Gartner.

Hype Cycle for Infrastructure Strategies, 2020

Published: 8 July 2020 **ID:** G00450228

Analyst(s): Philip Dawson, Nathan Hill

This Hype Cycle focuses on infrastructure architecture, automation/intelligence, AI/ML, IoT and hyperconverged innovations. Cloud- and container-based delivery and consumption models disrupt and provide opportunities for I&O leaders, while software-defined innovations are mature.

Table of Contents

List of Tables

List of Figures

Analysis

What You Need to Know

For 2020, the infrastructure strategies Hype Cycle shifts focus to consumption, intelligence/automation, architectures, cloud, network, resilience and infrastructure delivery. Technologies that are more specific and exclusive to edge and digital workspace situations have been moved to separate Hype Cycles. In addition, software-defined (SD) has not only rationalized to fewer innovation profiles, but also to an Obsolete Before Plateau status as SD is becoming tied to vendors' individual offerings and domains, with limited common terms, lack of common integration and more proprietary lock-in. Individual innovation profiles take this SD maturity and obsolescence further.

Integrated systems, hyperconvergence, composable programmable and intelligent infrastructure combine to create intelligent platforms. This trend captures the progression of infrastructure through administration, application and automation, and delivering intelligence as a function of infrastructure. This architectural approach continues to deliver a stream of new technologies requiring development, deployment and management where traditional approaches fall short or are incomplete.

In-memory computing (IMC; beyond DBMS) functions are driving bigger infrastructure storage pools and memory tiers, which, in turn, accelerate the growth of and demand for big data. Infrastructure analytics also drives intelligent platforms through the advent of machine learning tied to artificial intelligence (AI), from business analytics to infrastructure and workload analytics to hybrid delivery.

Containers driven by microservices are rapidly evolving, with new technologies and applications that quantify the measurement and consumption of resources, instead of just provisioning them. Rather than fade away, virtualization is being refreshed and reiterated with new "revirtualized" projects around storage and network virtualization, and continues to abstract software from hardware.

For more information about how peer I&O leaders view the technologies aligned with this Hype Cycle, see "2020-2022 Emerging Technology Roadmap for Large Enterprises."

The Hype Cycle

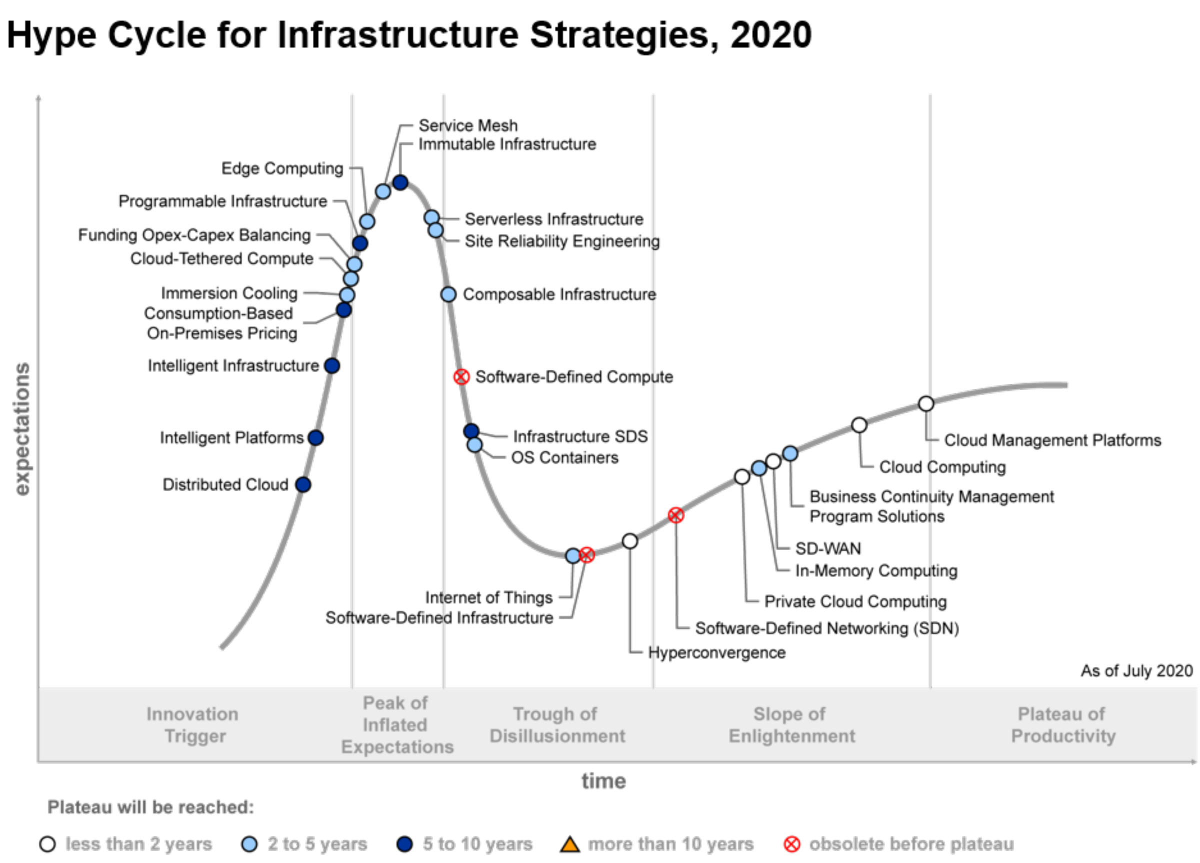
Infrastructure powers the digital business through data centers, cloud, edge, IoT and now, with COVID-19, the home. It extends the physical data center limits into remote office/branch office (ROBO) locations and beyond. Data centers are now "stretched" and elastic, blurring the boundaries between infrastructure, management, cloud, sourcing and virtualization. These hybrid delivery models mix I&O leaders' project silos with a product-driven strategy as they shift to repeatable services.

The appeal of the 2020 infrastructure strategies Hype Cycle is that similar technologies deployed in different use cases, such as ROBO/IoT/edge or hypervisors/containers, can be represented and compared separately, or viewed as a continuum across the Hype Cycle. Infrastructure simplification rationalizes core IT, while accommodating and planning IT expansion, keeping new things simple through repeatable standardization and being cloudlike even on-premises. The core IT and extending ROBO and edge help define thresholds of the new and old aspects of data center infrastructure. This hype is driving infrastructure modernization moving to IMC-scalable architectures, with agility and delivery IT initiatives optimizing cost and targeted SLAs. Software-defined infrastructure (SDI) technologies are sliding into the trough of the Hype Cycle, becoming obsolete. Today, serverless infrastructure, automation and intelligence edge computing, immutable, and composable are all driving innovation in the market. They are fueling the Hype Cycle where infrastructure, through machine learning, can automate and adapt infrastructure composability and programmability.

Innovation profiles new to this year's Hype Cycle are:

- Business continuity management program solutions
- Cloud-tethered compute
- Consumption-based on-premises pricing
- Distributed cloud
- Funding opex-capex balancing
- Immersion cooling
- Intelligent platforms

Figure 1. Hype Cycle for Infrastructure Strategies, 2020



Source: Gartner
ID: 450228

The Priority Matrix

Many of the technologies covered here are transformational, mainly centered around IoT, containers and machine learning (for example, edge computing and serverless infrastructure). High-benefit technologies include cloud, business continuity and consumption delivery models, composable infrastructure, intelligent infrastructure, and other data center initiatives. Other infrastructure technologies provide moderate impact, including intelligent platforms, service mesh, tethered computing and embedded devices.

Figure 2. Priority Matrix for Infrastructure Strategies, 2020

Priority Matrix for Infrastructure Strategies, 2020

benefit	years to mainstream adoption			
	less than two years	two to five years	five to 10 years	more than 10 years
transformational	Cloud Computing	Edge Computing In-Memory Computing Internet of Things OS Containers Serverless Infrastructure Site Reliability Engineering	Infrastructure SDS	
high	Hyperconvergence SD-WAN	Business Continuity Management Program Solutions Composable Infrastructure Funding Opex-Capex Balancing	Distributed Cloud Intelligent Infrastructure Programmable Infrastructure	
moderate	Private Cloud Computing	Cloud-Tethered Compute Immersion Cooling Service Mesh	Consumption-Based On-Premises Pricing Immutable Infrastructure Intelligent Platforms	
low	Cloud Management Platforms			

As of July 2020

Source: Gartner
ID: 450228

Off the Hype Cycle

Several technologies have been removed from this year's iteration of the infrastructure strategies Hype Cycles. They were removed due to their evolution and branching into other Hype Cycle domains, becoming less relevant to the core focus of infrastructure strategies.

Innovation profiles removed from this year's Hype Cycle are:

- Alternative architectures (non-x86)
- Application monitoring and protection
- Cloud application virtualization
- Data center interconnection fabric
- Desktop as a service
- Dynamic optimization technology
- Edge supercomputing
- Embedded devices
- Hardware-based security
- Hyperscale computing
- Micro data centers
- Micro operating systems
- ML-augmented data centers
- SD edge
- Software-defined security
- Virtualization software licensing
- Zero trust network access (software-defined perimeter)

On the Rise

Distributed Cloud

Analysis By: David Smith; Daryl Plummer; Milind Govekar

Definition: "Distributed cloud" refers to the distribution of public cloud services to different physical locations, while operation, governance, updates and evolution of the services are the responsibility of the originating public cloud provider.

Position and Adoption Speed Justification: Distributed cloud computing is a style of cloud computing where the location of the cloud services is a critical component of the model. Historically, location has not been relevant to cloud computing definitions. In fact, the variations on cloud (e.g., public, private, hybrid) exist because location can vary. While many people may claim that private cloud or hybrid cloud requires on-premises computing, this is a misconception. Private and hybrid cloud do not require that the private components are in any specific location. With the advent of distributed cloud, location formally enters the definition of a style of cloud services.

Distributed cloud supports tethered and untethered operation of like-for-like cloud services from the public cloud "distributed" out to specific and varied physical locations. This enables an important characteristic of distributed cloud operation — low-latency compute where the compute operations for the cloud services are closer to those who need the capabilities. This can deliver major improvements in performance as well as reduce the risk of global network-related outages.

User Advice: Begin identifying scenarios where a distributed cloud model will effectively obviate the need for a hybrid cloud model, and where hybrid cloud models, and connectivity and latency matter and will continue to be needed for years to come.

Business Impact: A major notion of the distributed cloud concept is that the provider is responsible for all aspects of the delivery. This restores cloud value propositions that are broken when customers are responsible for a part of the delivery as is true in some hybrid cloud scenarios. It should be noted that while the cloud provider does not need to own the hardware on which the distributed cloud substation is installed, it must take responsibility for how the system is managed and maintained. Otherwise, the value proposition of distributed cloud is compromised.

In hyperscale public cloud implementations, the public cloud is the center of the universe. There has been distribution of cloud services through worldwide regions in public cloud practically since its inception. The major hyperscale cloud providers have different geographic regions around the world, all are centrally controlled and managed and provided by the public cloud provider.

Now, with distributed cloud, we are extending that distributed concept out to the edge and into next-generation hybrid environments such as Microsoft's Azure Stack Hub, Oracle's Cloud at Customer, AWS Outposts, Google Anthos and IBM's forthcoming Satellite offering.

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: Amazon; Google; IBM; Microsoft; Oracle

Recommended Reading: "Top 10 Strategic Technology Trends for 2020: Distributed Cloud"

"'Distributed Cloud' Fixes What 'Hybrid Cloud' Breaks"

"The Cloud Strategy Cookbook, 2019"

Intelligent Platforms

Analysis By: Philip Dawson; Nathan Hill

Definition: Intelligent platforms provide the administration composability of infrastructure and programmable API functions with automated infrastructure intelligence. They integrate compute, storage and networking assets with some or the entire application software stack, creating dedicated workload architectures. Intelligent platform vendors also include components such as management tools, OSs and virtualization bought and consumed as a service.

Position and Adoption Speed Justification: Multiple vendors are driving the market for intelligent platforms around integrated systems, hyperconverged infrastructure (HCI), cloud and virtualization. Vendors such as Microsoft, Nutanix and VMware are promoting valid intelligent platform software, and the market momentum around HCI software in cloud is now creating a market for multiple hardware vendors to build software management and integration services. But the existing intelligent platforms market is most influenced by integrated systems' composable, programmable and intelligent functions, with variants that address multiple aspects of resilience and availability across on-premises, off-premises and cloud locations. Intelligent platforms are software integrated as a software service, with intelligent automation and management, and differ from integrated stack systems, which are hardware-integrated dedicated appliances.

User Advice: An intelligent platform provides a degree of balanced workload performance, application optimization and integration, but this comes at the expense of greater vendor dependency and inflexibility for future application customization and workload requirements. Successful intelligent platform implementations require close harmony among data center stakeholders and other vested interests such as procurement and the lines of business.

Business Impact: The promises of all integrated systems (over time) are lower operational costs and increased IT agility (confined to the limits of that system), via automated, near-real-time allocation/reallocation of pooled resources. Intelligent platforms usually add additional potential for differentiated workload performance or application manageability. Platforms running proprietary workloads rarely compete with each other, as the software choice predetermines the hardware options. However, users will frequently compare a more generic intelligent platform solution (with a separate software stack purchase) against a generic integrated system or HCI solution. This can be a hard comparison to make, because pricing strategies vary greatly between the two; integrated solution integrates pricing as part of the shift to consumption-based delivery.

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Embryonic

Sample Vendors: CU Coding; DDN; Hewlett Packard Enterprise (HPE); Microsoft; Nutanix; Oracle; VMware

Recommended Reading: "Drive Administration, Application and Automation Capabilities of Infrastructure-Led Disruption"

"The Road to Intelligent Infrastructure and Beyond"

Intelligent Infrastructure

Analysis By: Philip Dawson; Nathan Hill

Definition: Intelligent infrastructure is built from simple repeatable building block components from multiple sources, integrated and managed in a standardized automated manner. It optimizes infrastructure resources for application consumption through infrastructure machine learning and tuning as software overlays.

Position and Adoption Speed Justification: Intelligent infrastructure encapsulates intelligence and machine learning (ML) into the infrastructure configuration. Building on the capabilities of virtualization, it adds the dynamic hardware composition capability of a composable infrastructure to deliver a hardware configuration that is optimized for a specific application. Adding intelligence and ML on top of this infrastructure composition capability ensures that infrastructure is always optimized for the application load. Intelligent infrastructure additionally adds or feeds the AIOps and AI/ML functions to the intelligence plane. The intelligence plane automates infrastructure and workload provisioning to application consumption.

User Advice: I&O leaders contemplating intelligent infrastructure should accommodate three considerations:

- Select integrated systems infrastructure solutions based on their ability to meet the current business requirements while still offering the flexibility to exploit the intelligent infrastructure innovations delivered over the next five years.
- Increase agility and business alignment by integrating application asset management and sourcing information into the infrastructure intelligence and control planes.
- Prepare for the evolution of applications and workloads by incorporating intelligence/ML infrastructure functions and persistent memory into your future system requirements.

Business Impact: Intelligent infrastructure builds upon earlier hardware and software innovations, including CI, HCI, SDI and composable, but does not directly replace them. It also feeds off the application API-led programmable infrastructure that tunes infrastructure through system calls and requests. Intelligent infrastructure is the next innovation in delivering optimized systems for applications. In intelligent infrastructure, the "control plane" is enhanced with automation driven by infrastructure analytics ML, to become an "intelligence plane."

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Sample Vendors: Cisco Systems; CU Coding; DDN; Hewlett Packard Enterprise (HPE); IBM; Intel; Microsoft; VMware

Recommended Reading: "Simplify Intelligent Infrastructure by Using Workload Architectures"

"Drive Administration, Application and Automation Capabilities of Infrastructure-Led Disruption"

Consumption-Based On-Premises Pricing

Analysis By: Daniel Bowers

Definition: Consumption-based pricing for on-premises data center infrastructure is an acquisition model that includes a variable payment tied to measured usage.

Position and Adoption Speed Justification: Server and storage hardware vendors have launched or rebranded consumption-based pricing models in the last two years, positioning them as cloudlike or alternatives to public cloud. Examples include Dell Technologies' Flex on Demand, HPE's GreenLake Flex Capacity, and Pure Storage's Pure as-a-Service. These programs meet some users' desire to better align infrastructure costs with resource usage, and to shift infrastructure spending from capital expenditure to cloudlike operating expenditure. As early adopters learn financial and capacity planning lessons, vendors are evolving their offerings to offer a broader range of products and more flexible terms.

User Advice: While hardware-as-a-service offers appeal to organizations seeking to cut infrastructure costs, these programs are typically *not* cheaper than an outright purchase. Required minimum-usage commitment levels and three- to five-year contracts including mandatory services mean consumption-based options are not strictly pay per use. Pricing often ignores market changes during the contract period, such as the long-term industry trend toward lower storage cost per GB. However, these programs enable infrastructure without large upfront capital investment. Consumption-based pricing can also be part of IT's evolution from a cost center toward a service- or product-centric delivery model. IT leaders considering these programs must address any misalignment between variable hardware usage and perpetual- or subscription-based software licensing.

Business Impact: IT leaders — jointly between infrastructure and operations, vendor management and finance — should evaluate the total cost of ownership (TCO) of a range of potential consumption-based program scenarios. For centralized IT organizations with mature chargeback processes, consumption-based programs can improve the linking of costs to specific usage. Organizations must ensure that contract terms match company requirements for categorizing capital versus operating expenses, and that contracts include appropriate end-of-term options (for example, renewals or buyouts). Consumption-based programs change hardware life cycle management, resulting in both high renewal rates and high barriers to exit because it can be difficult to resume traditional acquisition. By removing friction from new infrastructure deployment, these programs also risk allowing unchecked growth of storage resources.

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: Cisco; Dell Technologies; Hewlett Packard Enterprise (HPE); Hitachi Vantara; Lenovo; NetApp; Pure Storage

Recommended Reading: "How to Use Consumption-Based Procurement Models for On-Premises Infrastructure"

"Key Considerations for CSOs Moving to a Consumption-Based Subscription Model"

"Create Consumption-Based On-Premises Infrastructure Bundles for Midsize Enterprises"

Immersion Cooling

Analysis By: Martin Reynolds

Definition: Immersion cooling is a type of data center server cooling system that immerses server boards in a nonconductive heat transfer liquid, typically built using an immersion container in a dense, closed system. Immersion systems deliver well-above-average cooling efficiency, enabling compute systems to run at relatively high performance and floor space density.

Position and Adoption Speed Justification: Open liquid cooling, in which the cooling fluid is in contact with the server components through spray or immersion, is an established technology in supercomputing and military applications. These solutions, however, are either too small or too large for typical data center applications. Furthermore, in some cases the systems use oil for cooling, an impractical messy solution for most businesses. Newer systems use liquids that do not wet the boards and fall back into the cooling pool, or liquids that evaporate quickly with little environmental impact. These systems eliminate the handling issues associated with liquid cooling and work well within existing data center and office environments. Several companies now offer commercially viable solutions that offer up to 10 times improvement in floor space compute density, offering data center managers more flexibility in server deployments.

The server boards use custom carriers for standard server motherboards, or are custom-designed motherboards. Immersion cooled systems are relatively quiet, as the fluids used are far more efficient at heat transfer than air. Furthermore, these servers eliminate the significant power consumption of system fans. These systems, in some cases, take the form of pods with robotic handling in a closed system, improving manageability and sealing the cooling system from the outside. These systems are viable now, but their high power density and novel configurations require customized space and power designs.

As the energy and power density of microprocessor systems advance, liquid cooling will become increasingly important in data center engineering. Immersion cooling may well become a preferred solution in data center environments constrained by space or power environments because of its high efficiency.

User Advice: Immersion cooled systems are significantly smaller, quieter and more energy efficient than traditional data center racks. Their initial value will most likely come from outside the data center, where they enable higher compute density at higher energy efficiency and lower noise.

Although the capital cost of the system is typically higher because of the mechanical and cooling infrastructure involved, there are environments where these systems outperform any alternative.

For example, a business might require that its computing take place in, say, a downtown tower block. In these buildings, there is often a floor power budget that covers both equipment and cooling. An immersion cooling system could easily improve power efficiency by 40%, theoretically enabling 66% more energy for computing within the same power budget. A closed system with robotic board handlers can operate in an otherwise open office with minimal sound baffling.

Another application is in medium-scale edge computing nodes or wireless telecom nodes, where immersion cooling helps the system manage the thermal, spatial and power constraints of a remote server bunker, pole or closet. Shipboard or truck-based mobile data centers also benefit from these space and power efficiencies. Some of these systems originate from the cryptocurrency mining space, with an extremely high GPU density. For certain GPU-centric small-scale supercomputing tasks, these systems represent a practical on-premises solution. Therefore, users operating data center equipment in environments constrained by power, space, or cooling should add immersion cooling systems to their options.

Business Impact: Immersion cooling systems enable enterprises to deploy higher levels of compute capability to strategic locations than is possible with conventional air-cooled racks. Key applications include factory automation; large-scale call center operations; edge data centers; data centers in remote or unattended locations. These systems allow greater business efficiency where latency or unsupervised operation are important parameters.

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: Asperitas; Green Revolution Cooling (GRC); Immersion Systems; QCooling; Submer Technologies

Recommended Reading: "How to Create a Data Center Cost Model Suitable for Public Cloud Comparison"

Cloud-Tethered Compute

Analysis By: Tony Harvey; David Wright

Definition: Cloud-tethered compute is a model where MaaS, IaaS, or PaaS are delivered in a customer-controlled environment but managed by the vendor via a network tether from a public or private cloud. The system may require the tether to be continuously connected, for billing, or be able to operate disconnected with periodic connectivity. Updates are managed by the vendor, removing the responsibility of maintenance of the platform from the I&O team.

Position and Adoption Speed Justification: Cloud-tethered compute solutions are starting to become more common. Initially, they were developed by public cloud vendors to provide public cloud services such as IaaS and PaaS in client-controlled environments, for example, Azure Stack Hub and AWS Outposts. Other vendors have started to develop products in this space. For example, companies like Hivecell deliver an edge-focused solution; Dell EMC delivers a VMware-based private cloud with VMware Cloud on Dell EMC; and other vendors like HPE and Lenovo deliver metal as a service and other capabilities through consumption-based models. Current adoption is relatively low, although the potential for growth due to data sovereignty, connectivity and latency issues at the edge is significant.

User Advice: Cloud-tethered compute systems are relatively new to the market and may be missing key capabilities; and some entrants do not have clear SLAs and the commercial terms are not yet fully developed. Be careful when evaluating tethered compute systems to ensure that they provide the features that the development teams require and that you clearly understand the SLAs provided, what happens at the end of the contract, and how upgrades and expansions midterm are handled.

Key areas to consider when evaluating cloud-tethered compute systems:

- Provided as a service: Some or all elements are provided as a service and remain the property of the vendor. You must have a clear understanding of who owns which element and what are the consequences of a service shutdown.
- Connectivity: Does the system require a permanent connection for the tether, or can it operate in disconnected mode? What are the limitations when operating in disconnected mode?
- Response times and hardware maintenance services: Users more used to a Tier 1 vendor service contract for on-premises maintenance may struggle to adapt to a service model that is based on next-business-day, whole-unit replacement or lower.
- Contract terms: Understand what will happen at the end of the term and how any midterm changes, upgrades or additions will affect the termination date and costs.
- Security: Cloud-tethered compute systems typically use a “shared security model,” where some responsibilities belong to the customer and some to the vendor. Although the security of the vendor may be excellent, the customer’s security team must be engaged to ensure all security risks are mutually addressed.
- Variable pricing: Similar to cloud services, monthly or other recurring charges for a cloud-tethered system could vary based on usage, making costs and budgeting less predictable than traditional infrastructure.
- Data sovereignty: Although the data is held in a location under control of the user, the system is not under user control. At a minimum, any data being used on these systems should be encrypted at rest. In addition, cloud-tethered compute vendors should be able deliver secure data deletion and even storage device retention services to ensure that sensitive data does not leave the control of the user.

Business Impact: Cloud-tethered compute represents a new form of computing in which the customer’s IT gives responsibility for and control over some of its data center assets to a cloud provider. It reduces the need to perform many of the routine tasks traditionally performed by the IT staff, especially as related to server maintenance. It will, however, create new needs for skills in contractual analysis, security and spend management for these systems. Application- and data-level security and backup will still remain an I&O responsibility. In many cases, the I&O function may welcome the removal of basic duties related to system maintenance, which will enable it to focus on delivering higher-level services.

Benefit Rating: Moderate

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Sample Vendors: AWS Outposts; Google Anthos; Hivecell; Microsoft Azure Stack; Oracle; VMware

Recommended Reading: "'Distributed Cloud' Fixes What 'Hybrid Cloud' Breaks"

"Prepare for AWS Outposts to Disrupt Your Hybrid Cloud Strategy"

"How to Bring the Public Cloud On-Premises With AWS Outposts, Azure Stack and Google Anthos"

"Best Practices for Tech CEOs to Manage Edge-to-Cloud Products"

"Top Emerging Trends in Cloud-Native Infrastructure"

Funding Opex-Capex Balancing

Analysis By: Kevin Ji; Philip Dawson

Definition: Funding opex-capex balancing, an evolving "as a service" model, is relatively new and increasingly challenging to many organizations. IT leaders are struggling to comprehend and determine the implications for their infrastructure budgets since consumption-based pricing models have greater variability, and this requires new spend management disciplines to be introduced. The implications are not always comprehended or anticipated by the CIO and the CFO and their organizations.

Position and Adoption Speed Justification: Adoption of cloud computing is expected to increase for the foreseeable future, and the shift to public cloud is evident across multiple quantitative analyses. This includes cloud shift, where calculations of the transition are based on grouping IT segments into four categories. Cloud shift uses a defined method of assessing Gartner forecasts for a particular period.

Among many non-IT executives, there is limited comprehension of what the public cloud actually is. Executives don't always appreciate the radically different contract and pricing structures and the associated nuances of the contracts. It is not always clear exactly what organizations are contracting for and what the full financial implications are. When doing cost comparisons, IT leaders cannot simply calculate the difference between buying and amortizing hardware over a three-year period and procuring those same resources by the hour in cloud IaaS. This is not feasible because while the architecture design change and server scalability drive the variable service cost, they do not impact the capex cost in data center design. Some I&O leaders promote this kind of cost estimation because it typically favors the status quo — buying is less expensive than renting. It is, unfortunately, a highly incomplete, and therefore, misleading approach to cost estimation. Instead, the business case for cloud IaaS adoption requires thorough modeling of multiple scenarios.

User Advice: Migrating resource-intensive systems to cloud-native solutions offers material longer-term savings. However, cloud migration for unmodified on-premises systems does not deliver the true cost savings possible from cloud. Clients must leverage vendors to reduce license instances and costs as part of their cloud migration for existing systems. Here is the action plan:

- Analyze your current infrastructure vendor agreements and related spend to fully budget for the true impact of “as a service” payment on your total operating costs.
- Investigate your application portfolio for public cloud IaaS migration and determine the impact on your future infrastructure cost forecast.
- Determine your accounting options by working with finance team including tax advisors.
- Prepare and maintain an opex-capex impact assessment to communicate to the finance team by investigating pricing and related commercial options for public IaaS/PaaS cloud services.
- Establish the appropriate accounting treatment for all of your enterprise cloud costs. Ensure your approach complies with defined cloud migration principles and maximizes your ability to secure funding for your budget.

Business Impact: Traditionally, organizations purchased infrastructure hardware from vendors, configured/customized it, deployed it and then operated it. From a financial standpoint, this meant the significant upfront costs were generally capitalized under infrastructure including hardware/system software accounting standards and then depreciated/amortized (expensed) over the life of the hardware or system software. This approach meant the significant upfront costs were defrayed over the life of the asset. This method was highly desirable for organizations where capital was preferred. This long-established model is typically embraced by most organizations and their CFOs and finance functions.

In the new “as a service” model, the customer contracts to pay a service fee in exchange for the right to access the cloud supplier’s infrastructure services for a specified term. Over time, SaaS usage tends to increase, leading to additional licensing costs for additional users, increases in data storage costs and the inevitable rise in consumption price as a result of a plateauing pattern investment that moves ever upward.

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Adolescent

Sample Vendors: Apptio; CloudCheckr; Flexera (RightScale); Hewlett Packard Enterprise (HPE); Oracle; Turbonomic; VMware

Recommended Reading: "Calculating and Comparing Data Center and Public Cloud IaaS Costs"

"The Impact of Public Cloud on Operating Budgets"

"Manage the Impact of Cloud Applications on Opex and Capex Budgets"

At the Peak

Programmable Infrastructure

Analysis By: Nathan Hill; Philip Dawson; Milind Govekar

Definition: Programmable infrastructure is the concept of using and applying methods and tooling from the software development area to management of IT infrastructure. This includes, but is not limited to APIs, immutability, resilient architectures and agile techniques. It is also referred to as "infrastructure as code."

Position and Adoption Speed Justification: Programmable infrastructure comprises a composable set of programmable building blocks. Programmable infrastructure goes beyond "aaS" (as a Service) offerings that expose programmable interfaces and enable new ways for delivering infrastructure services. Programmable infrastructure strategies can be applied to private cloud, hybrid cloud and infrastructure platforms, as well as public cloud. Its goal is managing the life cycle of infrastructure delivery from provisioning, resizing and reallocation to reclamation, or in the case of elastic external resources, the termination of consumption.

APIs provide programmatic access to I&O services and data (e.g., depending on the workload requirements, an API that fires off automation that sets up a compute environment with CPU, memory and storage; installs software; assigns IP addresses). These are implemented so that I&O consumers (such as developers) can consume services and data to create new business solutions. Thus, I&O staff should be trained in using web technologies (such as HTTP and JSON) to develop these APIs. I&O leaders also should manage APIs as a technology product and implement full life cycle management, including version control and roadmaps.

The maturity of APIs that enable integration across different infrastructure platforms, combined with the scarcity of programmatic skills within I&O, account for the current maturity of programmable infrastructure.

User Advice: Organizations cannot simply apply automation to existing monolithic infrastructure components. Doing so will result in frustration due to the awareness of agility and response demands without fundamental infrastructure components to deliver on requirements — in essence, automation without platform agility.

Infrastructure and operations leaders must:

- Prioritize agility as one of their top goals in pursuit of digital business outcomes.

- Implement a programmable infrastructure by investing in infrastructure automation tools and AIOps (example vendors for these markets are listed below, but no single vendor or platform can enable an organizationwide programmable infrastructure strategy).
- Invest in infrastructure and DevOps, and modernize legacy IT architectures to implement an API-driven infrastructure.
- Look for reusable programmable building blocks as they extend their programmable infrastructure strategy.

Moving to an API-driven infrastructure is the key first step to enabling anti-fragile and sustainable automation through programmatic techniques. Achieving platform agility is not just about refreshing data center infrastructure to modern platforms like HCIs, although this may form part of the strategy. I&O leaders should consider all areas of the platform — cloud-native architectures, public and private cloud, new infrastructure for new products and services, as well as the modernization of legacy infrastructure.

Business Impact: A continuous-delivery approach requires continuous insight and the ability to automate application responses. This ensures that (only) the right infrastructure resources are available at the right time and location, and this is achieved through a programmable infrastructure. Thus, programmable infrastructure ensures optimal resource utilization while driving cost-efficiencies. However, greater value (than cost reduction) can be achieved via programmable infrastructure's ability to drive adaptive automation — responding faster to new business infrastructure demands, driving service quality and freeing staff from manual operations. It helps reduce technical debt, and enables a sustainable and highly responsive IT infrastructure service to the business.

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Embryonic

Sample Vendors: Alibaba Cloud; Amazon Web Services (AWS); Google; IBM; Microsoft; Pivotal; Tencent Cloud; VMware

Recommended Reading: "Digital Platforms Need Programmable Infrastructure"

Edge Computing

Analysis By: Bob Gill; Philip Dawson

Definition: Edge computing describes a distributed computing topology in which data storage and processing are placed close to the things or people that produce and/or consume that information. Drawing from the concepts of mesh networking and distributed computing, edge computing strives to keep traffic and processing local and off the center of the network. Edge balances latency requirement and the bandwidth required for an application, allows for autonomous operation, enables the placement of workloads and data that satisfies regulatory/security demands.

Position and Adoption Speed Justification: Most of the technology for creating the physical infrastructure of edge data centers is readily available. However, widespread application of the topology and explicit application and networking architectures are not yet common outside of vertical applications, such as retail and manufacturing. As IoT demand and use cases proliferate, the acceptance of edge computing as the topological design pattern (namely, the "where" a "thing" is placed in an overall architecture) has dramatically increased interest in edge technologies and architectures. However, the still-nascent state of non-IoT edge applications has prevented more rapid movement along the Hype Cycle since 2018.

User Advice: We recommend the following:

- We urge enterprises to begin considering edge design patterns in their medium- to longer-term infrastructure architectures.
- Immediate actions include simple trials using colocation and edge-specific networking capabilities, or simply placing remote location or branch office compute functions in a standardized enclosure (for example, "data center in a box").
- Some applications, such as client-facing web properties and branch office solutions, will be simpler to integrate and deploy, while data thinning and cloud interconnection will take more planning and experimentation to get right.
- We are beginning to see viable offerings from hyperscale cloud providers in extending their programming models and management systems to on-premises and edge-located devices, complementing their mostly centralized computing model with a distributed analog.
- For distributed applications requiring a consistent, global infrastructure, with less emphasis on IoT or unique physical endpoints, consider an edge infrastructure as a service provider, such as Cloudflare or NetActuate.
- Enterprises must also become familiar with an emerging "Edge-IN" application model, in which edge gateways and hubs serve as the linchpins for deploying heterogeneous, multicloud and multiendpoint applications. These are often based on open-source frameworks and technologies, such as containers and orchestration systems like Kubernetes.

Business Impact: Edge computing has quickly become the decentralized complement to the largely centralized implementation of public cloud. Edge computing solves many pressing issues, such as unacceptable latency and bandwidth limitations, given a massive increase in edge-located data. The edge computing topology enables the specifics of IoT, digital business and distributed IT solutions, as a foundational element of next-generation applications of all kinds.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: Akamai; Amazon; Cisco; Cloudflare; HPE; IBM; Microsoft; Vapor IO; Verizon; ZEDEDA

Recommended Reading: "The Edge Completes the Cloud: A Gartner Trend Insight Report"

"Top 10 Strategic Technology Trends for 2019: The Empowered Edge"

"The Future Shape of Edge Computing: Five Imperatives"

"How Edge Computing Redefines Infrastructure"

Service Mesh

Analysis By: Anne Thomas

Definition: Service mesh is a distributed computing middleware that optimizes communications between application services. It provides proxy and/or lightweight mediation for service-to-service communications, and supports functions such as authentication, authorization, encryption, service discovery, request routing, load balancing, self-healing recovery and service instrumentation.

Position and Adoption Speed Justification: A service mesh addresses the lightweight middleware requirements of service-to-service communications (east-west), especially among microservices running in managed container systems. These technologies are evolving rapidly. Many commercial and open-source solutions are now generally available.

Hype surrounding service mesh technology accelerated in early 2017 when Google, IBM and Lyft launched the Istio open-source project to produce a service mesh for Kubernetes. Numerous vendors now contribute to the project and provide commercial Istio-based products, and many people associate the service mesh market exclusively with Istio, even though it isn't the most mature product in the market.

Many clients have expressed confusion about the relationship between service meshes and other API mediation technologies, such as API gateways and application delivery controllers (ADCs). A service mesh is lighter weight, and therefore doesn't replace traditional API mediators (see "How a Service Mesh Fits Into Your API Mediation Strategy"). Unfortunately, management, federation and interoperability between the various API mediators and service meshes haven't been addressed by the vendor community, yet.

User Advice: Application leaders responsible for API management and microservices middleware should:

- Deliver secure and resilient miniservices and microservices operations by adopting a service mesh.
- Limit code dependence on any particular service mesh technology by favoring approaches that reduce vendor lock-in, such as sidecar proxies (over library-based implementations).
- Reduce cultural issues and turf wars by assigning service mesh ownership to a cross-functional PlatformOps team that solicits input and collaborates with networking, security and development teams.

- Accelerate knowledge transfer and consistent application of security policies by collaborating with I&O and security teams that manage existing API gateways and application delivery controllers.

Business Impact: A service mesh is a powerful piece of middleware that improves development and operations of microservice-based applications. It ensures reliable, resilient and secure service-to-service communications. It provides deep visibility into the services, enabling proactive operations and faster diagnostics. It automates complex communication concerns, thereby improving developer productivity and ensuring that certain standards and policies are enforced consistently across applications.

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: Amazon Web Services; Buoyant; F5; Google; HashiCorp; Istio; Kong; Microsoft; Netflix; VMware

Recommended Reading: "How a Service Mesh Fits Into Your API Mediation Strategy"

Immutable Infrastructure

Analysis By: Steve Riley

Definition: Immutable infrastructure is not a technology capability, rather it is a process pattern in which the system and application infrastructure, once instantiated, is never updated in place. Instead, when changes are required, the infrastructure is simply replaced. Immutable infrastructure could encompass the entire application stack, with in-versioned templates provisioned via APIs, which are most commonly available in cloud IaaS and PaaS.

Position and Adoption Speed Justification: Immutable infrastructure is typically used by organizations that take a DevOps approach to managing cloud IaaS or PaaS; however, it can be used in any environment that supports infrastructure as code. It represents a significant change in process for traditional infrastructure and operations groups. It may manifest as:

- Native cloud capabilities, such as Amazon Web Services (AWS) CloudFormation or Microsoft Azure Resource Manager templates
- Cloud management platforms, such as Flexera
- Software tools, such as HashiCorp's Terraform
- The customer's own automation scripts

Some or all of an application stack will be instantiated in the form of virtual machine images or containers, combined with continuous configuration automation tools that run after initial boot. Containers can be quickly replaced during runtime, while VM replacement is slower and requires greater coordination among other workload components. Containers improve the practicality of implementing immutable infrastructure and will drive greater adoption.

User Advice: Immutable infrastructure ensures that the system and application environment is accurately deployed and remains in a predictable, known-good-configuration state. It simplifies change management, supports faster and safer upgrades, reduces operational errors, improves security, and simplifies troubleshooting. It also enables rapid replication of environments for disaster recovery, geographic redundancy or testing. Cloud-native workloads are more suitable for immutable infrastructure architecture than traditional on-premises workloads. And, because redundancy may be required by CSP terms of service to receive service-level agreement relief, workloads designed with an immutable infrastructure approach lend themselves to easier replication.

The application stack for immutable infrastructure is typically composed of layered components, each of which should be independently versioned and replaceable. The base OS for the master image may be updated using traditional patching tools, or automatically or manually updated. Automation is then used to bundle components into artifacts suitable for atomic deployment, including VM images, container images, storage objects, network connections, and other necessary resources. The scripts, recipes, and other code used for this purpose should be treated similarly to the application source code itself, which mandates good software engineering discipline.

Some organizations that use immutable infrastructure reprovision only when a change is necessary. Others automatically refresh the infrastructure at frequent intervals (known as systematic workload reprovisioning) to eliminate configuration drift, to update components in which vulnerabilities were discovered, or to possibly eliminate advanced persistent threats. Frequent refresh is only practical in environments with fast and reliable provisioning; thus, it benefits strongly from containers. Integrate with a ticketing system so that refreshes can be initiated and tracked to completion.

The use of immutable infrastructure requires strict operational discipline. IT administrators should eliminate the habit of making one-off or ad hoc modifications to avoid configuration drift. Updates must be made to the individual components, versioned in a source-code-control repository, then redeployed so that everything is entirely consistent. No software, including the OS, is ever patched in production. Organizations that use immutable infrastructure may turn off all normal administrative access to instantiated compute resources — for example, not permitting SSH or RDP access. IT leaders should set a hard date for when all new workloads will use immutable infrastructure if technically feasible; deadlines can be effective motivators of behavior change.

None of the vendors listed in this innovation profile sell a product called “immutable infrastructure.” Rather, they offer one or more elements that help to establish an immutable infrastructure style. Expect to purchase multiple tools.

Business Impact: Taking an immutable approach to server and compute instance management simplifies automated problem resolution by reducing the options for corrective action to, essentially, one. This is to destroy and recreate the compute instance from a source image containing updated software or configuration that addresses the problem. Although immutable infrastructure may appear simple, embracing it requires a mature automation framework, up-to-date blueprints and bills of materials, and confidence in your ability to arbitrarily recreate components without negative effects on user experience or loss of state. In other words, getting to that single corrective action is not without effort. Treating infrastructure immutably is an excellent test of the completeness of your automation framework and the confidence of your platform. The immutable approach is a management paradigm, not a technology capability. The long-term outcome is one in which the workload defines the infrastructure, which is the opposite of traditional scenarios.

Benefit Rating: Moderate

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: Amazon Web Services; Ansible; Chef; Fugue; Google; HashiCorp; Microsoft; Puppet; SaltStack; Turbot

Recommended Reading: “Top 10 Technologies That Will Drive the Future of Infrastructure and Operations”

“Programmable Infrastructure Is Foundational to Infrastructure-Led Disruption”

“Adapting Vulnerability Management to the Modern IT World of Containers and DevOps”

“Solution Path for Infrastructure Automation”

“How to Make Cloud More Secure Than Your Own Data Center”

Serverless Infrastructure

Analysis By: Arun Chandrasekaran

Definition: Serverless infrastructure is a model of IT service delivery in which the underlying enabling resources are used as an opaque, virtually unlimited, shared pool that is continuously available without advance provisioning and priced in the units of the consumed IT service. The runtime environment consisting of all the necessary underlying resources (specifically, the compute, storage, networking and language execution environment) required to execute an application or service are automatically provisioned and operated.

Position and Adoption Speed Justification: The term “serverless” is a misnomer, but serverless computing does transform how compute and associated resources are provisioned, operated and consumed. The most prominent manifestation of serverless computing is serverless functions or fPaaS. With fPaaS, application code is packaged into fine-grained units called “functions,” with the execution of these functions delivered as a managed service. The key benefits of serverless fPaaS are:

- Operational simplicity — It obviates the need for infrastructure setup, configuration, provisioning and management.

- "Built-in" scalability — In serverless functions, infrastructure scaling is automated and elastic, which makes it very appealing for unpredictable, spiky workloads.
- Cost-efficiency — In public cloud-based serverless environments, you only pay for infrastructure resources when the application code is running, which exemplifies the "pay as you go" model of the cloud.
- Developer productivity and business agility — Serverless architectures allow developers to focus on what they should be doing — writing code and focusing on application design.

Serverless delivery of IT services has gained broad notice after Amazon popularized its Amazon Web Services (AWS) Lambda function platform as a service (fPaaS). Although some associate the notion of serverless exclusively with fPaaS, the significance of serverless, as demonstrated by the leading vendors (including Amazon, Google and Microsoft), extends beyond functions to an operational model where all provisioning, scaling, monitoring and configuration of the compute infrastructure are delegated to the platform. Examples of such services include AWS Fargate, Amazon Simple Queue Service (SQS), Amazon Athena, Microsoft Azure Container Instances (ACI) and Google Cloud Run, to name a few. Hence, fPaaS is no longer the only form of serverless platform services.

User Advice: Serverless infrastructure does not spell the end of traditional I&O roles. However, it will significantly change the way I&O roles operate. Although perhaps counterintuitive, serverless does require operations but, instead of managing physical infrastructures, I&O leaders increasingly will have to adapt to new serverless realities by:

- Including the cost implications of event-driven application architectures and the pricing models of different vendors to ensure cost governance and budget control when planning for serverless deployments by considering API gateway, network egress and other costs.
- Revising data classification policies and controls to account for the fact that objects in a content store can now also represent code, as well as data.
- Rethinking IT operations from infrastructure management to application governance, with an emphasis on ensuring that security, monitoring, debugging and ensuring application SLAs are being met. In those cases where an on-premises deployment is merited, I&O teams can support fPaaS in the role of service provider.

Business Impact: New application architectures, such as microservice patterns, are enabling unique competitive differentiation for companies that can rapidly scale their applications with the continuous deployment of software features, a high level of resiliency and more automation. Serverless infrastructure, implemented on-premises or off-premises, enables applications to be built quickly and deployed at a large scale. As such, it is suitable for any customer or web-facing activity in which speed of response and dynamic scalability are concerns. For variable workloads, serverless can be economical, compared with alternatives, due to its ability to provision and consume infrastructure resources only when they're needed. On-premises implementations are uncommon today due to data integration and scalability challenges.

To reap the benefits of serverless, organizations must invest time upfront to build a proof of concept (POC) to validate assumptions about the application design, code, scalability, performance and total cost of ownership.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: Alibaba Cloud; Amazon Web Services; Cloudflare; Google Cloud Platform; IBM; Iguazio; Microsoft Azure; Oracle; Red Hat; VMware

Recommended Reading: "A CIO's Guide to Serverless Computing"

"2019 Strategic Roadmap for Compute Infrastructure"

"Evolution of Virtualization: VMs, Containers, Serverless — Which to Use When?"

Site Reliability Engineering

Analysis By: George Spafford; Daniel Betts

Definition: Site reliability engineering (SRE) is a collection of systems and software engineering principles used to design and operate scalable resilient systems. Site reliability engineers work with the customer or product owner to understand operational requirements and define service-level objectives (SLOs). The site reliability engineer then collaborates with IT stakeholders to design and continuously improve systems that will meet the SLOs. For products or platforms that meet SRE guidelines, the engineer may choose to provide operational support.

Position and Adoption Speed Justification: SRE is a discipline originally created by Google, and was described in the 2016 book, "Site Reliability Engineering: How Google Runs Production Systems." Adoption interest continues to grow both by digital-native organizations as well as traditional enterprises. SRE emphasizes the engineering disciplines that lead to resilience, but individual organizations implement SRE in widely varying ways. SRE is a complementary practice for organizations seeking to scale their DevOps activities.

SRE is intended to help manage the risks of rapid change, through the use of service-level objectives (SLOs), "error budgets," monitoring, automated rollback of changes and organizational learning. SRE teams are often involved in code review, looking for problems that commonly lead to operational issues (for instance, an application that does not do log cleanup and therefore may run out of storage). They also ensure that the application comes with appropriate monitoring and resilience mechanisms, and that the application meets SRE approved standards or guidelines set to achieve negotiated SLOs. SRE teams can serve as an operations function and nearly all such teams have a strong emphasis on blameless root-cause analysis. This is to decrease the probability and/or impact of future events and enable organizational learning, continual improvement and reductions in unplanned work.

SRE practices are being adopted by organizations that need to deliver digital business products reliably. These practices require a culture that supports learning and improvement, highly skilled automation practices (and usually DevOps), usage of infrastructure as code capabilities (which usually requires a cloud platform). SRE also uses automation to reduce manual processes, leverages resilient system engineering principles, and an agile development process that employs continuous integration/continuous deployment (CI/CD).

User Advice: Organizations can benefit from SRE principles even if they are not sufficiently mature, agility-focused, or large enough to adopt SRE as a primary operations model. The SRE principles for risk management, release engineering, handling service-level objectives, monitoring, automation, and self-healing can be applied to a broader range of products and platforms. SRE also represents a useful means to scale DevOps initiatives.

An SRE initiative should have an executive sponsor. The first opportunity to begin with should have the following characteristics:

- The target application must change rapidly yet maintain high availability in order to maximize business value. Stakeholders should be politically friendly.
- The pilot must demonstrate sufficient value to improve credibility and support, yet also have an acceptable level of risk, allowing the stakeholders to learn.
- The initial SRE team must have a collaborative engineering mindset, strive to continuously learn and improve, and desire to automate tasks to reduce repetitious manual work, which is known as “toil.” It is often easiest to move DevOps-skilled employees from different parts of the organization, due to the relative difficulty of hiring engineers with SRE experience. A site reliability engineer is typically a software engineer with an excellent understanding of operations, or, less frequently, an infrastructure and operations engineer with strong programming skills.
- There must be clear SLOs that can be continuously monitored and reported against.
- The SRE collaborates with developers to help them learn how to design and build their product to meet the defined SLOs — the SRE is not doing the actual development work or inspecting quality in.
- The application development team must collaborate with the SRE team to meet SLOs. Developers are responsible for a resilient architecture and reliable code. SREs should not spend more than 50% of their time on ad hoc operational activities. Any excess should go to the developers for support.

An iterative approach must be used to start and evolve SRE practices. The teams involved must share experiences and lessons learned.

Business Impact: The SRE approach to DevOps is intended for products and platforms that need to deliver customer value at speed at scale while managing risk. The two primary use cases are to improve reliability of existing products or platforms as well as to in creation of new products or platforms that warrant the investment in reliability.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Recommended Reading: "DevOps Teams Must Use Site Reliability Engineering to Maximize Customer Value"

"SRE and DevOps: End-to-End Accountability"

"Agile and DevOps Primer for 2020"

"Innovation Insight for Chaos Engineering"

"Maverick* Research: Software Testing and the Illusion of Exterminating Bugs"

Sliding Into the Trough

Composable Infrastructure

Analysis By: Daniel Bowers; Philip Dawson

Definition: Composable infrastructure creates physical systems from shared pools of resources using an API. The exemplary implementation uses disaggregated banks of processors, memory, storage devices and other resources, all connected by a fabric. However, composable infrastructure can also aggregate or subdivide resources in traditional servers or storage arrays.

Position and Adoption Speed Justification: Servers, storage and fabrics are traditionally deployed as discrete products with predefined capacities. Individual devices, or set amounts of resources from individual devices, are connected together manually and dedicated to specific applications. Composable infrastructure allows resources to be aggregated through software-defined intelligent automation, enabling infrastructure and operations leaders to achieve higher resource utilization and faster application deployment. Although some blade-based server infrastructures have long included composable networking features, composable infrastructure describes a broader spectrum of capabilities including disaggregation of accelerator, memory and storage resources.

Current implementations are limited in that resources are pooled or restricted to using hardware from a single vendor. We saw modest steps toward greater vendor collaboration in 2020; for example, an agreement between next-generation fabric consortia, Compute Express Link (CXL) and Gen-Z Consortium, to cooperate on standards. A key step in the maturity timeline for composable infrastructure will be technology that can disaggregate DRAM from compute.

User Advice: The deployment of composable infrastructure is appropriate where infrastructure must be resized frequently, or where composability increases the utilization of high-cost components. The majority of current use cases are in multitenant environments where composability allows efficient sharing of pools of accelerators or storage. Another current use case is in test and development environments where infrastructure with varying characteristics must be repeatedly deployed.

Don't replace existing infrastructure to obtain composable infrastructure unless you have sufficient mature automation tools and skills to implement composable features. Verify that your infrastructure management software supports composable system APIs, or that you have the resources to write your own management tools. However, don't avoid infrastructure with composable features. Rather, don't choose such infrastructure *because* of those features unless you are prepared to use them.

Business Impact: Composable infrastructure helps deliver next-generation agile infrastructure where fast development and delivery mandate rapid and continuous integration. Increased utilization of high-cost resources, such as GPU accelerators and storage-class memory, can yield financial savings in multitenant environments. However, a proliferation of vendor-specific APIs and the lack of off-the-shelf software for managing composable systems are headwinds to widespread adoption.

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: Dell Technologies; DriveScale; GigaIO; Hewlett Packard Enterprise (HPE); Intel; Liqid; Western Digital

Recommended Reading: "Understand the Hype, Hope and Reality of Composable Infrastructure"

"Drive Administration, Application and Automation Capabilities of Infrastructure-Led Disruption"

"Decision Point for Data Center Infrastructure: Converged, Hyperconverged, Composable or Dedicated?"

"The Road to Intelligent Infrastructure and Beyond"

Software-Defined Compute

Analysis By: Philip Dawson

Definition: Software-defined compute (SDC) takes the concepts and abstraction of virtualization and applies them across physical servers, hosts and virtual machines. This can happen at many levels, not just inclusive of hypervisors, but it requires an abstracted management and control plane from the data plane.

Position and Adoption Speed Justification: While SDC is similar to virtualization, it differentiates in that the abstraction may not be with virtualization or the hypervisor, but across many layers of the stack on the servers, with a unifying management and control function. The abstractions may be, for example, multiple DBMS instances on a physical DBMS, or a clustered environment. Like software-defined infrastructure (SDI), SDC is offering cloudlike infrastructure without the detailed intimacy and burden of repurposing, self-selecting, metering and charging back applications. SDC may also extend to consolidation of multiple containers of applications, as well as across virtual machines

(VMs) onto the hypervisor. Multiple applications and containers may be consolidated onto single hosts, offering a higher-level alternative in the stack to hypervisor-based virtualization managed by container management frameworks. In 2020, SDC is also being tied to vendor specifics like OS or hypervisor, and not client needs and requirements. Hence, like other software-defined anything (SDx) terms, it is moving toward obsolescence.

User Advice: Consider the SDC attributes of not only VM hypervisors and related high-availability (HA) functions, but also of physical or bare-metal alternatives, as well as lighter-weight containers with related vendor-tied management container frameworks. Virtualization and abstraction happen at many levels across all software-defined domains and within SDC functions. Fulfillment of virtual requests of physical resources can be provisioned through the virtualization of many abstracted SDC functions, which, in turn, are provisioned through a common logical management toolset or framework. New hardware technologies like NVRAM increasingly drive more SDC and related technologies between locations and to the edge tied to vendor-focused management and delivery.

Business Impact: Balance virtualization offerings — the more you virtualize, the more physical resources and hosts need to be measured holistically and managed together, and the more they need to share approaches with nonvirtualized bare-metal workloads. SDC implementations should be strong in managing vendor-tied physical resources, hosts, VMs and containers across the portfolio.

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Obsolete

Sample Vendors: Docker; Hewlett Packard Enterprise (HPE); Huawei; Intel; Microsoft; Oracle; Parallels; Red Hat; Unisys

Recommended Reading: “Market Insight: Extend Software-Defined Infrastructure Offerings to the Edge to Provide New Benefits”

“2019 Strategic Roadmap for Compute Infrastructure”

“Optimize Infrastructure Patterns to Extend Control to the Edge and IoT”

Infrastructure SDS

Analysis By: Julia Palmer; Chandra Mukhyala

Definition: Infrastructure software-defined storage (SDS) abstracts storage software from the underlying hardware providing common provisioning and data services across IT infrastructures regardless of locality and hardware technology. It can be deployed as a virtual machine, container or as storage software on a bare-metal industry standard server, allowing organizations to deploy a storage-as-software package on-premises, at the edge or in the public cloud. This creates a storage platform that can be accessed by file, block or object protocols.

Position and Adoption Speed Justification: Infrastructure SDS changes the delivery model and potentially the economics of enterprise storage infrastructures. Whether deployed independently, or as an element of a hyperconverged infrastructure, SDS alters how organizations buy and deploy enterprise storage. Following web-scale IT's lead, I&O leaders are deploying SDS as hardware-agnostic storage, and breaking the bond from proprietary, external-controller-based (ECB) storage hardware. The power of multicore Intel x86 processors, use of software-based RAID or erasure coding, use of flash and high throughput networking have essentially eliminated most hardware-associated differentiation, transferring the value to storage software. New infrastructure SDS vendors are targeting a broad range of delivery models and workloads, including backup, archiving, big data analytics, high-performance computing, AI supporting structured and unstructured data for virtual machines, containers, and bare-metal workloads. Calculating the total cost of ownership (TCO) benefits of SDS involves comprehensive analysis of savings from both capital expenditure and operating expenditure, including administration, verification, deployment, and ongoing management, maintenance and support, as well as a potential improvement in business agility.

User Advice: Infrastructure SDS is the delivery of data services and storage-array functionality on top of industry standard hardware. Enterprises choose a software-defined approach when they wish to accomplish some or all of the following goals:

- Build a storage solution at a low acquisition price point on commodity x86 platform.
- Decouple storage software and hardware to standardize their data center hardware platforms.
- Establish a scalable solution specifically geared toward modern workloads.
- Build an agile, "infrastructure as code" architecture, enabling storage to be a part of software-defined data center automation and orchestration framework that expands to the public cloud.
- Take advantage of latest innovations in storage hardware before they are supported in traditional ECB storage arrays.

Advice to end users:

- Recognize that infrastructure SDS remains a nascent, but growing, deployment model that will be primarily focused on web-scale deployment agility, but also has applicability at the edge and public cloud deployments.
- Identify use cases for SDS deployment by mapping IT strategy with the specific SDS solution business advantage framework.
- Select SDS vendors that provide support for multiple deployment options, and offer tight hardware reference designs and flexible pricing models.
- Deploy SDS as part of a cohesive software-defined infrastructure (SDI) design, with an emphasis on delivering uniform storage platforms across on-premises, public cloud and edge environments.
- Recognize that SDS may involve substantial work in sizing the underlying hardware and building the total solution on your own versus a plug-and-play appliance.

- Grade SDS products by their ability to be truly hardware-agnostic, API-driven, based on distributed architecture, and capable of supporting edge, core or public cloud deployments.

Business Impact: Infrastructure SDS is a hardware-agnostic platform. It breaks the dependency on proprietary storage hardware and lowers acquisition costs by utilizing the industry standard x86 server platform of the customer's choice. Some Gartner customers report up to 40% TCO reduction with infrastructure SDS that comes from the use of x86 industry standard hardware and lower cost upgrades and maintenance. However, the real value of infrastructure SDS in the long term is increased flexibility and the ability to have common provisioning and data services tools regardless of data locality. I&O leaders that successfully deployed and benefited from infrastructure SDS have usually belonged to large enterprises or cloud service providers that pursued web-scale-like efficiency, flexibility and scalability, and viewed SDS as a critical enablement technology for their IT initiatives. I&O leaders should look at infrastructure SDS not as another storage product but as an investment in improving storage economics and providing data mobility including hybrid cloud storage integration.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: DataCore Software; IBM; NetApp; Nutanix; Red Hat; Scality; StorMagic; SUSE; VMware; WekaIO

Recommended Reading: "The Future of Software-Defined Storage in Data Center, Edge and Hybrid Cloud"

"Top Five Approaches to Hybrid Cloud Storage — An Analysis of Use Cases, Benefits and Limitations"

"Market Guide for Hybrid Cloud Storage"

"An I&O Leader's Guide to Storage for Containerized Workloads"

"Competitive Landscape: Infrastructure Software-Defined Storage"

"Software-Defined Storage Enables Versatile Hybrid-Cloud Architectures"

OS Containers

Analysis By: Thomas Bittman

Definition: OS containers are a shared OS virtualization technology that enables multiple applications to share an OS kernel without conflicting. This is enabled by what is usually called a "container daemon," which provides logical isolation of processes. This enables several applications to share an OS kernel, while maintaining their own copies of specific OS libraries.

Position and Adoption Speed Justification: Containers were used for years to increase the density of lightly used workloads (e.g., Oracle Solaris Zones, Virtuozzo and FreeBSD jails), focused on infrastructure management. Now containers are focused on developer requirements, for agile development, rapid provisioning and real-time horizontal scaling — especially for microservices architecture applications and cloud computing. Docker has become a popular container packaging format for developers, and micro OSs are being developed by most major OS vendors. A container/ cluster ecosystem is maturing that includes open source, such as Kubernetes.

User Advice: Several container technologies are targeting agile development and cloud deployment, including Linux containers (LXC), Docker's libcontainer, Apache Mesos, Microsoft Windows Containers and VMware vSphere Integrated Containers. Several container technologies designed specifically for consolidation and density are mature (e.g., Oracle Solaris Zones and Virtuozzo); however, all container technologies will significantly improve resource utilization and densities — the difference will be in the richness of management tools available for them. Development teams focused on agile development and microservices can leverage container technologies.

OS containers do not replace hypervisor-based virtualization technologies — such as VMware or Kernel-based Virtual Machine (KVM) — because they essentially benefit different kinds of applications, and target different problems. Containers focus on application design and developer requirements, and hypervisors focus on capacity management and infrastructure and operations (I&O) professionals. Security and manageability concerns are often mitigated by deploying containers within VM architectures (either layered or explicitly integrated, such as the Microsoft and VMware offerings). Although VMs are inherently "heavier," most of the "weight" comes with the size of the OS itself. Balancing application needs, developer needs and operational needs will determine whether to run containers on bare metal, on VMs or integrated with VMs.

Business Impact: Container technologies are a part of a development architecture that will help enterprises become more agile overall, with applications that can change quickly, scale rapidly to demand and improve the density of capacity use. In production, containers will be used strategically for new applications designed for containers and agile development, rather than existing, monolithic application architectures. However, for developer ease of use, containers will also be used as wrappers for traditional workloads, pushing enterprises to develop their container management expertise quickly.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Early mainstream

Sample Vendors: Canonical; Docker; Mesosphere; Microsoft; Oracle; Red Hat; Virtuozzo; VMware

Recommended Reading: "Best Practices for Running Containers and Kubernetes in Production"

Internet of Things

Analysis By: Alfonso Velosa; Benoit Lheureux

Definition: The Internet of Things (IoT) is a core building block for digital business and digital platforms. IoT is the network of dedicated physical objects that contains embedded technology to communicate and sense or interact with their internal states and/or the external environment. IoT comprises an ecosystem that includes assets and products, communication protocols, applications, and data and analytics.

Position and Adoption Speed Justification: Gartner's CIO Survey 2020 shows that IoT is regarded by CIOs as one of the top five game-changing technologies, with enterprises vary widely on their IoT adoption depth and maturity. Enterprises on a global basis have ongoing IoT-enabled initiatives for use cases ranging from incremental benefits (for example, asset optimization or compliance reporting) to transformative benefits (for example, product as a service or guaranteed asset uptime). The more developed use cases center on fleet management and industrial equipment maintenance, where ROI is calculated from cost optimization such as reducing maintenance and fuel costs. Many enterprises are now exploring employee and citizen safety solutions using IoT enabled capabilities. Finally, Gartner's 2019 IoT Survey indicates that while enterprises expect a 3-year payback on average for their IoT projects, 42% expect payback in less than 2 years. In the 2020 economic downturn, many clients are pushing for even shorter project paybacks.

The hype has decreased from the highs in 2016 through 2019; we reflect this by moving the profile's position into the trough. Enterprises continue to address cost, complexity and scaling challenges implementing IoT-enabled business solutions, as well as increased adoption of contact-less monitoring solutions, drone inspections, etc. driven by the 2020 pandemic. Challenges include end-to-end integration complexity, the need to bridge cultural divides between IT and operations, confusing vendor marketing, especially as they increasingly shift to IoT-enabled business solutions, security concerns, and the 2020 pandemic disruption on IoT project schedules.

User Advice: CIOs should take action to address IoT concerns across the following areas:

- Business: Measure and deliver IoT value based on digital and strategic business objectives. If you are still experimenting, use a proof of business value approach. Build business cases with project payback of less than 18 months to account for implementation challenges and cultural resistance. Add employee and customer safety to your priority list of IoT projects and capabilities.
- Management: Build or contribute to an IoT center of excellence (COE) composed of IT, operational and business personnel. Use the COE to drive global alignment on best practices, alignment to business objectives, budgeting and people allocation. Remember that IoT is really about business process transformation, so focus on culture change first and technology second to ensure success.
- Architecture: Ensure the architecture teams focuses on both the IT and operational technology portfolio as well as the need to manage a multi-IoT platform approach. Ensure analytics and applications are part of the conversation.

- Skills: Invest in business and architecture skills to support project ideation and prioritization, as well as technical skills for IoT platforms, integration, analytics and security. Drive learning via projects with short-term outcomes, and include business leaders, IT leaders, and front-line workers.
- Vendors: Assess and select providers on how they lower project risk for your enterprise via their vertical market expertise, technical capabilities (including best-of-breed partners) and trained professional services partners. Ensure your vendors integrate into your IT infrastructure.
- Governance: Establish accountability, participation, predictability and transparency policies for IoT — addressing sponsorship, budgets, digital ethics, data ownership and rights to monetize IoT data, etc...
- Risk: Scan for threats from enterprises in your ecosystem who may use IoT capabilities to damage or limit your differentiation and competitiveness.

Business Impact: As an evolutionary business impact, IoT will impact most enterprises' internal operations, differentiation, competitive position, and product strategies. Connected things will help lower costs, drive revenue, and improve enterprise processes in these types of usage scenarios.

Optimization of a range of business processes:

- Cost optimization: Lower operating costs for energy reduction, maintenance minimization, minimizing inventory spoilage, lowering theft
- Operations optimization: Better productivity, increased efficiency, logistics and coordination
- Optimize assets: Asset utilization, health monitoring, reliability, predictive maintenance
- Conserve resources: Energy efficiency and pollution reduction

New revenue strategies:

- Generate revenue via improved products, contractual services, usage-based pricing, and monetizing IoT data.
- Increase engagement: Improved experiences of consumers, citizens and others in order to improve loyalty and increase customer lifetime value.

Safety focus:

- Drive employee and citizen safety by monitoring and checking people's health, shifting to over-the-air updates to avoid in person visits, fall monitoring for the elderly and remote workers.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: ABB; Alibaba Cloud; Altizon; GE Digital; Hitachi Vantara; Tencent; Vodafone

Recommended Reading: “Predicts 2020: As IoT Use Proliferates, So Do Signs of Its Increasing Maturity and Growing Pains”

“Toolkit: Enterprise Internet of Things Maturity”

“Survey: Manufacturers See Quick Return on IoT Projects”

“Forecast: Enterprise and Automotive IoT, Worldwide”

Software-Defined Infrastructure

Analysis By: Philip Dawson

Definition: Software-defined infrastructure (SDI) includes the broad set of software-defined anything (SDx) infrastructure components and the software-defined data center (SDDC). SDI also includes non-data-center infrastructure deployed in Internet of Things (IoT) applications and an SD edge of edge-based adapters, monitoring devices, gateways, appliances and machines.

Position and Adoption Speed Justification: Data center infrastructure is well-covered with compute (SDC), network (SDN) and storage (SDS), but SDI also extends to non-data-center infrastructure with the use of monitoring devices or machines that are software-defined. This is enabled through the use of sensors and adapters that are abstracted through software, becoming SDI in edge, IoT and operational technology (e.g., retail POS), rather than traditional, IT-driven SDI through data center or cloud. In 2020, we are seeing SDI move to vendor-specific silo technology (not heterogeneous service drive) and, hence, obsolete as multivendor interoperable standards.

User Advice: As SDI initiatives roll out, consider the integration and measurement of non-data-center edge infrastructure. Focus on core IT SDI for compute, network, storage and facilities, but consider the impact of SDI on IoT, edge computing, remote office/branch office (ROBO) and other operational technologies. Key verticals operating in multiple, geographically distributed locations (such as retail, manufacturing, retail banking, distribution and utilities) are extending IoT and non-data-center SDI initiatives for new IT operations and functions. Expect SDI to be tied to a specific vendor or technology silo.

Business Impact: With the increase of IoT touching edge-based operational technology, SDI reaches beyond and between SDDCs, and leverages SDI benefits and features for new multimode applications and edge IoT endpoints. However, SDI is now tied to vendor technology not interoperability.

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Obsolete

Sample Vendors: IBM; Intel; Microsoft; Red Hat; VMware; Wipro

Recommended Reading: “Simplify Intelligent Infrastructure by Using Workload Architectures”

"Drive Administration, Application and Automation Capabilities of Infrastructure-Led Disruption"

Hyperconvergence

Analysis By: Philip Dawson; Jeffrey Hewitt

Definition: Hyperconvergence is scale-out software-integrated infrastructure designed for IT leaders seeking operational simplification. Hyperconvergence provides a building block approach to compute, network and storage on standard hardware under unified management. Hyperconvergence vendors build appliances using off-the-shelf infrastructure, engage with system vendors that package software as an appliance, or sell software for use in a reference architecture or certified server. Hyperconvergence may also be delivered as a service or in a public cloud.

Position and Adoption Speed Justification: Hyperconvergence solutions are maturing and adoption is increasing as organizations seek management simplicity. VMware vSAN utilization among VMware ESXi customers, and Storage Spaces Direct utilization among Microsoft Windows Server 2016 and 2019 Data center edition customers are on the rise. Nutanix, an early innovator in HCIS appliances, has largely shifted to a software revenue model and continues to increase the number of OEM relationships. Hyperconvergence vendors are achieving certification for more-demanding workloads, including Oracle and SAP, and end users are beginning to consider hyperconvergence as an alternative to integrated infrastructure systems for some workloads. Meanwhile, suppliers are expanding hybrid cloud deployment offerings. Larger clusters are now in use, and midsize organizations are beginning to consider hyperconvergence as the preferred alternative for on-premises infrastructure for block storage. Meanwhile, a growing number of hyperconvergence suppliers are delivering scale-down solutions to address the needs of remote office/branch office (ROBO) and edge environments typically addressed by niche vendors.

User Advice: IT leaders should implement hyperconvergence when agility, modular growth and management simplicity are of greatest importance. The acquisition cost of hyperconvergence may be higher and the resource utilization rate lower than for three-tier architectures, but management efficiency is often superior. Hyperconvergence requires alignment of compute and storage refresh cycles, consolidation of budgets, operations and capacity planning roles, and retraining for organizations still operating separate silos of compute, storage and networking. Adopt for mission-critical workloads, only after developing knowledge with lower-risk deployments, such as test and development. Workload-specific proofs of concept are an important step in meeting the performance needs of applications. Consider the impact on DR and networking. Test under a variety of failure scenarios, as solutions vary greatly in performance under failure, their time to return to a fully protected state and the number of failures they can tolerate. Consider nonappliance options to enable scale-down optimization of resources for high-volume edge deployments. In product evaluations, consider the ability to independently scale storage and compute, retraining costs, and the ability to avoid additional operating system, application, database software and hypervisor license costs. In large deployments, plan for centralized management of multiple smaller clusters. For data center deployments, ensure that clusters are sufficiently large to meet performance and availability requirements during single and double node failures. While servers are perceived as commodities, they differ greatly in terms of power, cooling and floor space requirements, and performance, so evaluate hyperconvergence software on a variety of hardware platforms for lowest total cost of ownership and best performance.

Business Impact: The business impact of hyperconvergence is greatest in dynamic organizations with short business planning cycles and long IT planning cycles. Hyperconvergence enables IT leaders to be responsive to new business requirements in a modular, small-increment fashion, avoiding the big-increment upgrades typically found in three-tier infrastructure architectures. Hyperconvergence provides simplified management that decreases the pressure to hire hard-to-find specialists. It will, over time, lead to lower operating costs, especially as hyperconvergence supports a greater share of the compute and storage requirements of the data center. For large organizations, hyperconverged deployments will remain another silo to manage. Hyperconvergence is of particular value to midsize enterprises that can standardize on hyperconvergence and the remote sites of large organizations that need cloudlike management efficiency with on-premises edge infrastructure. As more vendors support public cloud deployments, hyperconvergence will also be a stepping stone toward public cloud agility.

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Mature mainstream

Sample Vendors: Cisco Systems; Dell Technologies; Hewlett Packard Enterprise (HPE); Huawei; Microsoft; Nutanix; Pivot3; Red Hat; Scale Computing; VMware

Recommended Reading: "Magic Quadrant for Hyperconverged Infrastructure"

"Critical Capabilities for Hyperconverged Infrastructure"

"Toolkit: Sample RFP for Hyperconverged Infrastructure"

"The Road to Intelligent Infrastructure and Beyond"

"Use Hyperconverged Infrastructure to Free Staff for Public Cloud Management"

Climbing the Slope

Software-Defined Networking (SDN)

Analysis By: Mark Fabbi; Joe Skorupa

Definition: Software-defined networking (SDN) is an architectural approach to designing, building and operating networks that promised increased agility and extensibility by abstracting the network topology and control plane. However, SDN products never made it to mainstream enterprise adoption. Rather, SDN spawned innovations in automation, orchestration, segmentation and the disaggregation of network hardware and software.

Position and Adoption Speed Justification: SDN remains a major topic of discussion across multiple network markets and in many vendor marketing efforts. However, true SDN technologies have not achieved any significant enterprise market traction and should not be considered by any enterprise networking organization. The hope that SDN would allow for the decoupling of the control plane from network hardware and foster independent software innovation never came to fruition and there are effectively no SDN technologies available in the mainstream marketplace today.

While true SDN solutions have not had any significant market adoption, the development of SDN and the threat to established market players had a profound, positive effect on subsequent market developments. SDN clearly influenced the increasing use of white-box switches, the open-source hardware and software movement (supported by the Open Compute Project) and the development of independent network switch software providers. More important to the enterprise market was a shift in focus of traditional network vendors innovation around operations and management. This has led to improvements in agility and automation, simplified operational requirements and a general adoption of function-wide configuration (i.e., management that spans devices in a data center, campus or WAN environments). Without the threat of SDN it is unlikely the operational advances from traditional vendors would have occurred.

User Advice: While SDN is clearly obsolete in the enterprise market there are still many organizations that site SDN as a cornerstone of their future strategy and architecture. What is critical for enterprises is to understand what they are trying to accomplish when they think “SDN.” We recommend the following:

- Don’t get caught up in the hype and vendor claims that commercial products are SDN or engage in any discussions or planning to deploy SDN. SDN is not the answer to any enterprise networking challenge today.
- Focus on the desired outcomes you are trying to achieve such as increased automation, virtual segmentation, external orchestration, control and programmability of the network or decoupling physical hardware from software switch operating systems.
- Select an operational/automation framework first — then decide on networking vendors and products. Decoupled hardware and software and independent network overlays provide a way of establishing long-term operational models that are independent of underlying hardware if that is a desired outcome.
- Evaluate both hardware infrastructure and software overlays approaches. Evaluate not only vendor promises but the operational requirements to actually achieve a stated benefit. Hold any considered or deployed vendor accountable to deliver your desired outcomes and don’t believe marketing hype surrounding SDN-related claims. Evaluate reference accounts and pay particular attention to implementation and ongoing operational costs and investments to ensure that benefits can be realistically achieved.
- Develop cross-functional collaboration and investigate methodologies to better integrate server, virtualization, network, security and application teams. These teams can help identify key use cases — both short-term, such as self-service development environments and

microsegmentation; and long-term, integrating networking more broadly into data center orchestration.

- Allocate time and resources to evaluate technologies and a shortlist of relevant vendors — both incumbent and nonincumbent — in order to arrive at a solution that best meets the needs of the cross-functional organization.

Business Impact: There is no direct commercial impact of full SDN solutions. However, downstream innovation can increase network agility, simplify management, improve security and lead to reductions in operational and capital costs while fostering cross-functional collaboration. New network solutions should focus on reducing or eliminating the "human middleware" problem that has plagued traditional network solutions for the past two decades. By bringing network operations into more streamlined and automated operational process that have been architected in the virtual compute environment, user organizations can bring application deployments in line with the increasing speed of business. Available network overlay technology can create new competitive environments shifting the focus from physical infrastructure to software and operational features.

Benefit Rating: Low

Market Penetration: Less than 1% of target audience

Maturity: Obsolete

Sample Vendors: NEC

Recommended Reading: "State of SDN: If You Think SDN Is the Answer, You're Asking the Wrong Question"

Private Cloud Computing

Analysis By: Thomas Bittman

Definition: Private cloud computing is a form of cloud computing used by only one organization, or one that ensures an organization is completely isolated from others. As a form of cloud computing, it has full self-service, full automation behind self-service and usage metering. It does not have to be on-premises or owned or managed by the enterprise.

Position and Adoption Speed Justification: Private and public cloud computing are at opposite ends of the "isolation" spectrum. As public cloud providers have offered virtual private cloud, dedicated instances and dedicated hosts, the gap between private and public has become a spectrum of isolation choices. Recent offerings from the major cloud providers for on-premises cloud footprints (tethered cloud) have created another, newer form of private cloud computing — and these immature alternatives are keeping private cloud from the Hype Cycle plateau.

Organizations that build a private cloud service are emulating public cloud computing providers to acquire similar benefits — mainly agility, mainly for new cloud-native applications, and mainly for business value and growth. This can be for infrastructure as a service (virtual machines or containers), platform as a service, or, in some situations, software as a service.

Due to cost and complexity, most successful private clouds are built and delivered by third parties.

This term is also used to describe a very different trend, where traditional infrastructures are being modernized with virtualization, some automation and some self-service. In this manner, they are leveraging only some valuable attributes of cloud computing, but are applying them to existing applications with traditional infrastructure requirements. However, because these are different trends, Gartner does not include this form of modernization in our definition of private cloud. But when the goal is IT efficiency or modernization for existing applications, these “just enough cloud” architectures can be beneficial.

User Advice:

- Evaluate third-party options first. These include hosted private cloud, managed services, virtual private cloud alternatives, tethered cloud or public cloud.
- Choose your private cloud strategy based on the necessary return on investment or business goals. If business growth or business value for new applications is the goal, consider a true cloud architecture. If IT efficiency or IT modernization for existing applications is the goal, choose cloud-inspired technologies and methods to implement. Just-enough cloud is often enough.
- Focus on business and application needs first; don’t start with the technology. One technology architecture and operational model cannot support all of the application needs of a typical enterprise. Either build multiple architectures and operational models, or leverage third parties.
- Focus on services that fit the cloud model: standard, high volume and self-service; those that require agility and horizontal scalability; and usages that might be short-lived.
- Consider the long-term roadmap for your private cloud services. Build with the potential to integrate, interoperate or migrate to public cloud alternatives at the appropriate time.
- Manage the scope of work — start small and expand based on the business case.
- Build expertise in managing multiple architectural and operational models, and multicloud — this is more valuable to an enterprise than expertise in building a single cloud architecture.

Business Impact: Cloud computing enables agility that an enterprise can use to react quickly to business requirements in functionality or scale. Due to economies of scale, cloud computing can also improve efficiency and lower costs. However, because leveraging a true cloud computing architecture requires applications and operational models designed for cloud computing, the cost of transformation for existing applications does not always justify the investment.

True private cloud computing is used when enterprises aren’t able to find cloud services that meet their needs in terms of regulatory requirements, functionality or intellectual property protection. True private cloud computing is almost always purpose-built for a specific set of new applications, and its success can be measured in revenue or market share.

When the primary goal of a private cloud is IT efficiency, businesses can reduce costs and improve overall operational efficiency for their existing application portfolios by leveraging cloud technologies where appropriate. They can then add manual or custom intervention, or customized changes as needed, to support those applications.

However, enterprises need to recognize that these are two different goals with different architectures, and trying to accomplish them in a single architecture usually achieves none of the goals well. Being bimodal based on business and application needs makes the most business sense.

Benefit Rating: Moderate

Market Penetration: More than 50% of target audience

Maturity: Mature mainstream

Sample Vendors: Hewlett Packard Enterprise (HPE); IBM; Microsoft; Red Hat; VMware

Recommended Reading: "Rethink Your Internal Private Cloud"

"Building 'Just Enough' Private Cloud With Virtualization Automation"

In-Memory Computing

Analysis By: Philip Dawson; Massimo Pezzini

Definition: In-memory computing (IMC) is an application architecture style which assumes that all the data required by applications for processing is located in the main memory of their computing platforms. In IMC-style applications, a persistent, nonvolatile data store (hard drive or solid-state drive) is used to permanently store in-memory data for recovery purposes, to manage overflow situations, to manage the information life cycle and to transport data to other locations, but is not used as the primary location for the application data.

Position and Adoption Speed Justification: IMC delivers dramatic improvements in performance, scalability and analytic sophistication over traditional architectures. IMC is enabled by a range of rapidly converging software technologies and advancements in hardware architectures.

The emergence of persistent memory in the form of 3D XPoint, Intel Optane and NVM Express (NVMe) has sparked renewed interest in IMC. Meanwhile, adoption has grown, driven by several factors including:

- Tens of thousands of SAP clients (of any size) have adopted the SAP HANA in-memory DBMS either to retrofit traditional workloads, such as SAP ERP and SAP Business Warehouse, or to power new applications such as SAP S/4HANA.
- Packaged application providers (such as Microsoft and Oracle) delivered IMC-enabled add-ons to their products.
- Traditional DBMS platforms from almost every vendor include IMC capabilities as standard features, or as options.

- Increasing availability of open-source IMC technologies that make adoption affordable for midsize organizations and startups.
- IMC technologies are increasingly embedded in mainstream software products and cloud services (for example, applications, business process management tools, application platforms and integration platforms).
- A growing number of providers offering cloud-based renditions of their IMC-enabling, or IMC-enabled, products.
- Convergence of several IMC technology streams into proto-IMC platforms.
- The emergence of a new generation of memory technologies that will combine the low data access latency of DRAM and the nonvolatility and affordable cost of flash memory.

The already notable use of IMC across vertical sectors, geographies and business sizes continues to expand. This is stimulated by the growing number of digital business initiatives, the mounting interest around real-time analytics and emerging use cases such as IoT, omnichannel (e-commerce or banking), event-driven decisions, hyperscale architectures and hybrid transactional/analytical processing (HTAP).

Factors that will be obstacles to even-faster adoption include:

- Market and technology fragmentation
- Cloud services enabling alternate IMC adoption at low cost and risk
- Security and IT operations challenges
- Disappointing results when just lifting and shifting traditional applications on top of IMC-enabling technologies
- Long development and maturity cycles for new IMC-enabled versions of packaged applications

User Advice: I&O and application leaders in charge of modernizing their architectures to support strategic initiatives must identify which of the following approaches is the best path for their IMC adoption:

- Developing new custom (or purchasing new packaged) groundbreaking applications, natively based on IMC design principles.
- Replatforming or reengineering traditional applications for IMC technologies. These approaches are less invasive than the IMC-native style, but usually lead to incremental benefits.

Application leaders should also:

- Monitor their strategic vendor roadmaps to identify how IMC impacts their investment plans.
- Make sure that the particular variants of IMC that are offered are suited to their needs.

IMC skills and best practices are rapidly becoming commonplace. Nonetheless, it is still advisable for mainstream organizations to incrementally become familiar with IMC architectures by successfully deploying a few IMC applications, based on the less disruptive approaches, before embarking on more ambitious, native IMC-style projects.

Business Impact: IMC has a long-term, disruptive impact by radically changing users' expectations, application design principles, product architectures and vendor strategies. IMC opens up a number of opportunities for digital innovation, such as hyperscale applications, HTAP, omnichannel customer experience, API economy, situation awareness, event-driven IT, and self-service and real-time analytics.

IMC-style applications drive transformational business benefits by enabling IT leaders to:

- Deliver orders of magnitude faster, analytical, transaction processing and HTAP applications.
- Support hyperscale business models serving very large numbers of globally distributed, mobile-enabled users or smart machines interacting in real time.
- Provide deeper and greater real-time business insights and situation awareness, via event stream processing.
- Radically reduce the time to value of new applications through near-interactive development processes.
- Improve price performance and resilience of complex systems.

Organizations leveraging IMC are better positioned to build defensible business differentiation than those sticking with traditional architectures. Organizations that fail to endorse IMC risk falling behind in the race for leadership in the digital era.

Benefit Rating: Transformational

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Sample Vendors: GigaSpaces; IBM; Intel; Microsoft; Oracle; Qlik; SAP; Software AG; Tableau; Workday

Recommended Reading: "Predicts 2019: In-Memory Computing at a Turning Point, Driven by Emerging Persistent-Memory Innovation"

"Forecast: In-Memory Computing Platforms (AIM Components), Worldwide, 2017-2022"

"SAP S/4HANA Research Roundup"

SD-WAN

Analysis By: Andrew Lerner

Definition: Software-defined wide-area network (SD-WAN) products replace traditional branch routers. They provide several features: dynamic path selection, based on business or application policy; centralized policy and management of WAN edge devices; and zero-touch configuration. SD-WAN products are WAN transport/carrier-agnostic, and can create secure paths across multiple WAN connections. SD-WAN products can be hardware- or software-based, and managed directly by enterprises or embedded in a managed service offering.

Position and Adoption Speed Justification: Rampant client interest in SD-WAN products continues, and we estimate that more than 25,000 customers have deployed SD-WAN products in production networks, which is over 600,000 branch locations. We expect continued rapid growth of SD-WAN deployments, and forecast vendor revenue to grow at a more than 23% compound annual growth rate (CAGR) for the next three years. In conjunction with a hybrid WAN topology, SD-WAN improves availability, cost and performance for enterprise WANs. Organizations moving to hybrid or internet-only WAN transport are driven toward SD-WAN products, because of their improved path selection functionality and manageability. Large numbers of vendors (several dozen) are competing in the market, including incumbent network and security vendors, startup vendors and smaller vendors with regional or vertical focus.

User Advice: Networking leaders should refresh their branch WAN equipment by implementing SD-WAN when they're migrating apps to the public cloud, building hybrid WANs, equipment is at end of life, or managed network service/MPLS contracts are up for renewal. Follow a comprehensive SD-WAN selection process by evaluating a diverse set of vendors and running a pilot. This is particularly important now, because not all offerings on the market are stable and scalable. Include network security teams in the design, planning and implementation, because SD-WAN-enabled hybrid WANs directly affect placement of security controls, such as firewalls and secure web gateways (SWGs).

Business Impact: The main purpose of emerging SD-WAN products is to create simpler and more cost-effective branch office WANs that map to modern application and cloud architectures. These products are significantly faster, easier to deploy and more manageable than traditional, router-based solutions. The benefits of an SD-WAN approach are substantial, compared with traditional, router-based WAN architectures, including reduced capital and operational expenditures (capex/opex) at the WAN edge, improved provisioning times, and the potential for enhanced branch availability.

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Early mainstream

Sample Vendors: Aryaka; Cato Networks; Cisco; Fortinet; Palo Alto Networks; Silver Peak; Versa Networks; VMware

Recommended Reading: "Technology Insight for SD-WAN"

"Magic Quadrant for WAN Edge Infrastructure"

"Solution Comparison for SD-WAN"

"Assessing the Strengths and Weaknesses of SD-WAN Technology"

Business Continuity Management Program Solutions

Analysis By: David Gregory; Roberta Witty

Definition: Business continuity management program (BCMP) solutions are the primary tools used to manage BCM programs and their artifacts for all phases of the life cycle, from planning to crisis activation. BCMP solutions provide capabilities for availability risk assessment, business impact analysis, business process and resource/asset dependency mapping, recovery plan management, exercise and crisis management, and BCM program management metrics and analysis.

Position and Adoption Speed Justification: Organizations need a consistent, repeatable process for all aspects of the BCM program development process. Also, the growing focus on BCM program metrics has increased BCMP solution sophistication by providing program analysis capabilities, which more mature BCM programs use to promote resilience in day-to-day business operations. The biggest competitors for BCMP solutions are "do-it-yourself" approaches leveraging Microsoft Office tools, with SharePoint as a document repository. Almost all BCMP solutions are available as SaaS solutions, which is extremely beneficial when there is a real disaster, and internal IT services (including potentially on-premises BCMP solution access) and systems are not available.

In the 2020 Hype Cycle, we retained BCMP solutions at its 2019 position. This is because there has been no significant change in the adoption rate over the last 12 months. There is, however, significant evidence that this may change over the coming months.

The COVID-19 pandemic has placed a direct spotlight on organizational resilience, and this is likely to be a significant factor in generating interest in the BCMP market during 2020 and 2021. This is likely to be particularly true in small and midsize businesses (SMBs), where BCM activities are often ad hoc, fragmented and practiced in departmental silos, which has led to a slow and disjointed response. Where this is the case, it is probable that organizations will be interested in BCMP solutions because they provide an easy way to jump-start a BCM program and deliver great benefit due to the standardization and centralization of BCM activities.

The BCMP solution market continues to enter the Plateau of Productivity level for large and extra-large organizations, regulated enterprises, government agencies, as well as for organizations with complex business operations.

We anticipate continued adoption during the next five years, given the increase in preparedness initiatives across all industries, especially due to COVID-19 and continued targeted cyberattacks.

User Advice: Consider a BCMP solution when you are:

- Starting a new BCM program and want to follow standard practices throughout the organization
- Updating or automating your current BCM program and processes, such as risk assessments and business impact analyses

- Maturing your BCM program, thus needing more program management analytics than traditional office management tools can provide
- Expanding your BCM program's automated footprint to gain more control over the execution phase of the BCM program life cycle
- Integrating plans from several departments, business units, divisions and legal entities into a consistent and easily updated set of recovery plans reflecting all elements of the organization
- Developing future pandemic plans and procedures

Do not overbuy; instead, focus on:

- Ease of use for business users, not IT or BCM program office users only
- Ease of configuration (not code customization) and reporting — by you, not the vendor — to your organization's continuity delivery framework, for example
- Ease of integration with other key business applications — such as enterprise directories, HR tools, business process management tools (whether internally developed or purchased), configuration management databases (CMDBs), IT asset management (ITAM) tools, BCM solutions that your organization may already have purchased (such as emergency/mass notification services [EMNS] or crisis/emergency management platform [C/EMP] solutions), and news feeds to a BCM program dashboard
- Mobile device (smartphone or tablet) support for recovery plan access and execution at the time of a business disruption

If your organization is already using an integrated risk management (IRM) solution, evaluate that product's BCMP capability to determine whether its functionality will support your BCM program.

Business Impact: BCMP solutions will benefit every organization wanting to perform a comprehensive analysis of its preparedness level to cope with business or IT interruptions. BCMP solutions also allow businesses to establish up-to-date, accessible plans that will help to facilitate robust response, recovery and restoration actions.

If used to its complete potential, a BCMP solution database becomes a "gold mine" that can be used to enhance business operations and resilience outside of recovery. It can be used in areas such as:

- HR management
- Business and IT reengineering (by finding duplication of services or rarely, if ever, used services)
- Mergers and acquisitions (by having a map of business services to other critical assets)

Benefit Rating: High

Market Penetration: More than 50% of target audience

Maturity: Early mainstream

Sample Vendors: Assurance Software; Avalution; Continuity Logic; Dell Technologies (RSA); Fusion Risk Management; Infinite Blue (BC in the Cloud); Premier Continuum; RecoveryPlanner; SAI Global (Strategic BCP); ServiceNow

Recommended Reading: "Market Guide for Emergency/Mass Notification Services"

"Market Guide for Crisis/Emergency Management Platforms"

"Magic Quadrant for Business Continuity Management Program Solutions, Worldwide"

"Critical Capabilities for Business Continuity Management Program Solutions, Worldwide"

Cloud Computing

Analysis By: David Smith

Definition: Cloud computing is a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service using internet technologies.

Position and Adoption Speed Justification: Cloud computing is a very visible and hyped technology, and has passed the Trough of Disillusionment. Cloud computing remains a major force in IT. Every IT vendor has a cloud strategy — although some strategies are better described as "cloud inspired." Users are unlikely to completely abandon on-premises models, but there is continued movement toward consuming more services from the cloud and enabling capabilities not easily accessible elsewhere. Much of the cloud focus is on agility, speed and other benefits beyond cost savings.

"Cloud computing" continues to be one of the most hyped terms in the history of IT. Its hype transcends the IT industry and has entered popular culture, which has had the effect of increasing hype and confusion around the term. In fact, cloud computing hype is literally "off the charts," as Gartner's Hype Cycle does not measure amplitude of hype (meaning that a heavily hyped term such as "cloud computing" rises no higher on the Hype Cycle than anything else).

Although the peak of hype has long since passed, cloud still has more hype than many other technologies that are at or near the Peak of Inflated Expectations. Variations, such as private cloud computing and hybrid approaches, compound the hype and reinforce the conclusion that one profile on a Hype Cycle cannot adequately represent all that is cloud computing. Some cloud variations (such as hybrid IT and now multicloud environments) are now at the center of where the cloud hype currently is. And, of course, there are different types of cloud services such as IaaS, PaaS and SaaS, each at various stages of industry hype.

New and advanced use cases for cloud introduce even more terms such as distributed cloud, multicloud and cloud-native. These add to the overall cloud hype as well as the applicability of cloud to more and more scenarios, including enabling next generation disruptions.

User Advice: User organizations must demand clarity from their vendors around cloud. Gartner's definitions and descriptions (which align with other useful ones such as NIST) of the attributes of cloud services can help with this. Users should look at specific usage scenarios and workloads, map their view of the cloud to that of potential providers, and focus more on specifics than on general cloud ideas. Understanding the service models involved is key — especially the need to understand the shared responsibility model for security.

Vendor organizations should focus their cloud strategies on more specific scenarios and unify them into high-level messages that encompass the breadth of their offerings. Differentiation in hybrid cloud strategies must be articulated. This will be challenging, as all are “talking the talk,” but many are taking advantage of the even broader leeway afforded by the term. “Cloudwashing” should be minimized. Gartner’s Cloud Spectrum can be helpful.

Adopting cloud for the wrong reasons can lead to disastrous results. There are many myths surrounding cloud computing as a result of the hype (see “Revisiting the Top 10 Cloud Myths for 2020” for details and advice).

Business Impact: The cloud computing model is changing the way the IT industry looks at user and vendor relationships. Vendors must become providers, or partner with service providers, to deliver technologies indirectly to users. User organizations will watch portfolios of owned technologies decline as their service portfolios grow.

Potential benefits of cloud include cost savings and capabilities related to the flexible and dynamic usage models of cloud (including concepts that go by names such as “agility,” “time to market” and “innovation”). Organizations should formulate cloud strategies that align business needs with those potential benefits. Agility is the driving factor for organizations embracing cloud most of the time.

Benefit Rating: Transformational

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Sample Vendors: Amazon; Google; IBM; Microsoft; Oracle; Red Hat; Salesforce; SAP

Recommended Reading: “Cloud Computing Primer for 2020”

“The Cloud Strategy Cookbook, 2019”

“Revisiting the Top 10 Cloud Myths for 2020”

“Four Types of Cloud Computing Define a Spectrum of Cloud Value”

Cloud Management Platforms

Analysis By: Dennis Smith

Definition: Cloud management platforms (CMPs) enable organizations to manage private, public and multicloud services and resources. Their specific functionality is a combination of provisioning and orchestration; service request management; inventory and classification; monitoring and analytics; cost management and resource optimization; cloud migration, backup and disaster recover; and identity, security and compliance. This functionality can be provided by a single product or a set of vendor offerings with some degree of integration.

Position and Adoption Speed Justification: While the CMP market is continually changing, vendors and enterprise customers are getting a better feel about where such tooling can and cannot be used. Vendors are still being challenged with evolving customer requirements (for example, interfacing with multiple public clouds, cost transparency with workload optimization to remediate cost overruns and handling newer functions like containers and serverless deployments). At the same time, major market consolidation will continue. For example, many vendors, that initially targeted cost management, have been acquired as this functionality is becoming a part of the basic CMP. Additionally many vendors in adjacent markets are acquiring CMP vendors and combining this functionality with asset management (software and hardware) and SaaS operational management. Cloud service providers (CSPs) and management service providers (MSPs) are also entering the market. Additionally, many long-standing vendors are introducing next-generation products, often targeting holes that their previous products had. Finally vendors in different markets (e.g., monitoring) are also entering the market. Some of the core CMP functionality is also being combined (for example, monitoring and analytics with cost management and resource optimization). The ability to serve both application developer and I&O personas is the key. This requires that CMPs be linked into the application development process without imposing a workflow that inhibits agility while also allowing infrastructure and operations (I&O) teams to enforce provisioning standards.

Organizations have an increasing need to address multicloud requirements. In some cases, they want to become internal cloud service brokers (CSBs) and manage public services that were previously acquired — often by lines of business (LOBs) outside the I&O organization — and have become difficult to manage operationally.

User Advice: As CMP market volatility increases, IT organizations must:

- Consider CMP vendor's viability along with evaluating features.
- First consider native cloud services as an alternative or option versus CMPs, particularly if you favor depth with an individual cloud provider versus breadth across different cloud providers.
- Consider functionally focused tools (e.g., cloud expense management tool) if you only require a limited set of functionalities.
- Augment, swap out or integrate additional cloud management or traditional management tools for many requirements, because no vendor provides a complete cloud management solution.
- Standardize, because deriving value from your CMP will depend heavily on the degree of standardization offered by the infrastructure, software and services.
- Set realistic expectations on deployment times, as mature organizations implement CMP in a relatively short period (one to two years); however, less mature organizations may require two or more years to design effective, repeatable, and automatable standards and processes.
- Plan for new roles, such as cloud architects and cloud service brokers (CSBs), including developing skills in the financial management and capacity management areas.

Business Impact: Enterprises will deploy CMPs (increasingly as a part of a larger product suite) to increase agility, reduce the cost of providing services and increase the likelihood of meeting service levels. Costs are reduced and service levels are met because CMP deployments require adherence to standards, as well as increased governance and accountability. Desirable IT outcomes include:

- Policy enforcement (e.g., on reusable standard infrastructure components).
- Reduced lock-in to public cloud providers, although at the cost of CMP vendor lock-in that can slow innovation.
- Enhanced ability to broker services from various cloud providers and to make informed business decisions on which providers to use.
- Ongoing optimization of SLAs and costs.
- Management of SLAs and enforcement of compliance requirements.
- Health and performance monitoring of cloud applications.
- Accelerated development, enabling setup/teardown of infrastructure that mimics production, resulting in lower overall infrastructure costs and higher quality. This can be in support of DevOps initiatives.

Benefit Rating: Low

Market Penetration: 5% to 20% of target audience

Maturity: Mature mainstream

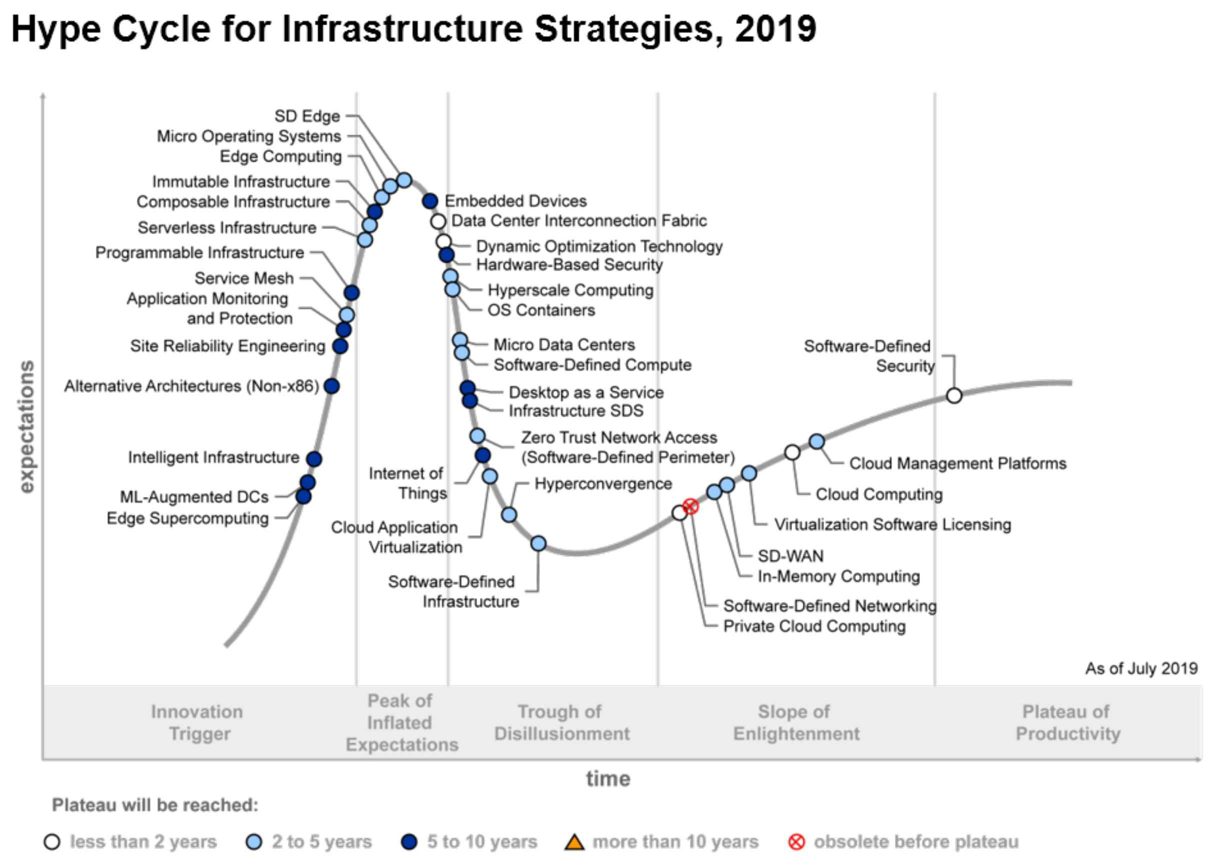
Sample Vendors: CloudBolt; Flexera; CloudSphere; Morpheus Data; Scalr; Snow Software; VMware

Recommended Reading: "Magic Quadrant for Cloud Management Platforms"

"Critical Capabilities for Cloud Management Platforms"

Appendixes

Figure 3. Hype Cycle for Infrastructure Strategies, 2019



Source: Gartner
ID: 388022

Hype Cycle Phases, Benefit Ratings and Maturity Levels

Table 1. Hype Cycle Phases

Phase	Definition
Innovation Trigger	A breakthrough, public demonstration, product launch or other event generates significant press and industry interest.
Peak of Inflated Expectations	During this phase of overenthusiasm and unrealistic projections, a flurry of well-publicized activity by technology leaders results in some successes, but more failures, as the technology is pushed to its limits. The only enterprises making money are conference organizers and magazine publishers.
Trough of Disillusionment	Because the technology does not live up to its overinflated expectations, it rapidly becomes unfashionable. Media interest wanes, except for a few cautionary tales.
Slope of Enlightenment	Focused experimentation and solid hard work by an increasingly diverse range of organizations lead to a true understanding of the technology's applicability, risks and benefits. Commercial off-the-shelf methodologies and tools ease the development process.
Plateau of Productivity	The real-world benefits of the technology are demonstrated and accepted. Tools and methodologies are increasingly stable as they enter their second and third generations. Growing numbers of organizations feel comfortable with the reduced level of risk; the rapid growth phase of adoption begins. Approximately 20% of the technology's target audience has adopted or is adopting the technology as it enters this phase.
Years to Mainstream Adoption	The time required for the technology to reach the Plateau of Productivity.

Source: Gartner (July 2020)

Table 2. Benefit Ratings

Benefit Rating	Definition
Transformational	Enables new ways of doing business across industries that will result in major shifts in industry dynamics
High	Enables new ways of performing horizontal or vertical processes that will result in significantly increased revenue or cost savings for an enterprise
Moderate	Provides incremental improvements to established processes that will result in increased revenue or cost savings for an enterprise
Low	Slightly improves processes (for example, improved user experience) that will be difficult to translate into increased revenue or cost savings

Source: Gartner (July 2020)

Table 3. Maturity Levels

Maturity Level	Status	Products/Vendors
Embryonic	▪ In labs	▪ None
Emerging	▪ Commercialization by vendors ▪ Pilots and deployments by industry leaders	▪ First generation ▪ High price ▪ Much customization
Adolescent	▪ Maturing technology capabilities and process understanding ▪ Uptake beyond early adopters	▪ Second generation ▪ Less customization
Early mainstream	▪ Proven technology ▪ Vendors, technology and adoption rapidly evolving	▪ Third generation ▪ More out-of-box methodologies
Mature mainstream	▪ Robust technology ▪ Not much evolution in vendors or technology	▪ Several dominant vendors
Legacy	▪ Not appropriate for new developments ▪ Cost of migration constrains replacement	▪ Maintenance revenue focus
Obsolete	▪ Rarely used	▪ Used/resale market only

Source: Gartner (July 2020)

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

Understanding Gartner's Hype Cycles

GARTNER HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters
AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit http://www.gartner.com/technology/about.jsp