only SSD-accelerated storage (where SSDs are used to cache data for higher performance), are noted as such.

All the providers offer object-based cloud storage, unless otherwise noted. In many cases, this service is integrated with a content delivery network (CDN).

All the providers offer customers a self-service ability to create complex hierarchical network topologies with multiple network segments, and to have multiple IP addresses per VM, including a public and a private IP, unless otherwise noted. Some providers use a software-defined network (SDN), which typically allows greater API-based control over the network.

All the providers have a private WAN that connects their data centers, unless otherwise noted. They offer an option for private network connectivity (usually in the form of Multiprotocol Label Switching [MPLS] or Ethernet purchased from the customer's choice of carrier), between their cloud environment and the customer's premises, unless otherwise noted. Providers for which we state "third-party connectivity is via partner exchanges" are ones where private connectivity is obtained via cross-connect in the data centers of select partners, such as Equinix and Interxion; this also meets the needs of customers who require colocated equipment. Some carriers may also have special products for connecting to specific cloud providers — for example, AT&T NetBond and Verizon Secure Cloud Interconnect.

Most of the providers support the use of Internet-based IPsec virtual private networks (VPNs). All the providers allow customers to have VMs with only private IP addresses (no public Internet connectivity), and also allow customers to use their own IP address ranges, unless otherwise noted. Some providers may enforce secure access to management consoles, restricting access to VPNs or private connectivity. All providers have a DNS service, unless otherwise noted.

All the providers claim to have high security standards. The extent of the security controls provided to customers varies significantly, though. All providers offer multifactor authentication (MFA) for the portal, unless otherwise noted. Most providers offer additional security services. All the providers evaluated can offer solutions that will meet common regulatory compliance needs, unless otherwise noted. Unless otherwise noted, all the providers have ISO 27001 audits for their public cloud IaaS offering (see Note 1). Many will have SOC 1, 2 and 3 (see Note 2). A few can meet FedRAMP requirements (see Note 3). Some can support PCI compliance with stored cardholder data. Some can support Health Insurance Portability and Accountability Act (HIPAA) compliance and will sign a business associate agreement (BAA). Audits should not be taken as indications of security.

All the providers conduct background checks on personnel, prohibit their personnel from logging into customer compute instances unless the provider is doing managed services on behalf of the customer, and log the provider's administrative access to systems, unless otherwise noted.

Most of the providers offer a stateful firewall (intrusion detection system/intrusion prevention system [IDS/IPS]) as part of their offering, although a few offer only access control lists (ACLs), and a few offer no self-service network security at all; this will always be explicitly noted. All providers provide distributed denial of service (DDoS) attack mitigation, unless otherwise noted.

All the providers offer self-service monitoring as an option, unless otherwise noted. A few offer trigger-based autoscaling, which allows provisioning-related actions to be taken based on a monitored event. Some providers offer schedule-based autoscaling, which allows provisioning-related actions to be executed at a particular time. Note that many providers have autoscaling that stops and starts compute instances in a preprovisioned pool ("static autoscaling"), rather than provisioning or deprovisioning them from scratch ("dynamic autoscaling"). If stopped instances continue to incur charges (storage charges may apply even if there is no compute charge), static autoscaling may offer lessened cost-savings. Dynamic autoscaling is more flexible, but it may be slower than static autoscaling when the provider does not have very fast provisioning times.

All the providers offer self-service, front-end load balancing, unless otherwise noted. All also offer back-end load balancing (used to distribute loads across the middle and back-end tiers of an application), unless otherwise noted.

All the providers offer a portal and self-service mechanism that is designed for multiple users and that offers hierarchical administration and role-based access control (RBAC). However, the degree of RBAC granularity varies greatly. From most to least control, RBAC can be per element, tag, group or account. Unless stated otherwise, a provider's RBAC applies across the account. Providers typically predefine some roles; the ability to have customer-defined roles offers more control, and is noted where available. We strongly recommend that customers that need these features, but that want to use a provider that does not have strong support for them, evaluate a third-party management tool, such as Dell Cloud Manager (formerly Enstratius), RightScale or ServiceMesh (owned by CSC).

All the providers log events such as resource provisioning and deprovisioning, VM start and stop, and account changes, and allow customers self-service access to those logs for at least 60 days, unless otherwise noted.

All the providers, unless otherwise noted, offer the ability to place metadata tags on provisioned resources, and to run reports based on them, which is useful for internal showback or chargeback. Some providers also offer cost control measures, such as quotas (limits on what a user can provision) and leases (time-limited provisioning of resources).

The capabilities listed above are by no means comprehensive, and they are considered industry norms for enterprise-class offerings; few such capabilities are differentiating. Providers need basic capabilities to be implemented in a robust fashion in order to score well on this critical capability evaluation, but they also need capabilities beyond the basics listed above.

Keep in mind that a provider that seems like a good fit for a particular category of use case might not be an ideal fit for a specific need, as individual technical and business requirements and priorities vary. Conversely, a provider with a mediocre score on the relevant use case may nevertheless have the best fit to your particular requirements, especially if you have unusual constraints.

Each of the public cloud IaaS offerings rated in this Critical Capabilities report is briefly summarized below.

For each service provider, we provide an overview that discusses the offering's fit to common use cases, as well as a set of service traits, which summarizes the offering's compute, storage and network capabilities compared with the list of expected characteristics above, along with notable management-related services (such as monitoring, autoscaling, templates or an integrated service catalog) and other services (specifically, cloud software infrastructure services, such as database as a service).

Offerings that use VMware's vSphere hypervisor are described as vCloud-based or VMware-virtualized. A vCloud-based offering uses vCloud Director (vCD) software and offer access to the vCloud API; these service providers may offer their own portal, the vCD portal or both.

Amazon Web Services

Offering evaluated:

Amazon Web Services (AWS) essentially created the cloud IaaS market with the 2006 introduction of its Elastic Compute Cloud (EC2), and it still offers the richest suite of public cloud IaaS capabilities, along with deep and broad PaaS-layer capabilities.

AWS is suitable for nearly all use cases that run well in a virtualized environment. Applications should not need to be modified to run on AWS, although customers may benefit from optimizing applications for the platform. Customers also need to pay attention to best practices for resiliency, performance and security. AWS is a complex platform, due to its extensive array of capabilities and options; it takes deep expertise to take optimal advantage of AWS.

AWS is an especially strong choice for digital business and other new applications, including customer-facing applications, big data and analytics, and back ends for mobile applications and the Internet of Things. Its extensive suite of services is useful for improving developer productivity and simplifying operations, and customers typically use a blend of AWS's IaaS and PaaS capabilities. It is best-suited to a DevOps style of operations.

AWS is also commonly used for legacy applications in a "lift and shift" approach, as well as transformation-oriented full data center migrations, due to its solid feature set and ability to meet most requirements for security and regulatory compliance. Its very fast provisioning times may make it attractive as a disaster recovery platform when using an identical hypervisor is not a priority. AWS's extensive software marketplace and strong ecosystem of managed and professional services partners facilitate adoption of its platform.

AWS has many customers that have a large number of users, such as developers, scientists, engineers or researchers; it has the most sophisticated capabilities for management and governance of many accounts, users and infrastructure components. Batch computing users may value the AWS Spot Market, which uses a reverse-auction style of bidding for compute instances.

AWS appeals most strongly to customers who value thought leadership, cutting-edge capabilities, or a "safe" provider that has a well-proven service and is likely to continue to be a long-term market leader.

See "In-Depth Assessment of Amazon Web Services" for a detailed technical evaluation.

Notable Service Traits:

Compute: Xen-virtualized, fixed-size and nonresizable, but with a very broad range of VM sizes. Maximum VM size of 36x244. Single-tenant VM option (Dedicated Instances). HPC option that includes GPUs and high-performance network interconnects. Option for burstable-CPU smaller VMs. Autorestart is not available in all regions. Docker container service integrated into EC2.

Storage: Ephemeral local storage, VM-independent block storage (Elastic Block Store [EBS]), object storage with integrated CDN (S3 with CloudFront), and archive storage (Glacier). Optional encryption. Optional Provisioned IOPS for EBS provides quality-of-service guarantees for storage performance.

Network: Highly configurable and sophisticated SDN (Amazon Virtual Private Cloud [VPC]). Global load-balancing service. Third-party connectivity via partner exchanges (Direct Connect).

Security: Can meet almost all common audits and common compliance requirements, including SOC 1, SOC 2, SOC 3, ISO 27001, FedRAMP, PCI, HIPAA and GxP (pharmaceutical industry). Encryption available for most data stores. Very granular RBAC (Identity and Access Management), including RBAC across multiple accounts. MFA also includes API. Active Directory integration. Key management service, including hardware security modules.

Management: Control plane is continuously available. No maintenance windows. Broad range of native capabilities, including monitoring (CloudWatch), dynamic autoscaling, configuration management (Config, CloudFormation, OpsWorks), service catalog, developer services (CodeDeploy and more) and billing management. Extensive marketplace and ecosystem.

Other services: Database (Relational Database Service [RDS], DynamoDB), caching (ElastiCache), data warehouse (Redshift), Hadoop (Elastic MapReduce), data ingest and event processing (Data Pipeline, Kinesis, Lambda), and many more.

CenturyLink

Offering evaluated: CenturyLink Cloud (CLC)

CenturyLink has a range of public and private cloud IaaS offerings, built on different platforms. Its primary public cloud IaaS offering is CenturyLink Cloud, which was obtained via the acquisition of Tier 3 in November 2013.

CLC attempts to tread the middle ground between the needs of IT operators and developers by providing a service that appeals to traditional IT operations teams, but that also offers ease of use and API-controllable capabilities to developers.

Although CLC can be used in an entirely self-service fashion, CenturyLink also provides many add-on managed services that are heavily automated but still have some human-guided elements. Most customers who choose CLC are likely to purchase some managed services; customers with strong security or regulatory compliance needs are likely to require managed security services.

CLC's availability-oriented features are of particular note as they include automatic replication of VMs into a second data center, and storage that is integrated with rolling backups and disaster recovery options. The service's scriptable template system, Cloud Blueprints, is capable of provisioning complex, multi-data-center infrastructure configurations, and it is integrated with a marketplace. Its capacity pool is, however, relatively small, and provisioning throughput is relatively low, resulting in lengthy provisioning times when many VMs need to be provisioned at the same time. This limits the usefulness of the service for applications that require rapid scalability, or for on-premises disaster recovery scenarios that demand short recovery time objectives.

CLC is a capable platform for general business applications that run well in a virtualized environment (or that are suited to CLC's bare-metal server configurations), and thus can be used to "lift and shift" applications without change. It has solid features for user governance, including the ability to provide a restricted service catalog, which may make it appealing as a lab environment for developers. The CLC portal also allows a customer's administrator to set the price that his or her subaccounts see for CLC services, thus surfacing a price for internal chargeback.

CLC will appeal to IT operations teams that are prioritizing their own operational requirements or want to use managed services, but need to appease developers who want self-service infrastructure.

Notable Service Traits:

Compute: VMware-virtualized with a maximum VM size of 16x128, and nonsimultaneous provisioning. Bare-metal servers in three configurations, up to a maximum size of 20x128.

Storage: Local storage, VM-independent block storage and object storage. Media disposal does not meet NIST standard.

Network: Does not support use of customer's private IP addresses. No back-end load balancing.

Security: SOC 2 and SOC 3 audits. Will support PCI and HIPAA with BAA. No MFA. Provider personnel can log into customer compute instances. Group-based RBAC.

Management: Monitoring. Horizontal and vertical static autoscaling. Patching. Templates (Blueprints). Service catalog.

Other services: Database (with broader capabilities available via Orchestrate, an acquisition). Cloud Foundry-based PaaS (AppFog). Disaster recovery (SafeHaven, for recovering on-premises VMs into CLC).

CSC

Offering evaluated: BizCloud Virtual Private Edition (VPE)

CSC entered the cloud IaaS market in 2009. It has a common platform for its public and private cloud IaaS offerings, which all come at a single price point, but with different minimum capacity and term commitment requirements. CSC's BizCloud VPE offering has single-tenant compute but multitenant storage and networking. However, the capabilities of CSC's CloudCompute offering (which uses multitenant compute) and BizCloud (which is fully single-tenant) are nearly identical. New capabilities are introduced into BizCloud before being rolled out into CloudCompute.

CSC's cloud IaaS offerings are vCloud Datacenter Services, and thus are VMware-certified as highly available, secure and capable of workload portability. This should ease the movement between CSC's cloud and other VMware-virtualized infrastructure. Most vCD capabilities are available to the customer, but CSC also offers its own easier-to-use portal, Agility Store, which is based on the Agility Platform from CSC's 2013 acquisition of ServiceMesh. Customers may optionally use the full Agility Platform, which provides additional governance and automation capabilities. CSC has also integrated a significant number of other ITOM tools into the service.

CSC offers a capable platform for general business applications that run well in a virtualized environment, and thus can be used to "lift and shift" applications without change. However, CSC is focused on enabling data center transformation. Its offering will appeal to traditional IT operations organizations that value enterprise-class availability, security and governance or VMware-based hybrid interoperability, or that would like to gradually transform toward a DevOps philosophy.

Notable Service Traits:

Compute: VMware-virtualized. Maximum VM size of 16x128. Optional automated disaster recovery capabilities.

Storage: Local storage and VM-independent block storage. Optional encryption.

Network: No DNS service.

Security: ISO 27001 audit. Portal is accessible only via VPN, with certificate-based authentication. No MFA. No DDoS mitigation. Group-based RBAC.

Management: Monitoring. Service catalog, templates and governance via the Agility Platform.

Other services: Database (CloudDB).

Dimension Data

Offering evaluated: Dimension Data Public Compute-as-a-Service (CaaS)

Dimension Data has a common platform for its public and private cloud IaaS offerings, obtained through the 2011 acquisition of OpSource. Although it is VMware-virtualized, the portal and API are Dimension Data's own.

Dimension Data's Public CaaS offers a solid set of core compute, storage and networking capabilities. However, it has a limited feature set for management and governance, and it lacks cloud software infrastructure services (such as database as a service).

While this offering is capable of running cloud-native applications, the Dimension Data API is proprietary and lacks an ecosystem of third-party tools, thus limiting its appeal to developers. It also lacks rapid scalability. However, it may be suitable for SaaS vendors that need only basic infrastructure and intend to use Dimension Data's other SaaS-enablement capabilities.

This offering may also appeal to those who want to "lift and shift" general business applications and would prefer to control their infrastructure via a portal, or intend to use Dimension Data's managed services.

Dimension Data is focused on providing an environment that can reliably run production applications. The offering will appeal to developers who want an enterprise-class, VMware-based infrastructure and do not have a heavy DevOps orientation.

Notable Service Traits:

Compute: VMware-virtualized. Maximum VM size of 32x256. Nonsimultaneous provisioning.

Storage: Local storage, VM-independent block storage and object storage. SSD-accelerated storage is optional; no all-SSD storage. Expandable volumes.

Network: Cisco-hardware-based. LAN encryption between compute and storage devices.

Security: SOC 1 and ISO 27001 audits. Can support PCI and HIPAA with BAA. Provider personnel can log into customer compute instances.

Management: Monitoring. Horizontal and vertical static autoscaling. Active Directory integration. Compute and storage leases.

Other services: None in the scope of this evaluation.

Fujitsu

Offering evaluated: Fujitsu Cloud IaaS Trusted Public S5

Fujitsu launched a common global platform for its public and private cloud IaaS offerings in 2010. S5 offers Fujitsu's own portal and API. It uses Fujitsu's Resource Orchestrator Cloud Edition to provide its visual designer for infrastructure, as well as a significant depth of portal features, including solid features for user management.

Although Fujitsu targets the enterprise market, S5 is Xen-virtualized. Furthermore, users are constrained by relatively low-performance compute and storage, limiting S5's suitability for use cases that require more compute power, larger amounts of RAM or faster I/O.

S5 may be suitable for IT operations teams that need to provide a well-governed infrastructure lab environment for developers. S5 may also be suitable for small general business applications that run well in a virtualized environment and do not have significant security or regulatory compliance requirements.

S5 offers an acceptable set of baseline capabilities that may potentially meet the needs of both developers and IT operations organizations, especially those in Asia/Pacific.

Notable Service Traits:

Compute: Xen-virtualized and fixed-size, with a very limited range of VM sizes. Maximum VM size of 8x60. Slow, nonsimultaneous provisioning.

Storage: Block storage, without an SSD option. Object storage (Japan only). Snapshots cannot be used as images.

Network: LAN encryption. Does not fully support self-service complex hierarchical network topologies. Does not support static IPs.

Security: ISO 27001 audit. Granular RBAC. No DDoS mitigation. No MFA. Very limited logging of portal and API actions.

Management: Monitoring. Service catalog. Quotas and leases. No enterprise directory integration.

Other services: None in the scope of this evaluation.

Google

Offering evaluated: Google Cloud Platform (GCP)

Google Cloud Platform combines IaaS and PaaS capabilities within an integrated solution portfolio. Google's VM offering, Google Compute Engine (GCE), became generally available in December 2013.

GCP places a strong emphasis on self-service, API-accessible capabilities. It has a relatively rich set of capabilities and should appeal strongly to developers building new cloud-native applications who want to do things the "Google way" and intend to adopt a broad range of Google APIs.

GCE is also a very attractive platform for batch computing; it has exceptionally fast provisioning times and a very large pool of available capacity, along with per-minute metering and low-cost "pre-emptible VMs" (which can be shut down by the platform at any time), making it especially well-suited to short-lived large-scale batch jobs.

Although GCE may be technically suited to run general business applications, GCP's ability to integrate with on-premises infrastructure is weak. Google is focusing its hybrid efforts on containers and its open-source Kubernetes orchestration software. Although it is not necessary to rearchitect applications to run on GCE, Google has little in the way of features that facilitate the "lift and shift" of traditional applications.

GCP's strengths lie in developer enablement and new applications. It will appeal to organizations that like cutting-edge capabilities and are willing to use an emerging offering.

See "In-Depth Assessment of Google Cloud Platform" for a detailed technical evaluation.

Notable Service Traits:

Compute: KVM-virtualized, fixed-size and nonresizable, with a broad range of VM sizes. Maximum VM size of 32x208.

Storage: Ephemeral and persistent local storage. VM-independent block storage. Object storage. All data is encrypted at rest and in motion; customers can provide their own keys (in beta).

Network: High-performance, very configurable SDN. LAN and WAN encryption. No back-end load balancing. Global load-balancing service.

Security: SOC 1, SOC 2, SOC 3 and ISO 27001 audits. Will support PCI and HIPAA with BAA.

Management: Control plane is continuously available. No maintenance windows. Monitoring. Templates (Cloud Deployment Manager). No enterprise directory integration.

Other services: Kubernetes-based Docker container service (Container Engine), database (Cloud SQL, Cloud Bigtable, Cloud Datastore), data ingest (Cloud Dataflow, Cloud Pub/Sub), data analytics (BigQuery).

IBM (SoftLayer)

SoftLayer was acquired by IBM in July 2013, and its services replaced IBM's SmartCloud Enterprise offering. SoftLayer offers dedicated hosting, as well as cloud IaaS. In this assessment, we evaluate only those capabilities that are available as a cloud service — those that are standardized, fully automated and metered by the hour (or less). SoftLayer's noncloud capabilities are typically provided via hardware rented by the month, rather than as abstracted services; we refer to these solutions as "hosted appliances." All accounts have a quota of compute instances, and provisioning additional capacity requires sales order approval.

SoftLayer's compute options include bare-metal servers, as well as VMs, and VMs can be on single-tenant or multitenant hosts. Although some capabilities require VMs, SoftLayer tries to minimize the differences between VMs and bare-metal servers; its strength lies in these bare-metal capabilities. However, SoftLayer's services were originally designed for the small-business market, and its greatest weaknesses are in the capabilities that are desired by larger organizations.

SoftLayer is best-suited to batch computing use cases that require bare metal. It may also be suitable for cloud-native applications that require API-provisioned bare metal, but do not require API control or on-demand capacity for other solution elements (such as a load balancer).

Organizations could potentially consider SoftLayer for "lift and shift" migrations, simply using SoftLayer bare-metal servers as a substitute for on-premises servers — renting rather than buying servers. SoftLayer is less-suited to other general business application use cases. SoftLayer lacks all key compute resilience features; furthermore, users are notified of maintenances, but customers need to carefully configure notifications, as SoftLayer can generate a large number of email alerts. VM import is limited to a narrow range of OS versions. SoftLayer's weak user management capabilities, with no enterprise directory integration or granular RBAC, makes it unsuitable for most large-scale application development use cases.

SoftLayer will appeal to organizations that like portal and API control over scalable infrastructure, but need bare-metal servers in order to meet requirements for performance, regulatory compliance or software licensing. Furthermore, organizations that are outsourcing their infrastructure to IBM, and are replatforming onto SoftLayer dedicated hosting, may find that SoftLayer's cloud services are a useful complement.

See "In-Depth Assessment of SoftLayer, an IBM Company" for a detailed technical evaluation.

Notable Service Traits:

Compute: Citrix Xen-virtualized, fixed-sized, multitenant or single-tenant VMs, in many possible sizes, up to a maximum VM size of 16x64. Bare-metal servers up to a maximum size of 32x512. No VM autorestart. No VM-preserving host or data center maintenance.

Storage: VM-specific and VM-independent block storage; no all-SSD option. Object storage (without multi-data-center replication). Cannot snapshot VM-specific storage. Media disposal does not meet NIST standard.

Network: Does not fully support complex hierarchical network topologies or customer-provided private IP addresses without use of a hosted appliance. No load balancing; customer can use a hosted appliance.

Security: SOC 1, SOC 2, SOC 3 and ISO 27001 audits. Will sign HIPAA BAA. Provider personnel can log into customer compute instances.

Management: Monitoring. Dynamic autoscaling. No enterprise directory integration.

Other services: None within the scope of this evaluation. Customers with a need for PaaS should consider the use of IBM's Bluemix, which provides Cloud Foundry-based PaaS and access to other PaaS-layer services, along with a Docker container service and (in beta) OpenStack-based VMs. Bluemix is physically located in SoftLayer data centers but is a wholly distinct set of services.

Interoute

Offering evaluated: Interoute Virtual Data Centre (VDC)

Interoute entered the cloud IaaS market with the 2012 launch of its VDC offering. VDC is offered in various tenancy models, including multitenant public cloud. Although many other network service providers also have cloud IaaS offerings connected to their WAN services, Interoute has done a unique degree of integration between its network services and its cloud infrastructure. Customers can integrate and API-control their VDC LAN in conjunction with their Interoute WAN services.

Interoute VDC is best-suited to deploying production applications that require integration with the WAN, such as multisite distributed applications and applications that replicate data across multiple data centers. These can be either cloud-native applications or legacy applications.

However, because Interoute lacks critical features for user governance, it is unlikely to be suitable for organizations with large numbers of users of the service (such as developers or scientists), and therefore it is less likely to be suitable for application development or batch computing use cases.

Interoute VDC will appeal to organizations for which network integration is a high priority, especially those seeking DevOps-oriented control over an SDN WAN, or that have Pan-European cloud IaaS data center needs.

Notable Service Traits:

Compute: CloudStack-based and multihypervisor (VMware, Citrix Xen or KVM) VMs. Maximum VM size of 40x500.

Storage: Ephemeral and persistent local storage. VM-independent block storage. Media disposal is compliant with EU Data Protection Directive (disposal that meets NIST standard available on request).

Network: Customers can integrate LAN topologies with Interoute's WAN services and configure multisite networks via its portal as well as an API. No DNS service.

Security: SOC 1 (ISAE 3402) and ISO 27001 audits. No MFA. No granular RBAC.

Management: Monitoring. Puppet integration.

Other services: None within the scope of this evaluation.

Joyent

Joyent entered the cloud IaaS market in 2007. It has recently rebranded its public cloud IaaS offering as the Joyent Triton Elastic Container Service. Customers have a choice of "infrastructure containers" (native OS-based containers), Docker containers or VMs. Joyent's architecture is container-native; compute resources run in Triton Zones (Joyent's SmartOS virtualization technology, similar to Solaris Zones), which offer additional security isolation to both containers and VMs.

Beyond containers, Joyent also has some unique capabilities, such as analytics within its portal that exposes DTrace-based instrumentation of application and infrastructure performance, and its Manta platform, which couples object-based storage with the unique ability to run batch jobs on compute that is local to that storage. Joyent has a modest but solidly supported library of prebuilt images for common infrastructure software, including commercially licensed software.

Joyent should be looked at primarily as a distinctive container-oriented offering in the market, rather than as a general-purpose cloud IaaS offering (where the breadth of its feature set lags the market as a whole). Joyent will appeal to organizations that are developing container-native applications, or are looking for an easy-to-use platform with which to experiment with container technologies. Joyent also sells its software and offers a private cloud service (including on-premises), which may make Joyent attractive for some hybrid cloud use cases that require identical capabilities between cloud deployments.

Notable Service Traits:

Compute: Compute hosts use Joyent's SmartOS (based on illumos Solaris) as the host OS. Virtualization is via Triton Zones (Windows guests also use KVM); VMs are fixed-size and nonresizable. Native resizable containers as a service. Docker containers as a service. Maximum VM size of 32x256. No VM image import or export.

Storage: Local storage with SSD is limited to three VM sizes and five native container sizes. No VM-independent block storage. Object storage (Manta). Snapshots cannot be used as images. No data import from physical media. Media disposal does not meet NIST standard.

Network: No load balancing (customers can deploy Brocade Virtual Traffic Manager virtual appliance, referred to in the documentation by its previous brands of Riverbed Stingray and SteelApp). No inter-data-center private WAN. No enterprise network integration. No DNS service.

Security: SOC 1 audit. Will support PCI and HIPAA with BAA. Personnel can log into customer compute instances. Very granular RBAC.

Management: Monitoring.

Other services: None within the scope of this evaluation.

Microsoft

Offering evaluated: Microsoft Azure

Microsoft Azure combines a rich suite of IaaS and PaaS capabilities within an integrated solution portfolio. Azure was initially a PaaS launched in 2010, and IaaS VMs (Azure Virtual Machines) became generally available in April 2013.

Although Microsoft Azure is not as mature or feature-rich as AWS, it is still suitable for a broad range of use cases that run well under virtualization, and it has its own distinct set of capabilities. Customers are likely to consider Microsoft Azure for hosting Microsoft applications such as SharePoint, as well as use cases where the application is Windows-based, is written in .NET, is developed by a team using Microsoft developer tools such as Visual Studio, or is dependent on Microsoft middleware. However, Microsoft is increasingly targeting applications that run on Linux (and claims 20% of Azure VMs run Linux), although the most common enterprise Linux (Red Hat) is not currently available on Azure.

Azure is a capable environment for digital business workloads and other cloud-native applications, including mobile back ends and Internet of Things applications. Customers are highly likely to mix IaaS VMs with PaaS-level compute capabilities when using Azure for new applications. The Azure Batch service, Azure's analytics-related services and VM configurations designed for HPC have made Azure an attractive environment for batch computing. Some midmarket customers have also begun to "lift and shift" existing applications to Azure.

Although complex network topologies are supported, customers report difficulties with the hierarchical networking necessary to deploy security-related virtual appliances, such as enterprise firewalls, which may make Azure unsuitable for some secure workloads. Also, Azure is still in the process of building governance capabilities, making it less suitable for organizations with many cloud users or large deployments, although some of these challenges can be addressed using external tools. Furthermore, because Azure does not have multiple data centers in the same region and previous Azure outages have impacted multiple regions simultaneously, it is more difficult to deploy highly resilient architectures on Azure, and customers may need to consider non-Azure disaster recovery for Azure-based applications.

Microsoft Azure will appeal to organizations that have existing investments in Microsoft technologies and that intend to do one or more of the following: Use Azure for cloud-native applications that are built on .NET, use Microsoft middleware or use Azure PaaS capabilities; host Windows applications (with attention paid to Azure's ability to meet availability, performance and security requirements); migrate a Microsoft-centric data center to the cloud over a multiyear period; augment Microsoft SaaS applications; or build a hybrid cloud environment with Azure Pack (or the forthcoming Azure Stack).

See "In-Depth Assessment of Microsoft Azure" for a detailed technical evaluation.

Notable Service Traits:

Compute: Hyper-V-virtualized, fixed-size and nonresizable, but with a broad range of VM sizes. Maximum VM size of 32x448. HPC options that include high-performance network interconnects. Alternatively, use the PaaS VM-based compute service (Cloud Services Web and Worker roles) or App Service.

Storage: Ephemeral local storage, VM-independent block storage and object storage (Blobs) with integrated CDN. Higher-performance and SSD-based storage (Premium Storage) is available only as local storage for specific VM types.

Network: SDN (Virtual Networks).Third-party connectivity via partner exchanges (ExpressRoute). Global load-balancing service.

Security: Can meet most common audits and common compliance requirements, including SOC 1, SOC 2, ISO 27001, FedRAMP, PCI, HIPAA and GxP (pharmaceutical industry). No stateful firewall. RBAC uses predefined roles. Active Directory service and integration. Key management service.

Management: Control plane is continuously available. No maintenance windows. Monitoring. Scheduling service. Run book service (Azure Automation). Significant marketplace.

Other services: Database (SQL Database, DocumentDB), caching (Redis Cache), data warehouse, Hadoop (HDInsight), data ingest and event processing (Data Factory, Event Hubs, Stream Analytics), and many more.

NTT Communications

Offering evaluated: NTT Cloud ^{n}

NTT Communications has multiple cloud IaaS offerings. The Cloud ^{n} offering was originally launched in 2012 by NTT's subsidiary Verio. Cloud ^{n} encompasses both IaaS elements and a Cloud Foundry-based PaaS.

Cloud ^{n} contains a basic set of core compute, storage and networking capabilities, along with other vital complementing capabilities, such as autoscaling and a relational database service.

Cloud ^{n} can be used for cloud-native applications where API-controllable infrastructure is desirable, or when the customer wants to mix a Cloud Foundry PaaS front end with VM back ends. However, Cloud ^{n} 's weak security and enterprise integration capabilities will limit its appeal to customers who want to run general business applications in the cloud. Its limited range of VM sizes constrains the technically feasible use cases, especially those related to batch computing.

Cloud ^{n} may be appealing to developers who want a CloudStack-based (and thus somewhat AWS API-compatible) offering, particularly in Asia/Pacific. Customers who want cloud IaaS from NTT but want a "rented virtualization" approach that appeals more to IT operations organizations should consider NTT's Enterprise Cloud offering instead.

Notable Service Traits:

Compute: CloudStack-based, KVM-virtualized, fixed-size and nonresizable VMs. VMs are divided into two types: Flat (no control over network topology) and VPC Type OpenNW (allows configurable networking). The maximum size of Flat-type VMs is 16x64; the maximum size of VPC-type VMs is 8x16. Only Flat-type VMs are available in the U.S. region.

Storage: Local storage and network-attached file storage, without SSDs. S3-compatible object storage. No import of data from physical media. Media disposal does not meet NIST standard.

Network: Only VPC-type VMs have a customer-controlled network topology. No back-end load balancing. CDN service.

Security: ISO 27001 audit. No stateful firewall. No DDoS mitigation. No granular RBAC.

Management: Monitoring. Dynamic autoscaling. Templates (Cloud ^{n} Provisioning). No enterprise directory integration.

Other services: Database (RDB).

Rackspace

Offering evaluated: Rackspace Public Cloud

Rackspace began offering cloud IaaS in 2008, when it acquired Slicehost. However, in August 2012, it launched an OpenStack-based offering into general availability. This assessment covers only the current-generation OpenStack-based offering and does not include any of Rackspace's managed hosting capabilities. Rackspace mandates bundling a service level with all compute resources, but customers who want to self-manage can choose the "managed infrastructure" level for technical support without managed services.

Rackspace has a solid set of core compute, storage and networking capabilities, presented in an easy-to-use portal (Rackspace's own, not OpenStack Horizon). It also has some useful PaaS-layer services; some were obtained through acquisitions and available to any customer, not just those using Rackspace Public Cloud.

Rackspace Public Cloud is most likely to be used for typical OpenStack use cases — scalable infrastructure for cloud-native applications and batch computing. However, Rackspace does not have the governance features needed to manage large numbers of users, nor the enterprise integration features desired for "lift and shift" of existing business applications.

Rackspace Public Cloud will appeal to organizations that are looking for an OpenStack-based public cloud offering, that value ease of use for individual developers, or that are Rackspace managed hosting customers and need some complementary cloud IaaS capabilities.

Notable Service Traits:

Compute: OpenStack-based, Citrix Xen-virtualized, fixed-sized, nonresizable VMs. No VM autorestart. Maximum VM size of 32x240. Bare metal servers (OnMetal Cloud Servers) in three configurations designed for large workloads. No import of Windows images.

Storage: Local storage; most VM types use SSDs. VM-independent block storage with SSD option. Object storage (Cloud Files) with integrated CDN. Media disposal does not meet NIST standard.

Network: SDN (Cloud Networks). Private connectivity requires the RackConnect service and a dedicated appliance.

Security: SOC 1, SOC 2 and ISO 27001 audits. MFA also includes API. RBAC roles are per service and limited to full access, create access and read-only access. No DDoS mitigation.

Management: Monitoring. Templates (Cloud Orchestration). No enterprise directory integration. Logs are generated via the Cloud Feeds service.

Other services: Database (Cloud Databases), Hadoop (Cloud Big Data). Rackspace-owned separate PaaS services for NoSQL databases and caching (ObjectRocket and RedisToGo).

Verizon

Offering evaluated: Verizon Cloud — Virtual Private Cloud

Verizon has multiple public and private cloud IaaS offerings. While Verizon has been in the IaaS business for a long time (it acquired Terremark, which launched its Enterprise Cloud in 2008), the current Verizon Cloud is a new offering seeded via core technology and personnel obtained through the acquisition of CloudSwitch in 2011. Verizon Cloud became generally available in late 2014. Verizon Cloud uses a single portal that encompasses multiple deployment options, all of which are technically distinct environments but which are collectively managed within the Verizon Cloud portal. Each service subscription to an environment is called a "CloudSpace"; one account can have multiple CloudSpaces. In this assessment, we evaluate only a single environment type: the Virtual Private Cloud.

This offering contains a solid core of compute, storage and networking capabilities, along with a marketplace. Unfortunately, the Verizon Cloud portal's performance is slow and prone to unexpected errors. Furthermore, although it has very useful features for managing large groups of users and larger amounts of infrastructure across multiple environments, the way the portal exposes this functionality significantly decreases the utility of these capabilities.

This offering can be reasonably used for lab environment use cases, where application developers simply need on-demand provisioning of infrastructure resources. It may also be a candidate for hosting non-mission-critical business applications. Relatively small VM sizes limit its possible range of use cases.

Verizon Cloud will appeal to IT operations organizations that need to provide infrastructure resources to multiple application development teams or projects, and that want to purchase cloud IaaS from their network service provider.

Notable Service Traits:

Compute: Xen-virtualized and fixed-size VMs. Maximum VM size of 8x64. No VM autorestart. Nonsimultaneous provisioning.

Storage: Local storage and block storage with expandable volumes. Object storage.

Network: SDN. Private connectivity requires the Verizon Secure Cloud Interconnect service. No back-end load balancing. No DNS service.

Security: No common audits. Group-based, cross-CloudSpace RBAC.

Management: Monitoring. Patching service. No enterprise directory integration. The customer's VM image catalog can span CloudSpaces within an account.

Other services: None within the scope of this evaluation.

Virtustream

Virtustream entered the cloud IaaS market in 2010. It was acquired by EMC in July 2015 but remains an independent entity within the EMC federation. xStream is Virtustream's common hypervisor-neutral platform for its public and private cloud IaaS offerings.

Virtustream has its own proprietary API, as well as two portals — a complex interface for IT administrators, and an easier-to-use, simplified interface for developers and other end users. It has significant capabilities in infrastructure resiliency, security and regulatory compliance, although not all such capabilities are self-service. Its Micro VM technology enables it to offer policy-based, service-level management that allows customers to pay for resources consumed rather than resources allocated. Virtustream also divides its multitenant infrastructure into physically separate hardware pools based on application type and security requirements; customer logical environments span multiple pools.

Virtustream's capabilities are focused on bringing agility and efficiency to traditional IT, rather than developer enablement or DevOps-style management. In contrast to most other cloud IaaS providers, whose services are optimized for scale-out applications, Virtustream specializes in scale-up, mission-critical enterprise applications. Virtustream has had a strong focus on SAP applications, where it has significant specialized automation capabilities. It is also suitable for application development use cases related to SAP and similar applications.

Virtustream will appeal to IT operations organizations that want to migrate mission-critical traditional enterprise applications into a cloud IaaS environment. Although managed services are not a requirement, most organizations will need Virtustream's assistance in transitioning applications onto the xStream platform.

Notable Service Traits:

Compute: VMware or KVM-virtualized, multitenant or single-tenant VMs. Maximum VM size of 32x1,024.

Storage: Local storage. VM-independent block storage with expandable multimountable volumes. Optional encryption. No object-based storage. Snapshots cannot be used as images.

Network: No back-end load balancing. No DNS service.

Security: SOC 1, SOC 2 and SOC 3 audits. FedRAMP. Will sign HIPAA BAA. Granular RBAC. Portal requires VPN or private connectivity.

Management: Monitoring. Compliance.

Other services: None within the scope of this evaluation.

VMware

Offering evaluated: vCloud Air

VMware's vCloud Air (previously named VMware vCloud Hybrid Service) offering became generally available in September 2013. It is available in two tenancy models: Virtual Private Cloud (multitenant compute and storage) and Dedicated Cloud (single-tenant compute and multitenant storage). Virtual Private Cloud is offered in two variants — a paid-by-the-month shared resource pool, and a pay-as-you-go per-VM service. vCloud Air most closely resembles a vCloud Datacenter Service. While it uses vCD, vCloud Air has its own more user-friendly portal.

vCloud Air is currently an infrastructure-focused offering, with few capabilities for developer enablement or DevOps-style management. VMware's hybrid cloud strategy emphasizes workload portability between vCloud Air and on-premises VMware-virtualized infrastructure. Consequently, vCloud Air is best-used for application development and general business application use cases, where the highest-priority requirement is the IT operations team's desire to use the same VMware-based infrastructure constructs in the cloud and on-premises.

vCloud Air will appeal to IT organizations that desire a cloud IaaS offering that allows them to continue to use their existing investments in VMware-based IT operations skills and management tools; that need to either move existing VMware-virtualized workloads to the cloud, or be able to move cloud-developed applications back on-premises; and that are focused on traditional IT operations needs, not development enablement or IT transformation.

See "In-Depth Assessment of VMware vCloud Air" for a detailed technical evaluation.

Notable Service Traits:

Compute: VMware-virtualized. Maximum VM size of 16x244.

Storage: Local storage. VM-independent block storage with expandable volumes; this storage can be SSD-accelerated, but there is no pure-SSD option. Object storage.

Network: SDN (NSX-based). No inter-data-center private WAN. Third-party connectivity via partner exchanges. No DNS service.

Security: SOC 1, SOC 2, SOC 3 and ISO 27001 audits. Will sign HIPAA BAA. No MFA.

Management: Monitoring. Templates (as vApps). No enterprise directory integration.

Other services: Database (vCloud Air SQL, in beta). Disaster recovery.

Context

New digital business opportunities are placing tremendous pressure on IT organizations not only to deliver IT quickly and efficiently, but to create business value using information and technology. Most IT operations organizations have struggled to keep up with the demands of the business. Not only are they asked to rapidly deliver infrastructure for new initiatives, but these initiatives are often not part of the year's IT budget, and the demands are made without warning. Moreover, some of these initiatives result in unpredictable capacity demands, because the business cannot accurately predict the degree of success it will experience. Finally, some of these initiatives use application architecture patterns that are new to the organization.

As a result of these challenges, many application development organizations have turned to public cloud IaaS. IT operations organizations have started to embrace public cloud IaaS as well, viewing it as a potential replacement for most of their infrastructure. Public cloud IaaS is used by organizations of all sizes, and Gartner estimates that it now accounts for almost 20% of all virtualized workloads. Gartner's 2015 CIO Survey indicates that 10% of CIOs now have a "cloud first" approach to infrastructure, and 85% of CIOs consider cloud IaaS for their infrastructure needs (see "Implications of the 2015 CIO Survey for Infrastructure and Operations Leaders"). Few IT organizations have succeeded with private clouds (see "Internal Private Cloud Is Not for Most Mainstream Enterprises"), and most developers prefer to use public cloud IaaS, usually because it offers greater control, flexibility, ease of use and more self-service automated features. This is particularly true when using a provider that can offer a much faster innovation cycle, with potentially hundreds of new capabilities and improvements each year. Consequently, many organizations now see the choice of a public cloud IaaS provider as strategic, even if they intend to continue to run most of their IT infrastructure internally.

In the early years of the market, public cloud IaaS was used primarily for application development and testing, for batch computing and for new applications designed with cloud-native architectures. It is now commonly used for mainstream business applications as well, including mission-critical enterprise applications. Small or midsize businesses have broadly adopted public cloud IaaS for their infrastructure needs. Some midmarket companies have begun the process of migrating a significant percentage of their workloads onto public cloud IaaS. Most enterprises are placing select projects on public cloud IaaS, and they may be deploying almost all new applications onto public cloud IaaS.

Gartner recommends that organizations adopt a bimodal approach to cloud IaaS (see "Best Practices for Planning a Cloud Infrastructure-as-a-Service Strategy — Bimodal IT, Not Hybrid Infrastructure"). Public cloud IaaS is most frequently adopted for Mode 2 agile IT, but it is increasingly being used for Mode 1 reliable IT as well. Organizations are most likely to adopt public cloud IaaS because they have a requirement for business agility, but they are increasingly turning to public cloud IaaS to obtain cost savings.

Many organizations initially adopt a provider in an ad hoc fashion — often on a credit card and click-through agreement — and then negotiate an enterprise agreement later. It is common to conduct a formal RFP after the organization already has some (or even a large amount of) public cloud IaaS adoption. This is typically done to reduce the number of suppliers, gain volume discounts and improve governance. Most organizations now choose their primary public cloud IaaS providers in a strategic fashion, with a strong emphasis on the provider's long-term viability and likelihood of being a market leader in five to 10 years.

The public cloud IaaS market has entered its second phase of maturity. The technology is now sufficiently mature for most needs, but it continues to improve rapidly as providers introduce innovative new capabilities. Many providers, however, are struggling to compete, as buyer expectations and competitive pressure continue to increase. For most such providers, the gap between their capabilities and the capabilities of the market leaders is growing, not shrinking.

IT organizations that have not adopted public cloud IaaS should conduct an in-depth competitive evaluation. Also, organizations with one or more incumbent providers should review the competitive landscape annually to ensure their providers are still the best ones for their requirements. While no single public cloud IaaS provider is an optimal fit for all use cases, several providers can now serve all use cases at a "good enough" level. Nevertheless, many organizations that make significant use of public cloud IaaS will maintain enterprise agreements with at least two such providers and will have policies that determine which provider is used, based on the use case and requirements.

Product/Service Class Definition

Cloud computing is a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service using Internet technologies. Cloud 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). In this document, we use the term "cloud IaaS" synonymously with "cloud compute IaaS."

Cloud IaaS is 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 are metered by use. Self-service interfaces are exposed directly to the customer, including a Web-based UI and an API. In public cloud IaaS, the resources are multitenant and hosted in the service provider's data centers.

Public cloud IaaS is typically used to substitute for VMs and associated infrastructure that are running within a customer's own data center (although public cloud IaaS is most frequently used for new workloads that have never resided within a traditional data center). Many buyers are attracted by the self-service capabilities, which require no interaction with the provider or other human intervention. Also, because the resources are metered by the hour and can usually be bought without any kind of contractual commitment, public cloud IaaS is often perceived as an inexpensive alternative to traditional IT infrastructure.

No private cloud IaaS offerings are evaluated, whether industrialized or customized.

Critical Capabilities Definition

Public cloud IaaS needs to be evaluated for its technical suitability to the needs of particular workloads, as well as the organization's governance needs. This report examines eight broad critical capability areas that IT organizations should consider when evaluating public cloud IaaS offerings.

It is important to note that these are broad categories, not granular capabilities. They are inclusive of a range of features, and we do not provide a comprehensive list of these features. Because each of the categories includes a large number of features, the scoring in each category is directional. In general, a score of 3 indicates that a provider is able to fulfill the most critical features in that category. However, it is possible that a provider may be missing some important features in that category, yet has other strengths that increase its score in that category. You will need to conduct an in-depth evaluation of your shortlisted providers to determine whether they meet your specific needs.

This Critical Capabilities report is not intended to be a granular evaluation of provider capabilities. If you are seeking an in-depth technical evaluation of providers, you should consult Gartner's "Evaluation Criteria for Cloud Infrastructure as a Service" and the associated evaluations of individual providers against those criteria.

If you are looking for an evaluation of providers in the broader context of the entire cloud IaaS market, including private cloud IaaS services, see "Magic Quadrant for Cloud Infrastructure as a Service, Worldwide." Keep in mind, however, that a Magic Quadrant is not a product evaluation. It considers many business factors as well, and it looks at providers' recent execution and vision for the future. Furthermore, the Product/Service rating of a provider on the Magic Quadrant can be significantly different from its rating in this critical capability evaluation, since this report takes into account only one specific public cloud IaaS offering for each provider, whereas the Magic Quadrant takes into account providers' entire cloud IaaS portfolios.

Note that this Critical Capabilities report considers only those features that are available strictly within the context of the provider's public cloud IaaS offering. Importantly, "hybrid" capabilities that require the use of dedicated servers (that cannot be directly provisioned via the GUI and API for public cloud IaaS) are not counted. All capabilities must be provided as part of the standard, industrialized, fully automated public cloud IaaS offering. Capabilities that require hosted equipment, hosted software, leases or monthly rentals, managed services, provider-customized implementations, partner-provided services, or other services that are not a fully integrated element of the offering do not count for the purpose of this evaluation. The sole exception is in the security and compliance portion of this evaluation; see the section on that critical capability for details. A provider's cloud IaaS offering may incorporate PaaS-level capabilities, as long as such capabilities are fully integrated, with a single self-service portal and catalog, unified billing, shared identity and access management, and integrated low-latency network and security contexts.

This market is changing extremely quickly. Some providers release features as often as several times per week, and many providers release features at least once a quarter. Providers do occasionally remove existing capabilities as well. When evaluating service providers, ensure that you understand the current state of each provider's offering. The quantitative assessment is current as of May 2015 (matching the Magic Quadrant) for features that are in general availability; the description of each vendor is accurate as of August 2015.

Because the market is evolving so quickly, the baseline capabilities expected from providers are increasing each year. Provider scores may change significantly with each new iteration of this Critical Capabilities research, since the expected minimum capabilities increase, and the capabilities of each provider may advance significantly. In many cases, provider scores may have decreased since 2014, despite improvements in the provider's offering; this is due to the rise in customer expectations and therefore an increase in the capabilities required, along with a stricter method of scoring that excludes all human-powered managed services.

The critical capability categories are as follows:

Compute resilience: VM availability

Architecture flexibility: Compute, storage and networking options

Security and compliance: Security controls, risk management and governance

User management: Governance of a large numbers of users

Enterprise integration: Network integration, data migration and workload portability

Automation and DevOps enablement: IT operations management and developer enablement

Scaling: Scalability of the service, scaling applications and workloads

Big data enablement: Large-scale data processing and batch computing

These categories are described in detail below.

Compute Resilience

This category is focused on features that are important for VM availability.

Capabilities in this category include:

Autorestart (rapid detection of physical host failure and automatic restart of the VMs — on another host, if necessary)

VM-preserving host maintenance (the ability to perform maintenance on the host server, such as host OS and kernel updates, without rebooting guest VMs)

VM-preserving data center maintenance (the ability to perform data center and hardware maintenance without impacting guest VMs, usually implemented via live migration of VMs)

Affinity and anti-affinity in VM placements

Automated replication across data centers

While the availability of the control plane and other resource elements are considered here, the emphasis is strongly on VM availability, which is important for workloads that assume infrastructure resilience. Most non-cloud-native applications are architected with the assumption of compute resilience, and most enterprise virtualization environments take advantage of the compute resilience features of the hypervisor.

To meet basic expectations in this category, a provider should have autorestart, VM-preserving host maintenance and VM-preserving data center maintenance.

Architecture Flexibility

This category encompasses features that provide a customer with a breadth of resource types and architectures.

For compute resources, flexibility means a broad range VM sizes, along with other options such as bare-metal (nonvirtualized) servers, VMs on single-tenant hosts, multiple hypervisor choices and container-based capabilities (such as a Docker container service). Ideally, a provider should allow flexible (nonfixed) VM sizes — VMs that can have an arbitrary combination of the number of vCPUs and the amount of RAM. If the provider offers specific (fixed) VM sizes instead, the broadest possible number of combinations, representing varying ratios of vCPUs to RAM, is desirable.

For storage resources, flexibility means different types of storage and multiple performance tiers.

For network resources, this means the ability to create complex network topologies, as well as support useful features such as static IP addresses and the ability to have multiple virtual network interfaces per VM.

To meet basic expectations in this category, a provider should offer all the following:

Flexible VM sizes, or a full range of vCPU-to-RAM ratios for fixed-size VMs

VM-independent block storage, including an option for SSD-based storage

Self-service creation of complex hierarchical network topologies

Security and Compliance

This category encompasses features that are important to security, compliance, risk management and governance.

Capabilities in this category include specific security measures, such as network ACLs, IDS/IPS, MFA and encryption. This category also includes aspects such as the availability of audits, logging and reporting, and the ability to use the service if you have regulatory compliance needs, such as those of the Payment Card Industry Data Security Standard (PCI DSS), the Federal Information Security Management Act (FISMA) and HIPAA BAA.

This category encompasses both provider-supplied capabilities that are inherent to the service and are the provider's responsibility to manage, and provider-supplied security controls that are the customer's responsibility to use appropriately. Security is a shared responsibility; customers need to enable and configure appropriate controls.

Some providers offer managed security services that cannot be consumed in an on-demand, self-service fashion. Such capabilities are not included in the scoring, but a provider may be able to provide a higher degree of security if such capabilities are used.

Because security is a major concern for most customers using public cloud IaaS, the capabilities necessary to meet basic expectations in this category are extensive and include all the following:

The service integrates a self-service stateful firewall.

DDoS attack mitigation is provided for all customers.

Traffic between the provider's cloud data centers is sent over a private WAN, not the Internet.

MFA is provided as an option.

Customer changes and provisioning actions are logged, and the logs are retained for 90 days.

Administrative credentials for VMs are issued in a secure fashion.

Provider's personnel are subject to background checks, provider's personnel cannot log into customer compute instances unless the customer has purchased managed services from provider, and all administrative access is logged.

Previously used storage is overwritten before it is reallocated to another customer.

Physical media is sanitized before disposal, in accordance with the NIST SP 800-88 standard.

ISO 27001 (see Note 1) and SOC 1, 2 and 3 (see Note 2) or equivalent audits are available for customers to review.

User Management

This category encompasses features that are necessary to provision and govern multiple users of the service.

User management and governance capabilities are particularly important if you have large development, engineering or research teams. This covers aspects such as RBAC, quotas, leases and integration with enterprise directory services.

To meet basic expectations in this category, a provider should support all the following:

Multiple users per account and multiple API keys per account

Granular RBAC for both users and API keys

Integration with Active Directory and SAML-based single sign-on

Billing alerts

Enterprise Integration

This category encompasses features that are needed to operate in a hybrid IT environment.

These capabilities include secure extension of the organization's WAN, data migration features and workload portability features, along with the ability of an enterprise to license needed software, including the provider's software marketplace capabilities.

To meet basic expectations in this category, a provider should offer all the following features:

- Import of customer-built VM images
- Customization of VM images offered by the provider
- Snapshot VMs and use of snapshots as images
- Import and export of data on physical media
- Allows customers to directly extend their enterprise WAN to the cloud infrastructure
- Uses the customer's choice of private IP addresses from RFC 1918 address allocations
- Open software marketplace of third-party and open-source software that can be deployed in a single click and billed through the provider

Automation and DevOps Enablement

This category encompasses ITOM features, particularly those necessary to manage infrastructure in a DevOps fashion. It also includes software infrastructure capabilities that enhance developer productivity.

Features in this category include monitoring, service catalogs, templates, configuration management, application life cycle management (ALM) and metadata tagging. Other automated capabilities, such as database as a service, are also considered here.

Because one of the key advantages of cloud IaaS is "infrastructure as code" — the ability to have programmatic access to infrastructure — API capabilities are considered in all the categories of capabilities. However, this category also includes the quality of API access, including continuous availability of the control plane and API, and responsiveness to a large number and high rate of API requests.

To meet basic expectations in this category, a provider should offer all the following:

- Access to all functionality via the API (customers can do anything via the API that they can do via the portal)
- No maintenance windows that result in the control plane or API being unavailable
- Metadata tagging for compute and storage elements
- Self-service monitoring, including the ability to generate alerts
- Relational database as a service for at least one common relational database platform
- NoSQL database as a service
- In-memory caching as a service that supports at least one common caching protocol

Scaling

This category encompasses capabilities related to scaling applications and workloads.

Features in this category include local and global load balancing, integrated CDN, autoscaling and the resizing of existing resources, such as VMs and storage volumes. Speed of provisioning is also very important.

To meet basic expectations in this category, a provider should offer all the following:

- Provisioning of a single Linux VM in five minutes or less
- Provisioning of 20 Linux VMs in five minutes or less
- Resizing a VM without reprovisioning it
- Autoscaling based on a schedule or monitoring trigger
- Front-end and back-end load balancing
- Global load balancing with latency-based request routing
- Sufficient available capacity to permit customers to burst provision up to 10 times their normal baseline of consumption, in real time and without prior notice

Big Data Enablement

This category encompasses features that are typically desired for large-scale data processing.

Capabilities in this category include access to large VM sizes, large quantities of capacity on demand and GPUs. This category also covers capabilities such as object storage and services for Hadoop, unstructured data stores such as NoSQL databases, data ingest and data flow (big data is used as a convenient catch-all label for this criterion, rather than literally encompassing big-data-specific capabilities).

To meet basic expectations in this category, a provider should offer all the following:

- VMs with up to 32 vCPUs and 256GB of RAM
- Storage volume sizes of at least 2TB
- Sufficient available capacity to provision up to 1,000 VMs, in real time and without prior notice

- Object-based cloud storage
- Hadoop as a service, or one-click provisioning of a curated Hadoop solution

Use Cases

All use cases have been constructed with the needs of enterprises and midsize businesses in mind — organizations that have existing IT environments, infrastructure and applications, along with security and compliance requirements. Technology startups and other "greenfield" projects are likely to have different needs and criteria.

Application Development

This use case is focused on the needs of large teams of developers that are building new applications.

Many organizations begin their use of public cloud IaaS with this use case, often with a single developer in ad hoc adoption. As usage grows from an individual developer to the entire development organization, however, so do the needed capabilities. In this use case, we consider an application development environment for a large team of developers that must have appropriate governance, security and interoperability with the organization's internal IT infrastructure — one that should enhance that team's productivity with self-service, automated capabilities.

Batch Computing

This use case includes HPC, data analytics and other one-time (but potentially recurring), short-term, large-scale, scale-out workloads.

Batch computing is particularly well-suited to, and may be an exceptionally cost-effective use of, cloud IaaS. Big data enablement capabilities are the majority of the weighting. Since many such workloads depend on a high degree of automation, consideration is given to those aspects as well. Enterprise integration also has some importance, because such workloads often use data that originates on-premises.

Cloud-Native Applications

This use case includes applications at any scale, which have been written with the strengths and weaknesses of public cloud IaaS in mind.

Cloud-native applications assume that resilience must reside in the application and not in the infrastructure (low "compute resilience" weighting), that the application can run well in a variety of underlying infrastructure configurations (low "architecture flexibility" weighting), that the customer's IT organization will attend to security concerns (low "security and compliance" weighting), and that there are only minimal integrations with existing on-premises infrastructure and applications (low "enterprise integration" weighting). Automation, API capabilities and scale-out capabilities are, however, extremely important. Because many such applications have big data aspects, the big data enablement capability also receives a high weighting in this use case.

General Business Applications

This use case includes all applications that were not designed with the cloud in mind, but that can run comfortably in virtualized environments.

In this use case, which can include mission-critical production, applications are designed with the expectation that the infrastructure is resilient and offers consistently good performance. An organization intending to move existing enterprise applications into the cloud typically places a strong emphasis on security, and the public cloud IaaS needs to interoperate smoothly with the existing internal IT infrastructure. To gain more benefit from moving to the cloud, the organization needs the service to deliver additional value-added automation, but the organization is unlikely to make much use of the API, except possibly via third-party tools.

Vendors Added and Dropped

Added

- Interoute
- NTT Communications

Dropped

GoGrid. GoGrid was acquired by Datapipe in January 2015. Due to the timing of the acquisition and integration-related uncertainty, we could not assess either vendor for this Critical Capabilities research.

HP. While HP continues to operate its cloud IaaS offering (HP Public Cloud), it is no longer actively seeking to market and sell this offering. It no longer has sufficient market share to qualify for inclusion in this Critical Capabilities research.

Inclusion Criteria

The vendor inclusion criteria for this report are identical to those for "Magic Quadrant for Cloud Infrastructure as a Service, Worldwide."

All the services in this evaluation meet the following criteria:

- They are public cloud IaaS (by Gartner's definition of the term).
- The service is in general availability and is offered globally.
- The service's data centers are in at least two metropolitan areas, separated by a minimum of 250 miles, on separate power grids, with SSAE 16, ISO 27001 or equivalent audits (see Notes 1 and 2).
- A Web services API is available to all customers.
- There can be multiple users or API keys per account, with RBAC.
- Provisioning occurs in real time, with the smallest available Linux VM available within 10 minutes.

Applications can be scaled beyond the capacity of a single physical server.

There is an allowable VM size of at least eight vCPUs and 32GB of RAM.

Customers can securely extend their network into the public cloud IaaS offering.

There is an SLA for compute, with a minimum of 99.5% availability.

Customers can receive an invoice, and multiple accounts can be consolidated under one bill.

Customers can negotiate a customized contract.

The provider offers 24/7 support, including phone support (in some cases, this is an add-on rather than being included in the base service).

All the providers in this evaluation are among the top 15 providers by Gartner-estimated market share or mind share for the relevant segments of the overall cloud IaaS market (public and industrialized private cloud IaaS, excluding small deployments of two or fewer VMs). If a provider has multiple offerings that meet our definition for public cloud IaaS, we have selected the offering that we expect Gartner clients are most likely to purchase.

There are many additional providers of public cloud IaaS that are worthy of your consideration, even though they are not included in this report. Providers that are regional or have less market share are not included in this report, even if they have offerings superior to those of included providers.

See Table 1 for critical capability weightings in use cases.

Table 1. Weighting for Critical Capabilities in Use Cases

Critical Capabilities	**Application Development**	**Batch Computing**	**Cloud-Native Applications**	**General Business Applications**
Compute Resilience	1%	1%	5%	15%
Architecture Flexibility	10%	6%	8%	15%
Security and Compliance	10%	1%	3%	20%
User Management	25%	1%	2%	5%
Enterprise Integration	15%	5%	2%	25%
Automation and DevOps Enablement	33%	14%	40%	10%
Scaling	5%	7%	20%	8%
Big Data Enablement	1%	65%	20%	2%
Total	**100%**	**100%**	**100%**	**100%**
				As of October 2015

Source: Gartner (October 2015)

This methodology requires analysts to identify the critical capabilities for a class of products/services. Each capability is then weighed in terms of its relative importance for specific product/service use cases.

Critical Capabilities Rating

Each of the products/services has been evaluated on the critical capabilities on a scale of 1 to 5; a score of 1 = Poor (most or all defined requirements are not achieved), while 5 = Outstanding (significantly exceeds requirements). See Table 2.

Table 2. Product/Service Rating on Critical Capabilities

Critical Capabilities	**Amazon Web Services**	**CenturyLink**	**CSC**	**Dimension Data**	**Fujitsu**	**Google**	**IBM (SoftLayer)**	**Interoute**	**Joyent**	**Microsoft**	**NTT Communications**	**Racks**
Compute Resilience	3.7	4.7	4.1	2.8	3.2	3.9	1.0	3.5	2.5	3.0	2.4	
Architecture Flexibility	4.7	3.6	3.2	3.7	1.8	3.0	2.7	3.9	2.8	2.7	1.7	
Security and Compliance	4.5	2.8	1.9	2.9	1.3	3.4	2.8	2.0	2.0	3.8	2.1	

User Management	4.9	3.0	2.8	4.0	2.1	2.0	1.9	1.0	3.6	2.1	1.4	
Enterprise Integration	4.5	2.5	3.6	3.3	2.3	1.9	2.9	2.9	1.4	4.4	2.2	
Automation and DevOps Enablement	5.0	3.0	2.5	1.7	2.7	3.4	2.5	2.5	2.0	4.1	2.2	
Scaling	5.0	3.0	3.0	1.7	1.9	4.0	2.0	2.7	2.2	4.4	3.1	
Big Data Enablement	4.8	2.4	1.8	2.8	1.8	4.0	2.5	2.5	3.1	4.3	2.0	

Source: Gartner (October 2015)

Table 3 shows the product/service scores for each use case. The scores, which are generated by multiplying the use-case weightings by the product/service ratings, summarize how well the critical capabilities are met for each use case.

Table 3. Product Score in Use Cases

Use Cases	Amazon Web Services	CenturyLink	CSC	Dimension Data	Fujitsu	Google	IBM (SoftLayer)	Interoute	Joyent	Microsoft	NTT Communications	Racks
Application Development	4.81	2.98	2.78	2.86	2.22	2.83	2.42	2.30	2.42	3.48	1.99	
Batch Computing	4.81	2.64	2.19	2.66	1.97	3.72	2.48	2.61	2.77	4.15	2.10	
Cloud-Native Applications	4.84	3.00	2.61	2.25	2.25	3.58	2.35	2.67	2.37	3.99	2.29	
General Business Applications	4.53	3.17	3.04	2.94	2.15	3.03	2.40	2.80	2.16	3.67	2.16	

Source: Gartner (October 2015)

To determine an overall score for each product/service in the use cases, multiply the ratings in Table 2 by the weightings shown in Table 1.

Acronym Key and Glossary Terms

ACL	access control list
ALM	application life cycle management
ATO	Authority to Operate
AWS	Amazon Web Services
BAA	business associate agreement
CaaS	Compute-as-a-Service
CDN	content delivery network
CLC	CenturyLink Cloud

DDoS	distributed denial of service
EBS	Elastic Block Store
EC2	Elastic Compute Cloud
FedRAMP	Federal Risk and Authorization Management Program
FISMA	Federal Information Security Management Act
GB	gigabyte
GCE	Google Compute Engine
GCP	Google Cloud Platform
GPU	graphics processing unit
GxP	good x practice
HIPAA	Health Insurance Portability and Accountability Act
HPC	high-performance computing
I/O	input/output
IaaS	infrastructure as a service
IDS	intrusion detection system
IEC	International Electrotechnical Commission
IPS	intrusion prevention system
ISAE	International Standard on Assurance Engagements
ISO	International Organization for Standardization
ITOM	IT operations management
JAB P-ATO	Joint Authorization Board Provisional ATO
KVM	Kernel-based Virtual Machine
MFA	multifactor authentication
MPLS	Multiprotocol Label Switching
NIST	National Institute of Standards and Technology
PaaS	platform as a service
PCI DSS	Payment Card Industry Data Security Standard
RBAC	role-based access control
RDS	Relational Database Service
SaaS	software as a service
SAS	Statement on Auditing Standards
SDN	software-defined network
SOC	Service Organization Control
SRP	shared resource pool
SSAE	Standards for Attestation Engagements
SSD	solid-state drive
vCD	vCloud Director
vCPU	virtual CPU

VDC	Virtual Data Centre
VM	virtual machine
VPC	Virtual Private Cloud
VPE	Virtual Private Edition
VPN	virtual private network

Evidence

Scoring for this Critical Capabilities report was derived from recent independent Gartner research on the cloud IaaS market. Each vendor responded in detail to an extensive primary-research questionnaire covering the business and the technical features of their cloud IaaS offerings. Gartner analysts tested services, reviewed service documentation, corresponded with the vendors on the details of certain offerings and conducted reference checks with end users. Gartner also conducted thousands of client inquiries with prospective and current customers of public cloud IaaS between 2013 and 2015, and it currently conducts more than 1,000 such inquiries each quarter.

Note 1
ISO 27001

International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 27001 is an international standard for information security management systems (see "Security Research Roundup for ISO 27001 Compliance").

Note 2
SSAE 16 and SOC 1, 2 and 3

In 2011, the well-known Statement on Auditing Standards No. 70 (SAS 70) standard was replaced by the Statement on Standards for Attestation Engagements (SSAE) 16 standard (aka Service Organization Control Reports 1 [SOC 1]) (see "SOC Attestation Might Be Assurance of Security … or It Might Not").

Note 3
FedRAMP

These providers possess a Federal Risk and Authorization Management Program (FedRAMP) Authority to Operate (ATO) for specific services within their portfolio. Some possess an agency ATO, and others a FedRAMP Joint Authorization Board Provisional ATO (JAB P-ATO). Either allows the covered services to be used for FedRAMP-compliant needs.

Critical Capabilities Methodology

This methodology requires analysts to identify the critical capabilities for a class of products or services. Each capability is then weighted in terms of its relative importance for specific product or service use cases. Next, products/services are rated in terms of how well they achieve each of the critical capabilities. A score that summarizes how well they meet the critical capabilities for each use case is then calculated for each product/service.

"Critical capabilities" are attributes that differentiate products/services in a class in terms of their quality and performance. Gartner recommends that users consider the set of critical capabilities as some of the most important criteria for acquisition decisions.

In defining the product/service category for evaluation, the analyst first identifies the leading uses for the products/services in this market. What needs are end-users looking to fulfill, when considering products/services in this market? Use cases should match common client deployment scenarios. These distinct client scenarios define the Use Cases.

The analyst then identifies the critical capabilities. These capabilities are generalized groups of features commonly required by this class of products/services. Each capability is assigned a level of importance in fulfilling that particular need; some sets of features are more important than others, depending on the use case being evaluated.

Each vendor’s product or service is evaluated in terms of how well it delivers each capability, on a five-point scale. These ratings are displayed side-by-side for all vendors, allowing easy comparisons between the different sets of features.

Ratings and summary scores range from 1.0 to 5.0:

1 = Poor: most or all defined requirements not achieved

2 = Fair: some requirements not achieved

3 = Good: meets requirements

4 = Excellent: meets or exceeds some requirements

5 = Outstanding: significantly exceeds requirements

To determine an overall score for each product in the use cases, the product ratings are multiplied by the weightings to come up with the product score in use cases.

The critical capabilities Gartner has selected do not represent all capabilities for any product; therefore, may not represent those most important for a specific use situation or business objective. Clients should use a critical capabilities analysis as one of several sources of input about a product before making a product/service decision.

(http://gtnr.it/1KsfgQX)

© 2015 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. or its affiliates. This publication may not 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 Usage Guidelines for Gartner Services (/technology/about/policies/usage_guidelines.jsp) 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. Gartner provides information technology research and advisory services to a wide range of technology consumers, manufacturers and sellers, and may have client relationships with, and derive revenues from, companies discussed herein. 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. (/technology/about/ombudsman/omb_guide2.jsp)"

About (http://www.gartner.com/technology/about.jsp)

Careers (http://www.gartner.com/technology/careers/)

Newsroom (http://www.gartner.com/newsroom/)

Policies (http://www.gartner.com/technology/about/policies/guidelines_ov.jsp)

Site Index (http://www.gartner.com/technology/site-index.jsp)

IT Glossary (http://www.gartner.com/it-glossary/)

Contact Gartner (http://www.gartner.com/technology/contact/contact_gartner.jsp)