

Secrets From Cloud Computing's First Stage: An Action Agenda for Government and Industry

BILL WHYMAN | JUNE 2021

Cloud computing drives innovation and productivity across the economy, just as the electric grid did a century ago—yet it is more capable and dynamic, and it still in its early stages. Cloud is important not just at the firm level, but also for economic growth and global competitiveness.

KEY TAKEAWAYS

- Cloud computing is a powerful and disruptive technical architecture that is now breaking out into the mainstream economy. Millions of organizations use it, but adoption is more broad than deep, accounting for just 7.2 percent of global IT spending.
- Cloud enables better operating and business models by lowering costs and increasing business agility, innovation, and productivity. By taking over the hard work of operating IT, it allows users to focus on their core strengths and customer needs.
- Cloud is a platform, providing a common architecture spanning over 200 IT services, including machine learning. Major clouds are ecosystems: Thousands of companies build products on them, which creates more value than any of them could alone.
- Policymakers should accelerate cloud adoption and improve citizen services with a federal cloud modernization moonshot.
- Other policy measures should include enhancing security, strengthening cross-border data flows, building workforce skills and inclusivity, and preserving competition.

CONTENTS

- I. Introduction2**
- II. What Is Cloud Computing?2**
- III. Why Cloud Computing Matters to Economic Success5**
- IV. Cloud Economics, Industry Structure, And Market Dynamics 7**
 - Cloud Adoption Is in Early Stages—Adoption Is More Broad Than Deep..... 7
 - There Are Multiple Market Segments With Different Customer Needs 8
 - Cloud Providers Have Large Economies of Scale and High Fixed Costs 8
 - Customers Benefit From Declining Prices and Improved Price Transparency 9
 - The Cloud Is a Platform With a Partner Innovation Ecosystem 12
 - Products Are Differentiated “Up the Stack” 13
 - Integration Across the Industry Value Chain Is Low, But Increasing..... 14
 - Profitability Is Similar to Large Software Companies 15
 - The Cloud Is a Good Fit for Most, But Not All, Applications..... 15
- V. Where Cloud Computing Excels 16**
 - Cloud Drives a Better Business and Technology Model..... 16
 - Cloud Is a Rich and Capable Platform, Not Just a Collection of Services..... 17
 - Cloud Lowers Switching Costs Despite Lock-In Concerns, Multi-cloud Best Practices..... 19
 - Cloud Boosts Agility, Speed, and Innovation 21
 - Cloud Infrastructure-as-Code Enables Automation, Productivity, and Specialization 22
- VI. Challenges for Cloud Success23**
 - Achieving Stronger Security and Compliance: Automation and Management 23
 - Enhancing Governance, Managing IT and Data, and Reducing Complexity 25
 - Building New Skills and Inclusivity 26
 - Making Cloud Adoption Easier..... 26
 - Adapting to Hybrid Cloud, the Emerging Edge, and 5G 28
- VII. A Policy Agenda for Cloud Computing29**
 - 1. Accelerating Adoption 29
 - 2. Improving Security and Resilience: Government-Industry Collaboration Program..... 32
 - 3. Preserving Competition 34
 - 4. Strengthening Data Governance: Cross-Border Data Confidence-Building Measures 35
 - 5. Promoting Skills and Inclusivity via a Public-Private Training Partnership 36
- Endnotes.....38**

I. INTRODUCTION

Cloud computing is a powerful and disruptive technology that is now breaking out into the mainstream economy, more than a decade after launching in 2006. Millions of companies are using some form of cloud computing, often from Amazon Web Services (AWS), Microsoft Azure, Google Cloud, or other providers. Successfully adopting cloud computing will be a key determinant of which countries will prosper in the global economy. Cloud computing lowers costs, creates technical and business agility, and enables innovation and digital transformation. In 2020, the global market for cloud services was \$270 billion and cloud companies listed on U.S. markets had more than \$1 trillion in capitalization.¹ Hundreds of new cloud services are being introduced and thousands of new venture-backed cloud companies are being formed.²

Yet, cloud computing is still in its early days, given more than \$3 trillion in annual global information technology (IT) spending and decades of accumulated investment in traditional IT infrastructure. In short, cloud adoption is more broad than deep. Most companies use the cloud for only a small share of their IT needs, and spending on cloud computing is just 7.2 percent of annual global IT spending.³ It's early in this disruptive shift, and it's important to better understand cloud computing's successes and challenges. Cloud computing is important not only for individual companies' success, but also for economic growth and global competitiveness. It's what economists call a "general-purpose technology" that is pervasively used across most sectors. The cloud is becoming a platform that drives innovation and productivity across the broader economy. It represents the digital equivalent of the electric grid, only one that is more capable and dynamic.

Cloud computing is important not only for individual companies' success, but also for economic growth and global competitiveness.

After a brief introduction to what cloud computing is, this report focuses on the economic impact of cloud. It addresses why cloud computing matters for a country's economic success—which has not received as much focus as the cloud's technical and business impact. It then analyzes the economics of the cloud sector, including industry structure and market dynamics. The report recommends five areas where cloud computing excels and five challenges that need to be met to fully capture the benefits of the cloud. It concludes with a policy agenda for cloud computing. Governments will inevitably address cloud-related policy issues as cloud's reach and power expand. Recommendations for policymakers include accelerating adoption with a federal cloud-modernization moon shot; improving security and resilience; preserving competition by encouraging application portability; strengthening governance and enabling cross-border data flows; and building workforce skills and inclusivity.

II. WHAT IS CLOUD COMPUTING?

Cloud computing is a powerful technical architecture for IT, driven by technologies such as software virtualization. Just as important, cloud architecture enables new operating and business models. As a starting point, cloud computing provides IT resources (e.g., compute, storage) as a pay-only-for-what-you-use service delivered over the communications network by a third party. This is in contrast to the predominant model in which users directly "own and operate" on-premises physical IT equipment. The U.S. National Institute of Standards and Technology (NIST) states

“cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources ... that can be rapidly provisioned and released with minimal management effort or service provider interaction.”⁴ An important capability of the cloud is that IT resources can be scaled up when demand increases, or scaled down and turned off, delivering agility gains and cost savings. We break down the cloud’s defining characteristics in figure 1.

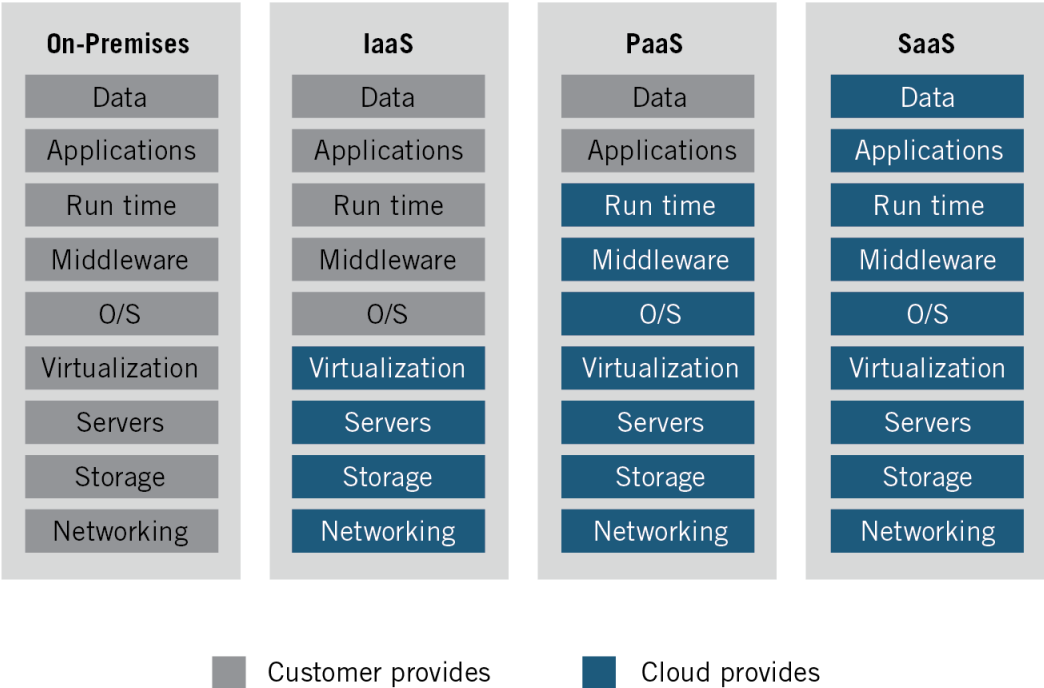
Figure 1: What is cloud computing?⁵

Defining Characteristic	Traditional On-Premises IT (or What the Cloud Is Not)	Cloud Computing
Pricing and cost structure	Pay up-front licensing or purchase. Fixed costs	Pay as you go, only for what you use. Variable costs
Asset ownership and maintenance	Customer owns and maintains IT assets	Customer “rents” cloud provider’s infrastructure. Customer still owns their data and applications
Service model	Customer provides their own IT service	Cloud provides the service to the customer
Capacity scaling and provisioning	Fixed capacity. Typically controlled by customer’s IT department	Variable. Elastic on-demand scale up and down, in near real time. Self-service provisioning by the customer, within limits
Geographic location	On the customers’ premises or leased data center	Cloud provider data center, accessed via the network. Location specific or agnostic
Dedicated or shared infrastructure	Typically dedicated only to the owner or purchaser	Shared with other customers. Software partitions create isolated resources

Three Layers of the Cloud: IaaS, PaaS, SaaS. Cloud computing is typically broken into three major parts, each covering a collection of services. Infrastructure as a service (IaaS) provides the underlying IT infrastructure such as compute, storage, and networking. Platform as a service (PaaS) provides middleware, databases, and developer and management tools to build and support applications. Software as a service (SaaS) provides a fully managed application that customers directly use such as email, financial applications, or supply chain management. While all three layers operate together, each represents a different scope of adoption of the cloud by the user (see figure 2). Each involves a different technical part (often called a “layer of the stack”) of the technology solution. As users move from IaaS to PaaS to SaaS, they shift more responsibility for IT to the cloud provider. This typically provides lower costs and more agility, but entails less control over individual IT resources. Users can and do adopt different services independently, as each user takes their own journey of cloud adoption.⁶ This reflects different customer priorities and skill levels. There is no one right starting point for adopting the cloud.

For example, at the IaaS level, customers store data, run web servers, and provide back-up and recovery services in the cloud. At the PaaS level, customers often run databases in the cloud for high scalability and availability, while also shifting database maintenance to the cloud. At the SaaS level, customers run applications such as email and collaboration, human resources management, and accounting systems in the cloud. In practice, the lines between these approaches and services blur, and cloud providers' offerings overlap across markets. There are also new so-called "XaaS" markets emerging such as Function as a Service (FaaS). These XaaS markets are important, but mostly smaller markets in earlier stages of adoption.

Figure 2: Different approaches to the cloud have different scopes of adoption⁷



Another distinction is between public and private clouds, which use similar technology.⁸ The major difference is public clouds offer services to the general public and the same cloud infrastructure is shared by many different entities, whereas private clouds are typically limited to a single organization. Private clouds are usually smaller in scale and can trail the cost and performance of public clouds. "Hybrid cloud" refers to combining traditional on-premises IT with public clouds. We discuss the benefits and limits of hybrid cloud.

This paper focuses on public cloud IaaS and PaaS. The primary providers of IaaS and PaaS are the same. IaaS and PaaS also share similar characteristics: Both are underlying technical resources that support the application, and both are predominantly used by IT managers and software developers rather than end-users. The IaaS and PaaS markets are moving closer together as providers offer services that combine their functionality.⁹ In contrast, SaaS is led by a different set of software providers (e.g., Salesforce, Workday, ServiceNow), and end users interact directly with the application. While SaaS takes advantage of and runs on top of IaaS and PaaS infrastructure, SaaS's core value added is the application itself.¹⁰

This paper also focuses on enterprise cloud markets that serve businesses, governments, and organizations—rather than individual consumers. While companies such as Facebook and Twitter use similar cloud technologies and run, in part, on top of enterprise cloud providers, they are consumer-focused companies. For the most part, they do not offer their services “wholesale” for other businesses to build on, but rather effectively offer SaaS applications to consumers. The consumer cloud companies also have different business models, often driven by advertising spending. This is not to say only enterprise providers define the cloud; there are a wide array of services offered by many others that compose the cloud. For the purposes of this paper, enterprise and consumer raise different issues.

III. WHY CLOUD COMPUTING MATTERS TO ECONOMIC SUCCESS

IT is the foundation of a modern economy. IT underpins the operations and workflows across every facet of an organization. Companies in every industry use IT to develop products, orchestrate supply chains, drive the selling process, run operations, and manage every step of the business. Every government uses IT to provide citizen services, from health care to education, and to manage back-office operations. In fact, 28 percent of U.S. business investment is in IT, and the U.S. federal government spends over \$100 billion annually on IT.¹¹

Cloud computing is an emerging technology architecture that is driving growth, productivity, and innovation. New computing architectures come in waves.¹² The mainframe dominated in the 1960s and 1970s, the PC in the 1980s, client-server in the 1990s, the Internet in the 2000s, and cloud computing took off in 2010 and is now gaining scale. Previously, new technology waves drove productivity gains and economic growth.¹³ Each new architecture builds on the prior one and delivers lower cost and more capability. Each wave deepens IT penetration of the economy, reaching more users and integrating IT with more economic activity. While cloud computing technology is new, it builds on a past technical lineage that includes network computing, grids, and distributed computing. Further, cloud computing is integral to new IT-driven business developments that have broad economic impact. For example, technologies such as artificial intelligence and machine learning (AI/ML) and Internet of Things (IoT) depend on robust computing and are being built directly into the cloud infrastructure and offered as cloud services. The cloud will become the platform for AI/ML.

Cloud computing is a better economic model for IT. It’s the digital equivalent of the electric grid—and is more powerful. Cloud computing finally delivers on the promise of IT by shifting the hard work of operating IT infrastructure to the cloud—with lower costs and better performance. It represents the industrialization of IT wherein cloud effectively standardizes and outsources the IT infrastructure to a cloud service provider. Cloud computing applies modern statistical process controls to IT similar to modern manufacturing. Customers no longer need to buy, deploy, operate, maintain, and own their IT infrastructure. The public cloud operates IT better, so customers don’t have to.

First, cloud computing has lower costs. An International Data Corporation (IDC) study sponsored by Amazon Web Services (AWS) shows cloud infrastructure has 31 percent lower operational costs than comparable on-premises infrastructure, and has even greater savings when people and downtime costs are included.¹⁴ Greater scale and technical innovation lead to rapid quality-adjusted price declines, according to economists Byrne, Corrado, and Sichel, of 7 percent *per year* for computing power, 12 percent *per year* for database, and 17 percent *per year* for

storage.¹⁵ The cloud is also more automated, enabling new capabilities and whole IT infrastructures to be deployed and managed with only a few key strokes. IDC estimated, in an AWS sponsored report, that IT staff efficiency is 62 percent greater in the cloud, while developer productivity is 25 percent greater.¹⁶ Second, cloud has better capital efficiency. It drives greater capacity utilization by aggregating demand across many customers and sectors, which each have their own usage profile. On-premises servers are typically used 20 percent of the time, whereas in the cloud, utilization typically well exceeds 50 percent. This reduces server over-provisioning and drives better asset efficiency. Third, the cloud has better energy efficiency and a lower carbon footprint. Google says its data centers use about six times less overhead energy for every unit of IT equipment than the average data center does.¹⁷ This progress is increasing.¹⁸

On an operating level, even large, sophisticated organizations have challenges implementing and upgrading IT.¹⁹ IT is not most companies' or governments' core competency, or even purpose. Each customer has to implement and operate their own IT system. For large mission-critical systems such as finance, customer records, and supply chains, implementation can take years. For example, HP implemented an application upgrade in 34 systems around the world, but the 35th failed, causing the company to miss its quarterly financial goals and its stock price collapsed.²⁰ In another example, the U.S. Department of Veterans Affairs initially spent \$1.1 billion over five years on an electronic health record system that it was not able to implement.²¹ In large part, this is because each enterprise's IT environment is different, complicated, and brittle. Over \$1 trillion globally is spent on people-based IT services each year to make these complex systems work. Even when successfully deployed and operated, traditional on-premises IT rapidly becomes out of date, causing a build-up of "technical debt" that needs to be modernized and replaced. Users then have to re-implement IT all over again. And IT implementation and operations are the easy part, relative to the challenge of getting real business value or government-mission value from IT. As discussed in more detail below, cloud computing is simply a better way for most of IT.

Cloud computing drives specialization and the division of labor, enabling the business, operations, services, and everything that uses IT to become much better. The cloud is not only a cheaper and better way of operating IT, it enables businesses and governments to focus on their customers, their core business, and the value added of satisfying needs. One of the most powerful ideas in economics is specialization and division of labor.²² The cloud takes specialization and division of labor to a new level in IT, wherein the cloud provider specializes in the IT infrastructure so users doesn't have to. The JPMorgan Chase bank has stronger IT capabilities than most, but it requires an army of 40,000 people in its IT organization with a \$9 billion annual budget.²³ As an apt historical analogy, in the early industrial revolution, manufacturing mills were built on rivers to provide water power.²⁴ Each manufacturer had to generate, distribute, maintain, and own their power, similar to how most enterprises today own and operate their IT. However, after electric-power technology became good enough, manufacturing mills began outsourcing electricity to the power grid, so they could focus on running their core manufacturing processes. They specialized in what they were good at boosting economic productivity.²⁵

Cloud computing drives innovation across the economy. Because cloud is a variable cost that can be turned on and off when needed, it lowers the cost of experimentation and failure. This reduces risk and enables faster time to market for new products. The entire economy benefits, but this is especially important for smaller companies that may not be able to afford larger up-front capital

purchases. The cloud also drops barriers to company formation, a key source of innovation and job growth.²⁶ Indeed, start-ups were some of the first users of the cloud. Most of the new Internet-based companies are cloud-native and, because they are starting fresh, their entire businesses are built on top of cloud IT infrastructure.

The larger significance of cloud computing is it's becoming the IT platform on which the economy is built. The distributed electric grid enabled the creation of hundreds of new products such as consumer appliances, and whole new industries such as radio and television broadcasting. As we will discuss, companies today (e.g., Twitter, Snap, Netflix) are building their whole business on the cloud, and large enterprises (e.g., Bosch, Providence Health, Procter & Gamble) are migrating existing IT to the cloud.²⁷ We'll show how cloud lowers cost, boosts productivity, and shifts scarce talent to higher value-added applications and business processes. It enables enterprises to move faster and be more agile. Cloud computing is spreading across every industry, providing flexible enabling technologies from basic computing to ML. The cloud enables businesses to reinvent each step of the business, including marketing, product development, manufacturing, selling, and financial management. Cloud computing is more than just the latest IT innovation. It underpins the broader economy directly enabling the “digital transformation” of the economy. *The digital economy is powered by the cloud.*

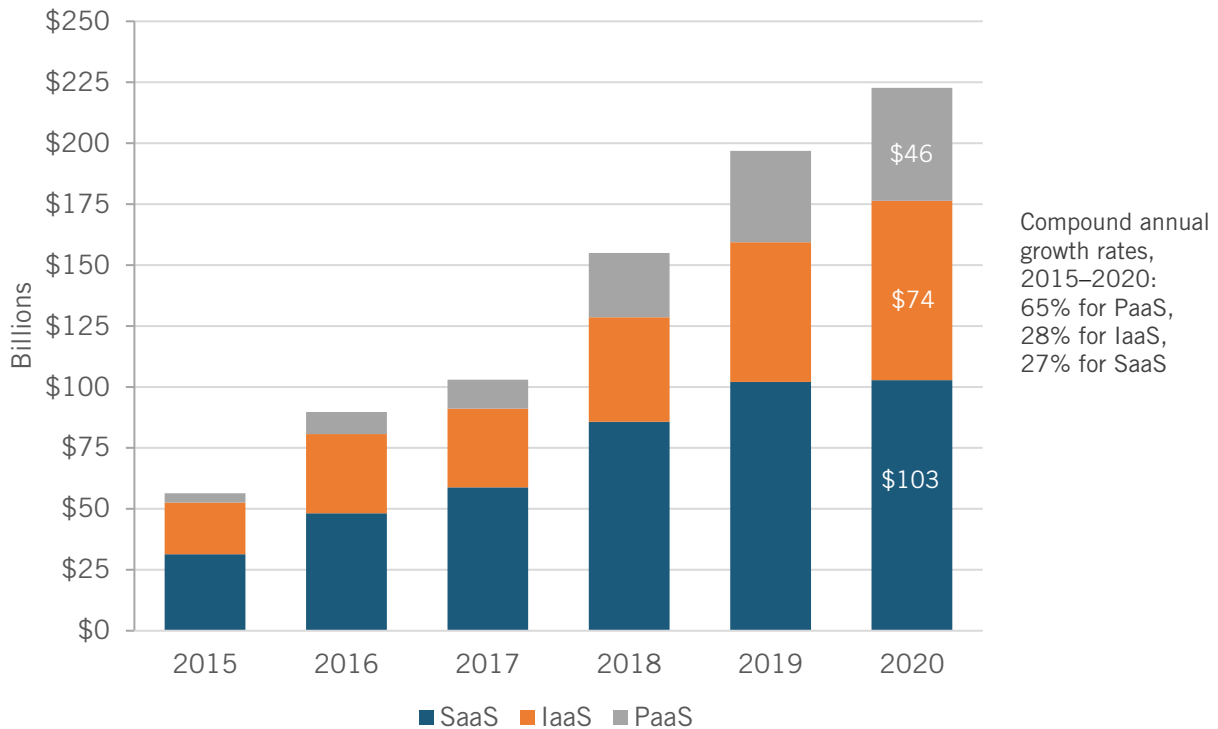
IV. CLOUD ECONOMICS, INDUSTRY STRUCTURE, AND MARKET DYNAMICS

Cloud Adoption Is in Early Stages—Adoption Is More Broad Than Deep

The market for cloud services is developing quickly but is still in the early stages. IaaS officially started in 2006, and in the last 10 years, adoption has accelerated beyond early adopters and is now breaking into mainstream users. Millions of organizations are using the cloud for specific applications and workloads. By number count, cloud computing is crossing over into the mainstream. However, measured by share of total use, market penetration is low, especially for new or advanced services. AWS CEO Andy Jassy recently pegged overall cloud adoption at 4 percent of global IT spending, and Gartner data shows 7.2 percent.²⁸ The latest National Science Foundation (NSF) business survey data shows similar low penetration.²⁹

SaaS (\$103 billion globally in 2020) and IaaS (\$74 billion) are the two largest markets by revenue, with PaaS the smallest (\$46 billion).³⁰ The SaaS market scaled first, building on the prior Application Service Provider model. SaaS adoption is much higher than PaaS or IaaS, exceeding 25 percent of the software application market, led by cloud-native companies such as Salesforce.com, Workday, and ServiceNow.³¹ Most traditional software companies such as Microsoft, Oracle, SAP, and Adobe are moving from traditional licensed, on-premises software to the SaaS model.³² All three markets are growing significantly faster (32 percent combined global growth rate) than is global IT spending (approximately 4 percent growth). As customers become more skilled at using the cloud, adoption grows, in turn increasing the spending, scale, and number of services used. Most customers use foundational services such as compute and storage. However, more advanced services such as ML, Internet of Things, and specialized databases still have low adoption.³³

Figure 3: Spending on global cloud computing, 2015–2020³⁴



There Are Multiple Market Segments With Different Customer Needs

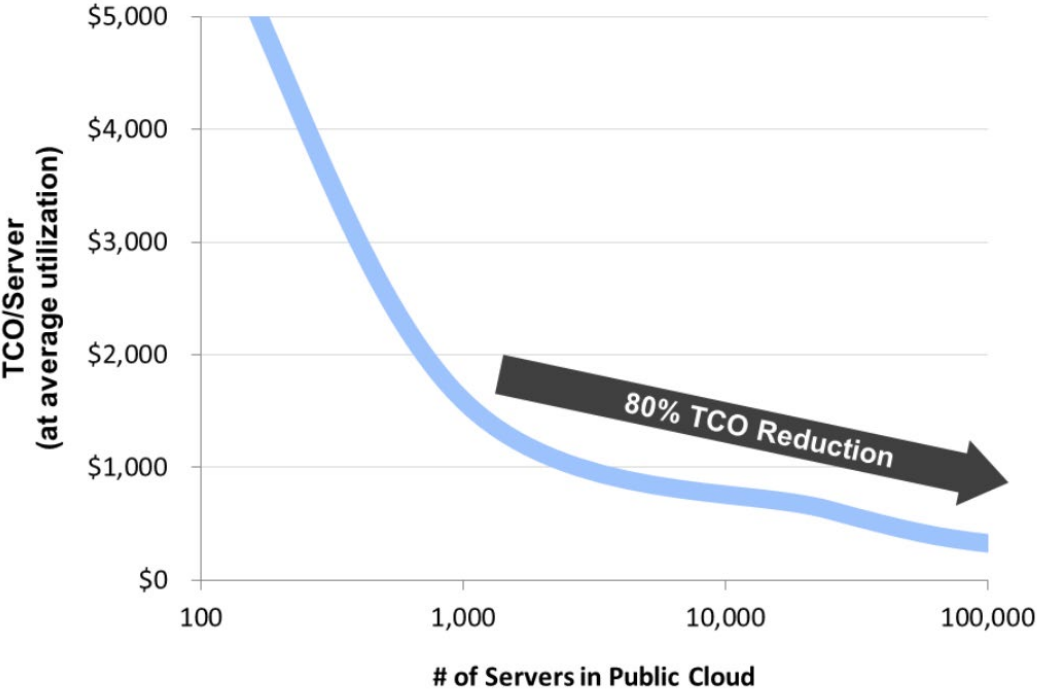
Different customers are at different stages of adopting the cloud and have reached different levels of cloud maturity. The key buying factors and customer needs vary across customer maturity and market segments. Some of the earliest cloud adopters were start-ups that valued the speed of near instant provisioning of IT, cash-conserving pay-as-you-go pricing, and access to world-class technologies. For some users, core capabilities such as lower cost, business continuity, and remote access matter most. For others, the ability to outsource IT at global scale and achieve granular management and security is key. While for advanced users, cloud-native architectures and software development practices enable them to rapidly test products, innovate, and achieve superior business agility. Ultimately, cloud computing is integral to organizational change and culture, becoming a source of competitive advantage. This is most identifiable in start-ups that built cloud-native architectures from the start (Airbnb, SmugMug, Pinterest). Yet, established firms are also going all in on the cloud in sectors ranging from banking, where Capital One has closed eight data centers, to education content providers such as Blackboard, Ellucian, and Instructure.

Cloud Providers Have Large Economies of Scale and High Fixed Costs

Large economies of scale are a defining characteristic of the cloud computing businesses. IDC tracks \$74 billion of investment in global cloud infrastructure in 2020, and the leading cloud providers each spend billions annually.³⁵ Microsoft and Amazon each have dozens of geographic regions and more than 150 data centers globally.³⁶ They don't disclose cloud-only CapEx data but it takes tens of billions of dollars to build cloud data centers globally, provision gigabit networks, and develop the software infrastructure. Opening a new region is typically a >\$1 billion investment.³⁷ The competition for scarce technical talent is also fierce.

Once providers incur this big fixed cost, the average cost for each unit of compute and storage declines as volume and utilization grows (figure 4). Cloud infrastructure is heavily automated, and providers have developed proprietary knowledge about design and management. Providers scale to millions of computing cores and petabytes of data without adding people, blowing past traditional industry benchmarks for the number of system administrators per server or database. This ultimately results in lower prices for users. Cloud providers also have greater economies of scale in purchasing, becoming the largest buyers of storage, memory, and compute. For example, the cloud is one-third of the global server market, up from just single digits a few years ago.³⁸ In addition, there are large soft costs such as compliance, which take years to achieve for the dozens of security and compliance regimes needed. All of these create substantial entry barriers for IaaS platform providers. However, lower-cost, large-scale IaaS providers reduce entry barriers for PaaS and SaaS, which can more quickly and cheaply leverage the massive cloud IaaS infrastructure. PaaS and SaaS can run their products on top of IaaS, choosing to buy IaaS rather than build it.

Figure 4: Cloud computing economies of scale³⁹



Customers Benefit From Declining Prices and Improved Price Transparency

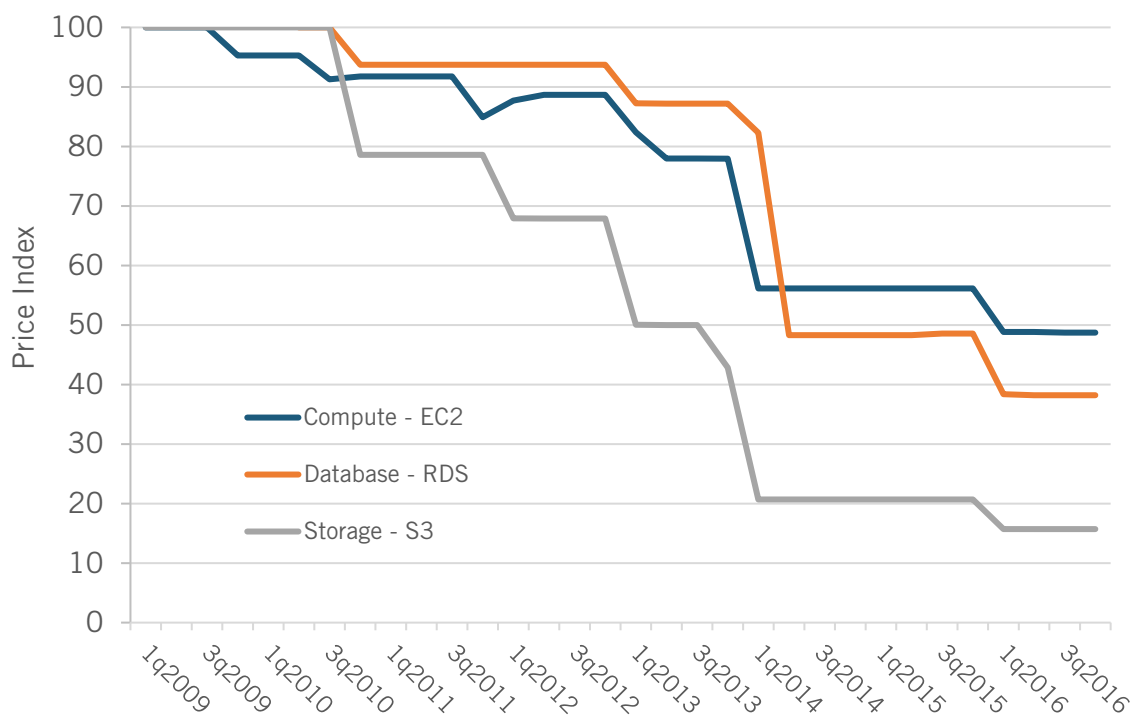
Compute, storage, and database are the largest segments of the market and show large price declines. AWS alone has cut prices over 70 times across services and regions since launching. Economists Byrne, Corrado, and Sichel analyzed published AWS prices from 2009 to 2016 and found that quality-adjusted prices fell rapidly (see figure 5).⁴⁰

- Compute prices (EC2) fell on average by 6.9 percent a year from 2009 to 2016 and, more rapidly, by 10.5 percent a year on average from 2014-2016 once Microsoft and Google started publishing prices publicly in 2014.

- Storage (S3) pricing fell on average by 17.3 percent a year from 2009 to 2016 and on average 25.1 percent a year from 2014 to 2016.
- Database (RDS) pricing fell on average by 11.6 percent a year from 2010 to 2016 and on average -22.6 percent a year from 2014-2016.

Falling prices are driven by 1) cost declines from economies of scale and technical innovation, 2) competition to win over customers from traditional on-premises providers that capture >90 percent of IT spending, and 3) competitive rivalry among the major cloud providers. For example, in a six-month period in 2012–2013 when cloud adoption started to pick up, AWS, Microsoft, and Google made over 20 price cuts.⁴¹

Figure 5: Declining prices for cloud computing⁴²



Cloud service pricing has become more complicated as the number of services and pricing models have grown. But this comes with greater choice and visibility. Prices for specific services are publicly published on cloud provider websites. Pricing differs by type of service (e.g., compute) and within a service (e.g., size and mix of computing power, memory, and networking optimized for different workloads). Cloud providers also offer a growing number of pricing models in addition to the standard no-up-front-charges, pay-only-for-what-you-use model. Customers achieve even greater discounts for committing to larger volumes and longer time periods, and pre-paying. There is even a “spot market” for computing power and a secondary market with cloud brokers and resellers. As users get more confident and capable managing their cloud spending, they are able to shift their mix of usage to even lower cost “reserved” capacity, thereby reducing their overall costs even further than standard price discounts. These pricing options are more transparent than most traditional software licensing. However, with more pricing choices,

customers must know both their usage patterns and how to use cost management tools to get the most benefit.

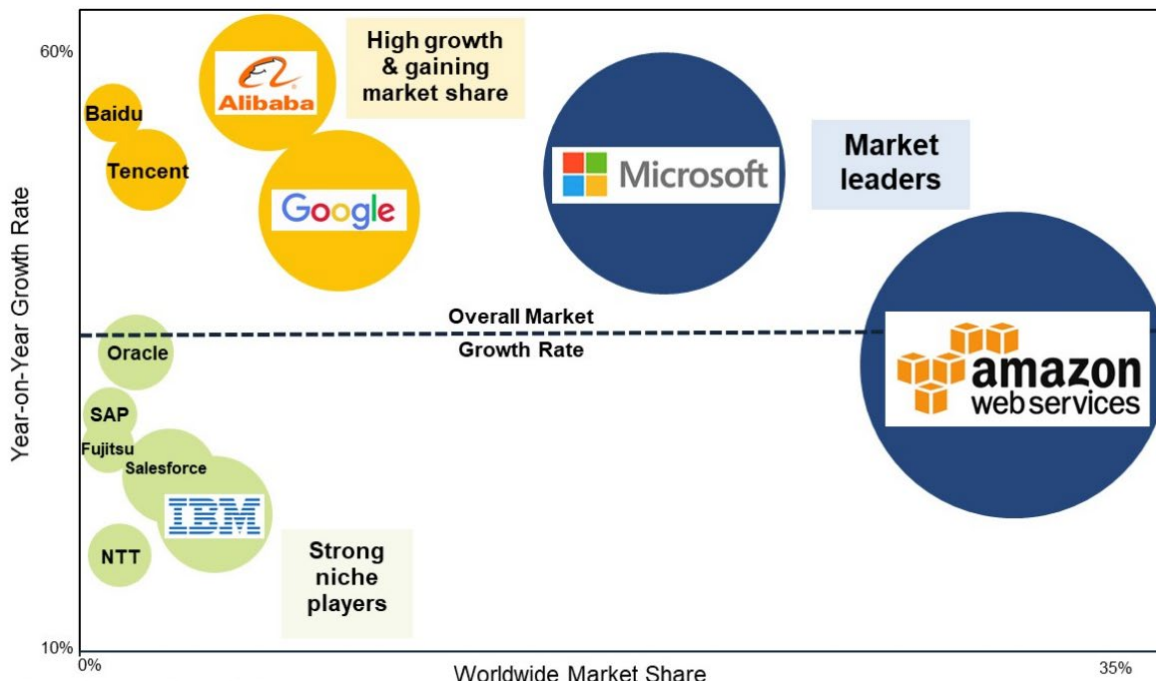
The Top Three Providers Have 60 Percent Global Market Share

Synergy Research data measuring the combined IaaS, PaaS, and hosted private cloud revenues for Q4 2020 shows that Amazon has 32 percent share, Microsoft has 20 percent, Google has 9 percent, followed by Alibaba and IBM each at around 5 percent, and then Oracle.⁴³ This concentration reflects the large economies of scale previously noted. Our estimated Herfindahl-Hirschman Index (HHI)—a measure of market concentration—is around 1,600, which is at the low end of “moderately concentrated” markets, according to Justice Department guidelines.⁴⁴ Note, HHI is a static snapshot and doesn’t capture market dynamics. The top three market share leaders have stayed in the same order. AWS invented the market for public cloud computing, is the leader measured by revenue, and has led the Gartner magic quadrant for 10 successive years.⁴⁵ Microsoft is number two, gaining 10 points of share from 10 percent to 20 percent from 2017-2020, as it pivoted to the cloud. Google and Alibaba are also making smaller share gains. Providers outside the top 10 are consistently losing share, and are down to a combined 20 percent share. We examine switching costs in the next section. Each company has its own focus:

- **AWS** emphasizes technical performance, a customer-driven product roadmap, and offers unmatched breadth and depth of its services. It has led the push into new cloud services from analytics to satellite data.
- **Microsoft** has built out its cloud service portfolio to narrow the gap to AWS. It leverages its existing relationships with customers, installed base of Office/Office 365 and Windows operating systems, and extensive go-to-market ecosystem of sellers and partners.
- **Google** Cloud has technical strengths, including its heritage in data and analytics. Under CEO Thomas Kurian, Google is improving its enterprise selling and support capabilities, and targeting vertical sectors.
- **IBM** emphasizes hybrid cloud computing, leveraging its large on-premises customer base. Especially with its Red Hat acquisition, IBM is focusing on multi-cloud and hybrid cloud management.
- **Oracle** is focused on large enterprises and mission-critical workloads. It leverages its proprietary and widely deployed database and enterprise application (ERP) software.
- **VMWare** runs on top of AWS, Azure, Google, IBM, and other clouds but no longer offers IaaS data center hardware infrastructure. VMWare leverages its leading position in on-premises virtualization and offers a common solution across the private cloud, the hybrid cloud, and the public cloud.
- **Major telecoms** (VZ, AT&T) and **legacy IT** providers (HPE, Cisco, Sun (now part of Oracle)) have exited the market, unable to keep up with the large investments required in services and infrastructure. Telecoms still have a presence in “hosting” customer IT equipment.
- **China’s** state-led economy has systematically nurtured and protected Alibaba, Tencent, Huawei, and Baidu. Alibaba dominates in China and is a major player in Asia. Tencent is growing rapidly too. The Chinese government requires non-Chinese cloud providers to operate through a Chinese partner in China.

- **Europe** is trying to build its own European data infrastructure. The EU is driving the GAIA-X project and the European Cloud Initiative, in which it expects to invest €2 billion.⁴⁶

Figure 6: Cloud provider market share and growth rate, Q1 2021 (IaaS, PaaS, and hosted private cloud)⁴⁷



Source: Synergy Research Group

The Cloud Is a Platform With a Partner Innovation Ecosystem

AWS, Microsoft, Google, Alibaba, Tencent, and others enable thousands of competitors and tens-of-thousands of partners to leverage their underlying cloud platforms and provide more specialized services. This includes hundreds of independent software vendors (ISVs) which offer additional services where the cloud platforms have no offering as well as competitive offerings. For example, data analytics ISV Snowflake has a \$60 billion market cap and competes directly with and runs on top of AWS, Microsoft Azure, and Google clouds. Similarly, VMware offers its widely used virtual machine software in partnerships with AWS, Microsoft Azure, Google, IBM, and directly competes with them.

While cloud computing is disrupting on-premises IT environments, partners are building cloud practices to both stay relevant to their customers and participate in cloud growth. For example, Rackspace manages its customers' accounts on AWS, Azure, and Google Cloud. All the major system integrators have built large businesses advising customers on how to move to the cloud and performing migrations. Accenture has created a dedicated business group called Accenture Cloud First with 70,000 employees, and it has \$11 billion in cloud-related revenue.⁴⁸ Cloud providers are encouraging this migration by building formal partner ecosystems that offer joint market-adoption programs, technical assistance, and business incentives. They also provide resources to software developers to incentivize adoption of their tools and technologies. AWS, Microsoft, and Google have built online marketplaces wherein customers can find software, data, and services from pre-approved partners and pay for pre-tested and pre-integrated software. AWS

has over 7,000 listed products and 300,000 active customers that have purchased 1.6 million subscriptions.⁴⁹ All of this gives government and enterprise customers access to innovative solutions from new and small companies they may not normally interact with, while giving small companies access to new customers.

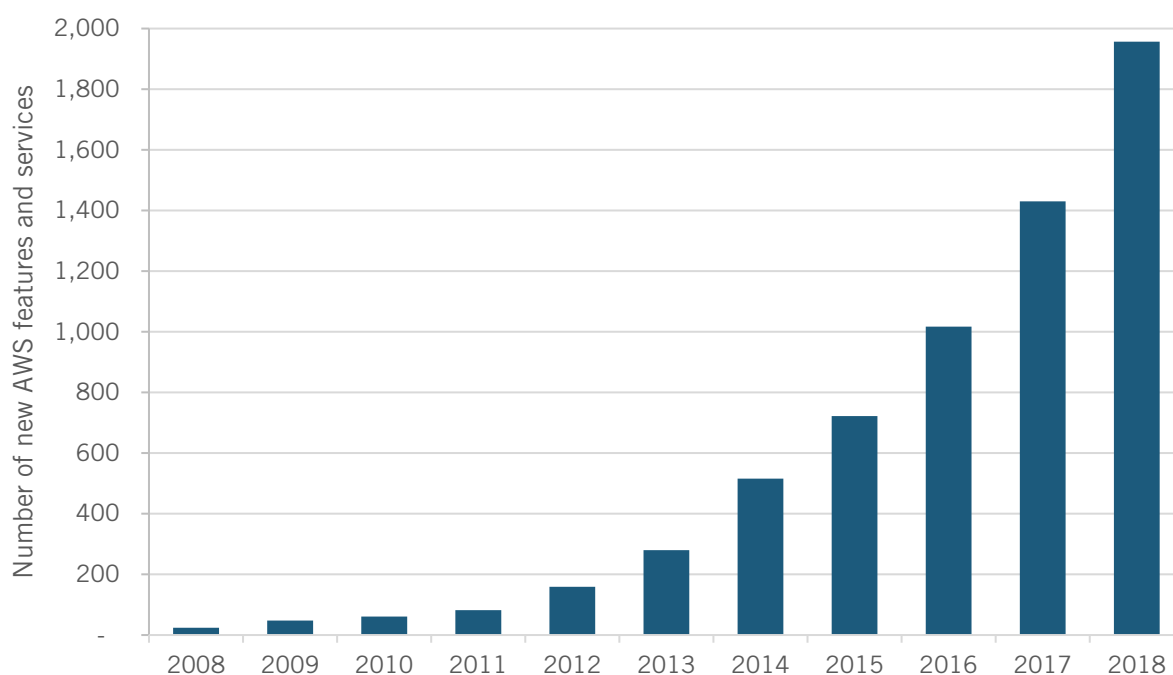
As we'll discuss in more detail, the major clouds are expansive platforms that each have a common architecture and management. They are developing network effects and attracting partners. The ecosystems with the most scale and best services win. Cloud computing also reduces lock-in by providing lower switching costs and more portability than traditional IT environments. In a rapidly changing IT industry, partners have to innovate and move “higher up the stack” to products important to their customers’ businesses. History shows that products (e.g., the browser) commoditize over time and can be absorbed as features in the underlying platform. The biggest competitor to the cloud is the more than 90 percent of the enterprise IT market that remains on-premises. In addition, given the early stage of cloud adoption, venture capitalists are pouring billions of dollars into new company formation, and there is an active market for cloud mergers and acquisitions. Leading VC firm Bessemer Ventures in 2020 tracked \$186 billion in private cloud investments and “record breaking cloud M&A activity.”⁵⁰

Products Are Differentiated “Up the Stack”

Computing is often characterized as a commodity, despite the complexities of managing and securing IT. Core services such as compute and storage are more homogenous and portable between IT environments. Cloud computing technologies such as virtual machines, containers (software that packages application code and its dependencies so that they run across IT environments), and de facto standards help create interoperability and portability. For an example, see the widely used open source Kubernetes containers.⁵¹ Published cloud prices for similar compute offerings are typically very similar between Azure and AWS, indicating homogeneity. Yet, there are differences even in these core services. For example, AWS’s Nitro system moves virtualization and management to a separate chip and uses a custom Graviton processor, thereby improving security and performance and lowering costs by freeing up resources for the customer that were previously used by the system. In addition, Nitro’s modularity enables faster innovation. Cloud providers are also trying to differentiate by creating 40 types and sizes of compute offerings that are optimized for different workloads.

Another example is object storage, which is the native technology to store different kinds of data in the cloud—and is often considered a cheap commodity. However, Azure, AWS, and Google are adding dozens of features and services, many of which are sufficiently differentiated to be separately priced, including security, compliance, availability, retrieval performance, and content management. Add database, analytics, ML, messaging, media services, security, developer tools, Internet of Things, AR/VR, and robotics, and cloud computing shows real differentiation. These higher-level differentiated services will in time become a growing part of the cloud. Though counting services is arbitrary, Azure and AWS offer >200 different named services. In addition, providers are adding many new features to these services, which itself points to differentiation (figure 7).

Figure 7: Cloud innovation

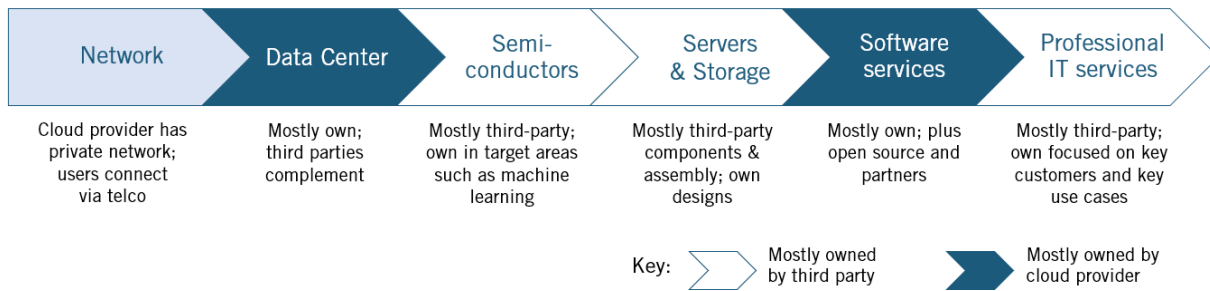


Integration Across the Industry Value Chain Is Low, But Increasing

Cloud providers purchase servers, semiconductors, memory, and storage, choosing to “buy rather than build” as their primary approach (figure 8). However, given their scale, the big three are increasing their role in the value chain, using their own designs and components. A good example is in semiconductors, where they design custom chips for ML to provide better performance, costs, and product differentiation (e.g., Google Tensor Processor; AWS Inferentia and Trainium). They also heavily customize servers, storage, and switches for better energy and operational performance, though they are built from largely standard industry components by original equipment manufacturers. Microsoft and AWS have also designed their own branded computing hardware for edge applications, such as Azure Stack and AWS’s Outposts and Snowball, which are still in the early stages of adoption.

In order to improve the performance of their services, cloud providers largely own or lease and operate their own global internal communications networks linking their data centers and regions. Their customers, however, purchase bandwidth from telecom carriers in order to reach cloud connection points. In addition to the networks themselves, the big three providers also primarily own and operate their own data centers, seeing them as a source of differentiation. Yet, they also use third-party data centers to supplement capacity and enter international markets quickly. Cloud providers’ core value-added is their software. Open source also plays an important role, and for some services (e.g., Kubernetes containers, Cassandra databases, Kafka event streams), cloud providers build a “managed service” based on open source code. Cloud providers rely heavily on third parties for professional IT services such as consulting, workload migration, and application modernization. While providers have their own professional services staff, they are often used for key customers or technically challenging deployments.

Figure 8: Cloud value chain



Profitability Is Similar to Large Software Companies

We estimate cloud operating margins at scale are in the 30 percent range, similar to the largest software companies.⁵² *Operating* profit margins for the largest scale cloud (AWS) were 29.8 percent in 2020. Adjusting Microsoft's "Commercial Cloud" business's *gross margin* to operating margins (so it is comparable to AWS) shows 36.2 percent.⁵³ However, this may be overstated due to cross subsidization between Microsoft's Azure cloud and its traditional on-premises software. Microsoft is also able to decide how to categorize revenue within large, multi-product enterprise agreements that include both cloud services and traditional software. Google Cloud is not yet profitable, with -32.4 percent operating margins in Q4 2020, though it has improved over the last 3 years. Losses reflect both the big investments Google is making to compete with AWS and Microsoft and that it has not yet achieved the scale for profitability.⁵⁴ Note, cloud services have a different revenue recognition model than traditional licensed software, which may make it appear less profitable in margin percentage terms with GAAP (generally accepted accounting principles) accounting. Most cloud revenue is on a usage basis, so revenue recognition is spread out over time—whereas most traditional software has a large license revenue component that's recognized at the start.

Cloud economics are still evolving. Cloud services may have lower gross margins than traditional software, because it has substantial cost of goods sold (COGS) due to energy, data centers, and network bandwidth, plus the aforementioned usage-based revenue recognition. However, the step down to operating margins may be less severe. Traditional software companies have large sales and marketing costs, typically 20–25 percent of revenue. The cloud has potentially lower sales and marketing costs given the cross-selling of higher value-added services from the same platform, more scalability, and greater self-service use. From an industry value chain perspective, we expect cloud to accelerate the shift in revenue and profits from hardware and components to the cloud, and from legacy software to the cloud. Long term we also expect cloud to shift profits away from people-based IT services, since those services (e.g., IT outsourcing, maintaining servers) are increasingly provided by cloud automation. However, over the next three to five years, we expect cloud computing to create offsetting demand for people-based IT services to help users migrate to the cloud, refactor applications, and manage and secure their cloud environment.

The Cloud Is a Good Fit for Most, But Not All, Applications

Cloud computing is a general-purpose technology, but there are always exceptions where the cloud is not the best fit. For example, some specialized applications in high-performance computing (HPC) are not designed for the scale-out horizontal architecture of the cloud and still

run best on HPC-optimized infrastructure. While most applications can be run in the cloud, they may not take advantage of cloud-native architectures and get the full benefits of the cloud. Some workloads are very sensitive to latency and are best located close to large on-premises datasets or in an edge device such as a self-driving car. These are not yet ideal workloads for the cloud. Some legacy applications may simply not have the business case to move to the cloud. Historical regulatory models may also keep some workloads on premises. However, the bulk of IT workloads are well suited to the cloud. Further, like many newer technologies, as cloud performance increases, the workloads not suited to the cloud will become an increasingly small “island.” The cloud is not perfect for everything, but be skeptical of statements claiming a given workload won’t run well in the cloud.

V. WHERE CLOUD COMPUTING EXCELS

Cloud Drives a Better Business and Technology Model

Cloud computing uses new technologies, but just as importantly, it employs a new business model that provides compelling value for businesses and governments. The cloud is not right for all applications. However, with millions of customers using cloud computing, it has proven it can provide superior value in cost, productivity, and business agility. Government customers with the most stringent security needs, such as the CIA, are heavy users of cloud computing. Media customers with the biggest high-bandwidth video needs such as Netflix and ViacomCBS are “all in” on the cloud. Banks and brokerages with the biggest need for transactional certainty such as Capital One, JP Morgan, and Goldman Sachs are heavy users of the cloud. Gaming customers with the biggest need for real-time interactivity such as Zynga are built on the cloud. Early on, companies thought they could innovate faster and operate IT better than the cloud. Today, they’re realizing that even if they have the money and technical skills, they’re unlikely to be better at IT than the cloud. This is similar to the way they realized that the electric grid is better at providing power. Zynga’s story is especially compelling, as they moved off the cloud to their own data centers, believing they could customize their own infrastructure for technically demanding game-specific requirements. They ended up moving back onto the public cloud when they found it had greater scaling, better economics, and faster innovation.⁵⁵ While cloud computing is still maturing, its direction is clear. Over the long term, we’re moving from a world wherein most enterprise IT is on premises to a world in which most enterprise IT is in the cloud. History shows spending time and resources to maintain past models is unlikely to succeed. It’s better to embrace the “cloud-future” than fight it.

Cost savings is often the starting point for new technology adoption, and enables executives to make the required business case to move to the cloud. In addition to the double-digit annual price-decline data previously shown, an IDC study sponsored by AWS shows 31 percent operational cost savings comparing cloud computing with similar on-premises infrastructure, and 51 percent savings when including IT staff costs and unplanned down time.⁵⁶ McKinsey has shown that the cloud saves costs across each step of the business system, including 5–10 percent of research and development (R&D) costs, 10–20 percent in supply chain and manufacturing, and 15–20 percent in business support costs.⁵⁷ These cost savings are backed up by thousands of customer case studies and testimonials.⁵⁸ To fully achieve these savings users must put in place proper governance and management. Nonetheless, the major drivers of these cost savings are clear:

1. Greater economies of scale and scope, boosted by technical innovation.

2. Better capacity utilization from aggregating millions of customers with different usage profiles. Traditional servers have a 20 percent average utilization rate and, even with virtualization, typically achieve ~50 percent rates, according to IDC. With the public cloud, utilization is well beyond that, and will increase as cloud spot markets grow.
3. Lower management costs from programmable infrastructure and automated tools that drive 62 percent better IT staff productivity.⁵⁹
4. Variable costs that can scale up and down with business need rather than pay-up-front fixed costs. This avoids capacity shortages or expensive excess capacity, as well as “lumpy” upgrade cycles that are risky and often delayed.
5. Better availability, durability, and resilience through horizontal “scale out” architectures that dynamically shift workloads to resources that run in parallel. Workloads can be spread across different availability zones (60 miles apart with separate telecom, power, flood plains, earthquake zones) and, if desired, in different countries.⁶⁰ Durability for objects stored in Azure and AWS is at least 99.999999999 percent.

The bigger value from the cloud comes from greater agility, speed, and focus on the core business. At its simplest, the cloud outsources IT functionality and IT management that is not a core competency. Users choose what and how much or little IT they shift to the cloud, depending on their business needs and IT skills. They can outsource a full end-to-end application to the cloud via a SaaS solution, or they can just move back-up and recovery for select workloads. This enables users to move faster, innovate more quickly, and focus on their customers. We’ll discuss how cloud enables agility and speed in more detail below.

The cloud provides global reach. Even the smallest organizations can tap billions of dollars of investment in global infrastructure. Google Cloud service is “available in 200 countries and territories;” AWS “serves over a million active customers in more than 240 countries and territories;” and Microsoft Azure says it is “available in over 140 countries” and is likely well above that.⁶¹ Organizations can provide services close to their target markets in order to improve network latency, and store data in designated countries for regulatory purposes. Users can specify where their data is stored and their computing operations run, and can secure them with encryption and a broad range of security tools.

Cloud Is a Rich and Capable Platform, Not Just a Collection of Services

Like many newer technologies, cloud computing is rapidly getting better as it matures. Further, the cloud is not just a “product.” It’s becoming a richer, more capable platform that provides all the major IT services, rapid access to new services such as Internet of Things and ML, and thousands of complimentary services from a large partner ecosystem (figure 9).

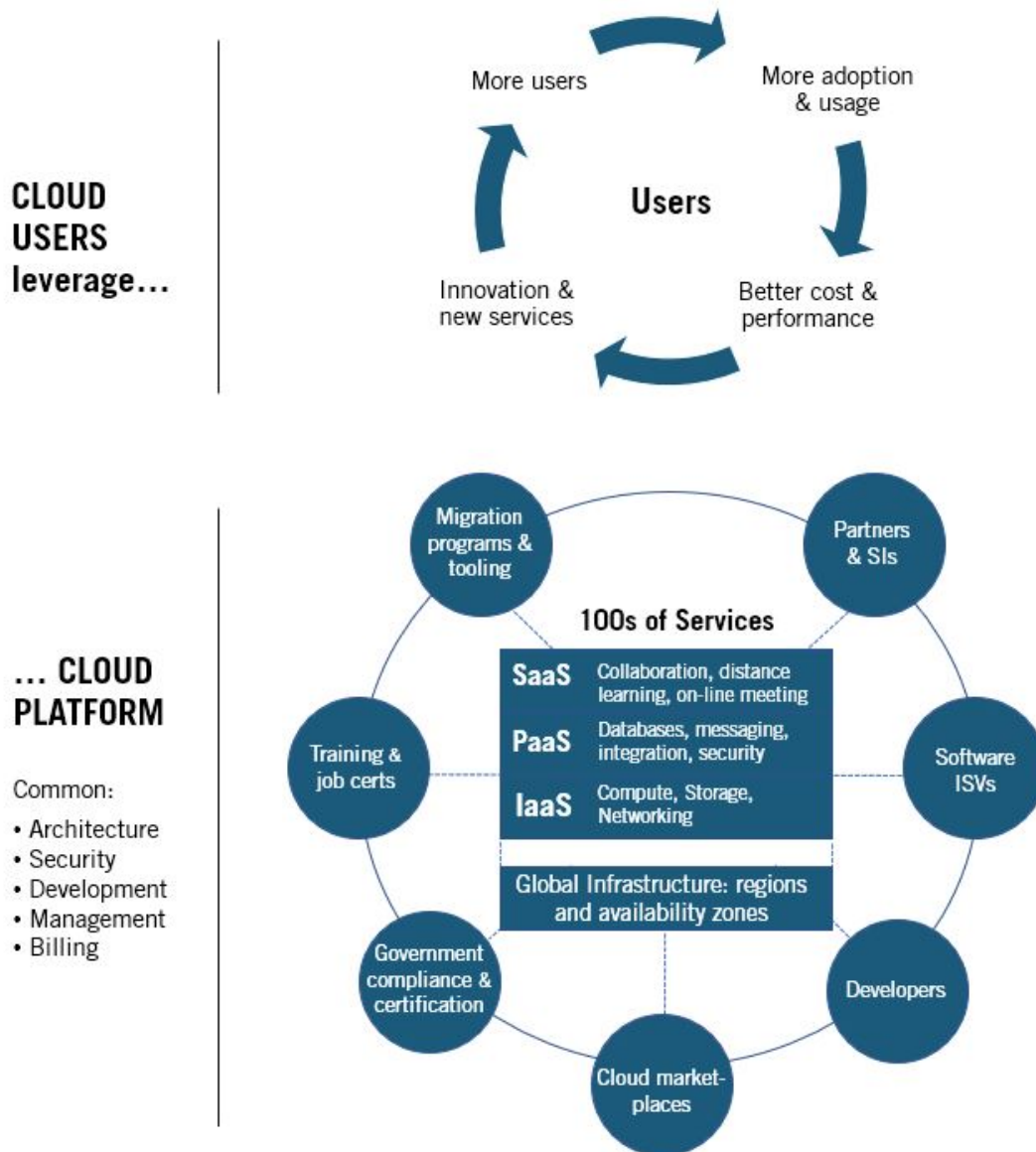
Expansive services: IT and beyond. The cloud increasingly provides just about every IT-based service customers need today—and maybe tomorrow. The cloud goes far beyond basic services such as compute, storage and networking. Initially, in services such as relational databases, cloud providers didn’t provide all the features and high-end performance needed by the most advanced users. However, now cloud providers can scale relational databases (from Oracle, Microsoft, MySQL, Postgres) often at lower cost and comparable availability (e.g., multi-region, multi-master services). In addition to traditional relational databases, AWS offers new database

technologies optimized for different kinds of workloads, including NoSQL (unstructured data in web applications), graph relationships (social media), document stores, block chain ledgers, and time-series databases. Even more, cloud providers offer solutions for new application domains, including Internet of things, content delivery and media services, satellite ground stations, robotics, ML, and mixed reality (AR/VR).

Common architecture, security, and management across the platform. The major clouds are not just a collection of services. These services leverage common software development architectures, operational models, and management tools within each cloud. They are built on common core technologies and provide common services, including integration, data management, security, and compliance. Cloud platforms are built with application programming interfaces (APIs), which are key to modern service-oriented architectures. APIs expose functionality in standard, reusable ways so other services in the cloud can use them. Cloud platforms also use common billing and cost-management tools. All this lowers costs, increases innovation, and expands capabilities. As a result, customers are able to add new services in just hours or days, without having to learn new skills or enduring slow and expensive deployments.

Partner ecosystems build on the cloud platform. The major clouds are platforms that are so rich and deep that they can tackle IT challenges that were not doable before, or were cost-prohibitive.⁶² In addition to hundreds of first-party cloud services, thousands of partners use APIs to “plug in” to the major clouds. Because APIs expose functionality in standard, reusable ways, partners build on top of and re-use the functionality of the underlying cloud platform. Keeping with the electric grid analogy, an electric socket is analogous to an API: Any device can plug in to the service and receive electricity so long as its plug is built according to the standard (size, dimensions, load). Cloud APIs enable a huge ecosystem of third parties that offer additional services to plug in to the platform. (Note, APIs also mitigate lock-in by making it easier to unplug). These partner offerings increase the depth and richness of cloud offerings, thereby making the cloud more useful.⁶³ The cloud platform and partner services are economic complements—each supports the other. *The major clouds are ecosystems that offer superior value from the combined offering of the platform, which is greater than what any single company can offer.* As a result, the major clouds are starting to exhibit demand-side economies of scale called “network effects,” where by customers and partners go to the platforms that offer the most services and biggest market.

Figure 9: Power and richness of the cloud platform



SI = System integrator. ISV = Independent software vendor.

Cloud Lowers Switching Costs Despite Lock-In Concerns, Multi-cloud Best Practices

Customers have concerns about their ability to leave clouds or move workloads across providers. This stems in part from IT industry historical practices. Yet, cloud computing breaks down lock-in. Cloud lowers switching costs (to move to a new provider), increases portability (to move applications and data), and provides more interoperability (to exchange and use information). While concerns are real, cloud technology, economics, and licensing lower lock-in. Customers should use the five approaches below, combined with exit planning, to mitigate risks.

First, the cloud model of pay as you go, only for-what-you-use, ramp-up, and ramp-down, lowers switching costs and increases cost transparency. This stands in contrast to most traditional software agreements that typically require license payments at the start, regardless of actual usage. Traditional multi-year software enterprise agreements also have non-transparent cross subsidization among multiple products. Cloud users receive granular billing that can show what resources were used by which users, in what region, and at what time. Customers should make sure they use the cloud's built-in tools and cost alarms to manage and lower costs. As an incentive, cloud providers offer volume discounts at the users' choice, such as Google's straightforward "Sustained Use Discounts." To be clear, *there still are costs* to moving users' data to exit the cloud, but they are transparent and there are tools to move the data and manage the costs.⁶⁴ Customers should also consider the likelihood of actually needing to switch versus a general concern about lock-in.

Second, cloud providers and partners offer server migration and data-transfer tools to move workloads to and from the cloud, diminishing lock-in. These tools automate and simplify the migration process. Cloud providers also offer tools to export users' virtual machine images. These tools are widely used by customers and partners on a weekly basis. For example, VMware's leading virtual machine technology works on AWS, Microsoft Azure, Google Cloud, IBM, and others with "push button" automation moving VMware workloads between premises and the cloud. The cloud providers themselves also have nascent tools that aim to provide more ambitious multi-cloud management. This is in part to win workloads from each other. For example, Google's Anthos and Microsoft's Azure Arc help provide an integrated view of users' IT resources across clouds. IBM and Red Hat also offers multi-cloud tools, technology (e.g., open source frameworks such as Open Stack), and professional services.

Three, users can architect applications using services such as containers, microservices, and APIs to make their workloads more cloud agnostic. These technologies, such as Kubernetes containers, run a workload in smaller independent "packages" that are not dependent on specific features of any one cloud. In that way, they can be moved across different clouds and on-premises. By breaking large applications into smaller pieces (e.g., microservices), applications can be designed so that proprietary services are limited to smaller, clearly defined pieces that can be "designed around" or, if needed, re-implemented. Architecting IT workloads into independent "loosely coupled" services with well-defined interfaces (APIs) makes portability a lot easier. Further, key cloud services are based on publicly available open source alternatives, such as Linux VMs, Kubernetes containers, and Postgres databases.

Fourth, what one customer sees as a proprietary service with switching costs, another customer chooses as value-added innovation. Cloud providers offer higher-level services such as ML that leverage proprietary capabilities. Cloud providers invest in new functionality that is specific to their cloud. While this creates new capabilities, when customers *choose* to use new proprietary services this also creates some switching costs. However, some level of switching costs (which can be mitigated) is not the same thing as lock-in, and should be weighed against the value of new innovation.⁶⁵ Customers can choose what they value most.

Fifth, portability across clouds has costs as well. Designing cloud-agnostic architectures typically creates more complexity and costs for customers. And re-factoring applications so they can be moved in portable containers also has costs, which not every application may justify. Further,

most organizations don't have enough skills to build and manage cloud computing. Adding multiple cloud providers increases this skill challenge. While the major clouds offer similar services for core capabilities, they are implemented differently. Multiple clouds create a more complex environment, which is one reason why most customers are not using multi-cloud management tools.⁶⁶ Multiple clouds can also lead to “least common denominator” approaches wherein customers don't take full advantage of advanced services, which can have big opportunity costs for the business. Customers should evaluate whether multi-cloud tools achieve their goal of provisioning, securing, and managing operations across clouds and on-premises. It remains to be seen whether multi-cloud management offerings such as Google Anthos, Azure Arc, and IBM Red Hat will add sufficient value to offset the additional cost and complexity.

Primary cloud and multi-cloud best practices. Many customers concentrate usage on one primary cloud for IaaS and PaaS. They benefit from innovative features, volume discounts, and being able to build skills and IT productivity in their primary cloud. They should design for portability and create exit plans to mitigate lock-in concerns. Exit plans should include clarifying data formats, database schemas, and system dependencies, as well as the actual business need for portability. Deploying cost and usage tools are key to managing costs. These tools are available from cloud providers and from partners such as CloudCheckr, Apptio Cloudability, and CloudChomp. For basic needs, users should avoid operating the same workload across multiple clouds, as the cost and complexity are likely to yield disappointing savings in practice. However, it likely makes sense to add services from other clouds that provide additional capabilities important for their business. Users with distinct business and mission needs are likely to use multiple IaaS-PaaS cloud providers. This is an opportunity to place workloads in the “best of breed” cloud for that service or domain, for example, adding an additional cloud to use their ML service or medical applications from Microsoft Cloud for Healthcare. In this case, it's better to run a new workload all in one cloud and use different clouds for different workloads. Similarly, it's more difficult to architect and operate the same workload across multiple clouds. If we include SaaS applications, most customers will be multi-cloud because they are likely to use many SaaS applications, for example, Microsoft Teams for collaboration, SAP Concur for expense management, or Salesforce.com for customer relationship management.

In summary, multi-cloud should be less about spreading basic compute tasks across multiple clouds in an effort to save money and reduce lock in—these goals can be better achieved in other ways. Rather, multi-cloud is better used to capture best-of-breed services across multiple clouds to maximize clouds' value.

Cloud Boosts Agility, Speed, and Innovation

Cost savings are often a starting point for moving to the cloud—and cost is easier to quantify. However, customers say agility, speed, and innovation matter most. Cloud computing enables users to bring products to market faster by experimenting quickly at low cost. An on-demand general-purpose server (8 cores, Linux, 32 GiB memory) costs just \$0.384 per hour, lowering the cost of trials and experimentation.⁶⁷ For example, biotech company Moderna completed the computing for its vaccine in 42 days, compared with the normal 20 months.⁶⁸ Supersonic-aircraft designer Boom tested hundreds of designs with thousands of flight simulations, using 53 million core hours. Because the cloud can scale-up capacity “horizontally” (using 1,000 machines for 1 hour rather than 1 machine for a 1,000 hours), it can reduce time spent on

product design. Boom is a start-up—the cloud democratizes access to the latest capabilities and scale putting small companies on more equal footing with large corporations and governments.

The cloud enables a business to be more agile by rapidly standing up and tearing down IT capabilities that underpin their operations and workflows. With the cloud, it no longer takes weeks or months to order servers and have IT administrators rack, install, and provision new resources. Instead, compute resources can be provisioned via software tools in just hours or days through the cloud management console. IT is no longer a fixed cost, as it becomes “temporary,” shut off when the task is done. Cloud software development and “DevOps” also bring speed and automation to building and deploying new code, thereby increasing developer productivity. It brings a new culture and organizational approach that enables more rapid software release cycles and rapid product development. This extends into IT management and operations, which are automated via programmatic software controls. The contrast from traditional on-premises environments could not be greater. One customer, following completion of a robust-but-expensive year-long ERP-application on-premises installation said, “Thanks, you’ve just poured concrete over my business.” The cloud is the opposite. It enables speed, scalability, and rapid development that brings a more-nimble culture and organizational approach.

Cloud Infrastructure-as-Code Enables Automation, Productivity, and Specialization

We’ve emphasized that in the cloud, IT resources such as servers, storage, and databases are not fixed but can be added or taken down at scale in near real time through software. In contrast, manually provisioned systems with hard-wired dependencies between components are increasingly legacy practices. This has significant implications for how users manage their infrastructure. It drives productivity and flexibility in IT operations, as well as improved security. Software templates digitally describe collections of IT resources (databases, storage, compute virtual machines), their parameters, and any dependencies among them. These templates become “configuration files” the cloud treats as executable software code. That means the templates can be launched, deleted, and managed as a single complete system or “stack.” Further, IT administrators can use these templates to change the IT infrastructure in repeatable, predictable, and auditable ways. *In the cloud, we no longer just program an individual application. The whole IT infrastructure itself becomes programmable.*

An entire infrastructure with thousands of user accounts can be provisioned across geographic regions at huge scale with only a few key strokes. IT administrators can test, tune, and launch the IT infrastructure in hours from the management console. IT architectures can be consistently cloned with version controls and re-used. The infrastructure itself can also be programmed to automatically respond to events. Run-time metrics can be set to trigger alarms or automated actions. For example, an action command might be: “If increased demand causes performance to fall by 20 percent, then automatically add 5 large servers and load-balance across them.” Performance and compliance can be monitored and optimized in an automated fashion. For advanced users, the software development process itself can be incorporated into this. Software code-pipelines can be built and deployed, and then integrated into the existing IT infrastructure. *Notice the specialization and division of labor:* The cloud automates the underlying IT so that businesses and governments can devote resources and focus to applications and business processes that help customers and improve outcomes. Once again, the cloud brings cost savings, IT productivity, and business agility. Cloud IT becomes a scalable service in a way that was simply not possible before.

VI. CHALLENGES FOR CLOUD SUCCESS

Like most newer technologies, there are challenges that get in the way of more robust and deep adoption of cloud computing. Some of these issues are not unique to the cloud. Nevertheless, we call out how the cloud heightens their importance or changes how they are addressed. Recurring themes include making cloud simpler and easier to use and providing best-practice guidance and better tooling and automation.⁶⁹ The industry is starting to recognize and address these issues. Over time, this will help accelerate adoption and ensure customers get full value from the cloud.

Achieving Stronger Security and Compliance: Automation and Management

Security and compliance can be better in the cloud than on premises, but automation and management are needed to achieve these benefits. The first question many customers ask when moving to the cloud is about security. Security in the cloud is now as good, or in many cases better, than on premises, as the CIO of CIA recently concluded.⁷⁰ However, the cloud has a different “shared security” model between the cloud provider and the customer. The cloud provider is responsible for securing the underlying infrastructure of the cloud, while the customer is responsible for what they put in the cloud (e.g., guest operating system, applications, data) or connect to the cloud. The most security-conscious users around the world from intelligence agencies, defense ministries, and banks now rely on cloud computing. Security remains a challenge for all, but these large and sophisticated organizations are demonstrating their confidence in cloud security by moving their workloads to the cloud. They also rely on 90 compliance regimes around the world to prove security is implemented in production, for example, FedRAMP for unclassified U.S. government workloads, SRG IL-6 for secret U.S. Defense Department data, C5 in Germany, G-Cloud in the United Kingdom, GDPR for data privacy in Europe, plus industry-specific regimes such as HIPPA in health care and PCI-DCC for financial payments.⁷¹

Cloud computing presents new patterns of security, whereas many organizations are comfortable with security models they already know. The cloud provider takes care of physical security (data-center perimeter, rack access, back-up power, etc.) and security controls are implemented in software (hypervisor, operating software, access controls). This is different from users owning or leasing a data center they operate with their own personnel. However, it also offers advantages. Security is now provided by a hardened cloud infrastructure, wherein “best in class” security is a core competency and crucial to the business. Cloud security is the top priority and receives billions of dollars in investment. And customers now only have to worry about a smaller scope (“attack surface”) of IT.⁷² They can focus on securing their data, applications, and intellectual property. It also means security is built by design into the cloud platform and each service.

Granular and scalable security policies. Customers use identity and access management (IAM) to set policies that, for example, allow only named users in a specific geographic region to have access to specific IT resources down to specific rows in specific databases at certain times of day. Multifactor authentication validates users not just with a name and password but with an additional code, typically generated from a small hardware device. Users decide which geography to store their data in and where their workloads run. Only users move their data.

Multi-layered security and zero trust. Cloud providers use many layers of security, starting with virtual private networks, sub-network isolation, and web application firewalls. Cloud-scale infrastructure is well positioned to protect against distributed denial of service (DDOS) attacks.

Perimeter defense is insufficient and security extends beyond the network to applications and the data itself. For example, new data loss prevention services use ML to identify anomalous access patterns, such as data that is being accessed by a non-typical user or in an unusual way. Cloud services continuously verify whether a user, machine, or application is “known” and what access and privileges they have. Data is encrypted both at rest when stored and in transit. Further, customers can generate and hold their own encryption keys the cloud provider does not have access to. These approaches are especially important in a cloud world, because the physical location of the resource is not a primary determinant of security.

Platform security, continuous monitoring, and transparency. The cloud is well positioned to improve security because it is a platform with common security models, IAM policies, API models, encryption, logging, and monitoring. Cloud configuration tools can ensure predictable and consistent infrastructure builds with pretested security. Continuous monitoring software enforces policies about who has access to what IT resources. This prevents unauthorized configuration changes to software. It also tracks events down to individual API calls, showing who requested what service and when, which is key to security incident response. Monitoring software also provides alerts of common security misconfigurations, such as leaving open certain network ports, allowing public access to storage buckets, or not using strong multifactor authentication on root accounts. Just as infrastructure becomes code, security becomes code.

Improving security integration, automation, and management. The richness of the platform can create a stronger security posture. However, it can also be complicated and, as with on-premises, requires dedicated skills. Security was identified as the top cloud challenge by 81 percent of users in the Flexera 2021 State of the Cloud Report.⁷³ Many common vulnerabilities are driven by misconfigured software, not keeping up with a changing IT environment, or customers simply not using available security features. Higher-level services that integrate prescriptive guidance, best practices, and security artifacts into easier-to-use packages would help customers take full advantage of cloud security. Prior efforts can be rationalized into a simpler coherent approach. At premium is simple, specific guidance about who is responsible for what in a shared-responsibility model—and how to best implement it with cloud tools and processes. Cloud providers can work with NIST and others to curate and publish these best practices, and where appropriate, develop them into common industry frameworks.

The cloud is also well positioned to provide better tools that automatically detect and correct vulnerabilities. As the industry has created new defenses, adversaries have responded with new counter-techniques. This “security leapfrog” has created dozens of tools that have their own data formats, interfaces, and management tools. Integrating across these different tools will give customers a more complete view of their security posture. Automation built in to the cloud should also take over more of the continuous monitoring and management. This can make security more consistently deployed, up to date, and easier to use. There are also more opportunities to apply ML (already embedded in the cloud) to sort through the thousands of security events captured in system logs. The cloud itself can proactively correct straightforward vulnerabilities, and surface complex situations for experts to intervene in. Just like in storage and servers, the cloud should manage more of the lower-level security and enable security experts to focus on the most-serious challenges.

Enhancing Governance, Managing IT and Data, and Reducing Complexity

Successful cloud adopters have found that strong governance and management is still required after moving to the cloud. Customer business leaders and IT executives need to jointly govern IT and data, paying special attention to the new methods of the cloud. Surveys show 75 percent of cloud users call out “governance” as a cloud challenge.⁷⁴ Customer IT leaders need to set access and security policies and ensure compliance, sometimes across thousands of users. Who gets access to what IT resources and on what terms is especially critical in a cloud environment because thousands of servers can be provisioned in minutes and there is quick console access to powerful services such as facial recognition. IT leaders need to establish baseline infrastructure configurations and set guard rails. Operational performance against service levels should also be managed and improved. This takes on new importance in the cloud, where IT is a service, not a physical asset. IT costs, including any cost-recovery, can be tracked, enforced with automated controls, and allocated across different business owners. Best practices need to be set so IT resources are not unintentionally left running. Again, the cloud provides enhanced visibility and granularity of spending, but also risks from more rapid access via the management console. Software development and operational best practices (DevOps)—the traditional preserve of the IT organization—also should be recast to take advantage of cloud-native approaches such as continuous integration and continuous delivery (CI/CD). Training and certifications in new techniques will be required in order to use the cloud well. The cloud is well positioned to bring together the many separate tools and services into easier-to-use packages. They can be more automated, have built-in diagnostics and alerts, and come with best-practice instructions.

Data governance is critical to capturing data’s value, and deserves special focus. Data is increasingly recognized as a valuable asset. Customers need to set data-classification tiers, access policies, back-up and recovery (RTO/RPO), and retention policies.⁷⁵ Data governance is needed to manage disparate data formats, indexing, and tooling requirements. Technical and organizational guidelines need to ensure data is not bottled up in silos. Especially with cloud scaling, data volumes can get big quickly. Reference architectures that show customers how to stage data differently in the cloud and increase shared access can help. Cloud providers are also maturing new services called “data lakes” that address these needs. Techniques that move compute to the data, rather than the increasingly difficult approach of moving data to compute, are promising.

Reducing complexity. As cloud capabilities grow, complexity grows with it. Complexity is at the root of many challenges. 451 Research’s Cloud Price Index tracks 2 million product offerings (SKUs) across AWS, Google, Microsoft, Alibaba, and IBM.⁷⁶ Even basic functions like billing can be complicated. Integrating sprawling “point” features and reframing them into higher level services or suites that solve real business needs would help many customers. Cloud services can also be easier for customers to use, with priority focus on better automation. Best-practice content exists in many forms, and can be curated and packaged so it’s easier to use. Promising areas include service-level measurement, service-development frameworks, and migration templates. Reducing complexity will also help harden the cloud as it becomes integral to the economy.

More broadly, for all of these needs, the cloud is a new operating model with a distinct culture. To fully capture cloud benefits and mitigate risks, successful adopters have found that strong IT governance is key to building a cloud culture and organization. This involves change, which can challenge existing organizations and depreciate hard-won skills. These concerns need to be

addressed so that the change brought on by cloud can be implemented. Culture and organizational change require special attention and are key enablers in the shift to a cloud world.

Building New Skills and Inclusivity

Cloud computing requires different skills, ranging from cloud-native application development to data architecture to operations. With high unemployment and industry clamoring for skilled employees, this is win-win. Demand for cloud skills far exceeds supply. However, these jobs require new certifications and new experiences. The IT industry and cloud providers offer free training in many formats including online courses, interactive labs, virtual day-long training sessions, and job-based learning paths, in multiple languages. Yet, training doesn't scale well and requires substantial investment in time and resources. Cloud requires not only new technical skills but also new domain skills and modern workflows throughout the organization. In 2020, Microsoft announced a Global Skills Initiative to reach 25 million people by year-end. AWS announced a goal of training 29 million people globally by 2025, supported by free training.⁷⁷ From a business perspective, the providers recognize that the shift to the cloud can be quickened by training customers and partners to better adopt and use the cloud.

From an economic perspective, technology's impact on the level and mix of employment is not new.⁷⁸ Like prior skill-based change, the cloud will cause labor market disruptions as demand falls for jobs in traditional areas (server and storage administrators) where cloud automation is more productive (patching, updates). At least since early 19th century England, when "Luddites" smashed textile machinery, technology change has challenged existing skills and organizations. Yet, over time, technology also creates new jobs and organizations adapt. Governments and universities have studied technology-driven job losses and skill shortages and launched workforce training programs to help. Prior Information Technology and Information Foundation (ITIF) research on workforce adjustment also recommends actions to be taken.⁷⁹ At the same time, companies and governments need to invest more in training and reskilling. In particular, training can focus on gaps in priority areas such as IT security, data and analytics, and ML. The benefits of the cloud need to be equitably shared and all communities need to be able to participate. Industry can be part of the solution. Building skills and inclusivity can be improved if not quickly solved. We return to this in the government policy section.

Making Cloud Adoption Easier

To accelerate cloud adoption, moving workloads to the cloud needs to be simpler and faster. While cloud automation can make technical deployment easier, for example, provisioning new servers and user accounts, it can be a complicated year-long initiative. The State of the Cloud Report shows understanding application dependencies (51 percent of users), assessing technical feasibility (48 percent), and assessing costs (44 percent) are the top challenges.⁸⁰

Challenges to IT projects extend to cloud workloads, but in new ways. Building a clear business case and funding model is a required part of every IT migration. However, the shift from an on-premises capital expenditure model to the cloud's pay-as-you go model doesn't always fit with existing budget procedures and can make "apples to apples" cost comparisons difficult.⁸¹ More-automated planning templates can help. Further, customers often don't have a full view of the resource dependencies (e.g., directories, databases) that are required to move a workload. IT inventory and mapping tools should handle this more quickly and accommodate cloud-specific architectures. Similarly, the cloud's distinctive approach to security typically requires substantial

advance work with the security organization for approvals. Making the many security artifacts and frameworks simpler to use would help. While there's no panacea, better integration and automation are a key part of this. Cloud providers also offer free migration tools for servers, databases, and data itself. Expanding their reach to comprehensively cover the end-to-end system (not just pieces) will ease the implementation. Further, better integration across multiple tools and data formats that are part of a migration will help produce a more seamless process and high-performing outcome. More broadly, customers need to better capitalize on the cloud's capability to quickly stand up pilot workloads to cut through adoption issues. *Customers can pilot it rather than debate it.* These can be operational pilots with customer data and cloned applications, not just mock-ups. Cloud providers could expand offers for a simple pre-approved pilot as part of every migration, for example for one month or other defined value.

Balancing “lift and shift” versus re-factoring and re-architecting for the cloud. There is no one right answer. Lift and shift, or moving an existing workload in its current structure, is simpler and gets customers to the cloud quickly. However, many important benefits of the cloud (agility, resiliency, managed services) are not captured. Re-architecting and re-factoring applications are more difficult, costly, and time consuming. However, they build the long-term foundations to capture the full cost and business value of the cloud. Customers should use both approaches. Providers and partners can help by developing better tooling to determine which approach to use when moving applications to the cloud. Lift and shift may be sufficient for static, low-value applications, while re-architecting can be the right approach for higher-value applications.

Ensuring organizational incentives and alignment. Aligning business or mission owners and IT leaders is a best practice, and becomes crucial given the cloud brings a new operating model. Hence, customer executive sponsorship is critical to resolve the inevitable implementation challenges. Many cloud initiatives stall after an initial project. Executive leadership helps drive broader adoption of production workloads and encourages enterprise-wide impact. This is not just a technology discussion but a conversation about business value and the accompanying business transformation that is needed. Early adopters have found that a cloud “center of excellence” creates an organizational home to support adoption and capture best practices. It's worth exploring whether elements of this could be translated into more standardized repeatable templates.

Engaging partners to help address the skills gaps that are a big part of cloud adoption challenges. Migration at scale requires “softer” organizational transformation and operational skills, in addition to “hard” technical skills. This is not unique to the cloud, but the newness of the cloud can exacerbate this challenge. Partners often play a key role in the actual migration process. However, many partners are still building the same new cloud capabilities as customers. This will resolve in time as partners invest to capture the business opportunity. Today, many partners have business conflicts because the cloud is cannibalizing existing revenue streams even as it's opening up opportunities. Cloud providers understand the importance of partners. Yet, they can help ease the transition through partner programs that more substantially improve the cost-benefit for partners and grow partner migration capacities. This can target key areas that often block migrations such as databases, mainframes, and security.

Adapting to Hybrid Cloud, the Emerging Edge, and 5G

The cloud is becoming a continuum of computing choices. The cloud is not just the central core. It is extending to local city facilities, on-premises data centers, and edge devices. Data is growing particularly at the edge of the network, generated by machines and devices, which is creating new computing demands. These large data volumes will need to be managed and secured. Some edge applications, such as autonomous vehicles, will require near-real-time processing, and IT systems will need to determine what data will be locally processed and stored, and what will be sent to the central cloud. The key is each location and workload presents choices and trade-offs about what is most important for an application and for the customer's business priorities.

Hybrid cloud computing is more a transition than a long-term design goal. The hybrid cloud approach emerged from contrasting a central remote cloud utility with on-premises customer computing. In part, this reflects the history of IT models shifting from centralized computing, such as mainframes, to distributed models such as PCs and mobile devices. The idea of hybrid cloud combines central cloud computing and traditional on-premises IT so that they work together. Industry politics and competitive positioning also play into this debate. The reality is the bulk of 20-plus years of IT spending remains on premises, making the hybrid cloud a part of the IT landscape. Yet, many workloads are being moved to, or newly built in, the cloud. It will be a hybrid world for many years, given the time and effort to migrate to the cloud. Hence, many organizations will need some kind of hybrid-cloud approach to manage the transition. *We caution that hybrid is more a transition than a long-term design goal for most IT and most users.*⁸² Both cloud providers and traditional IT vendors have invested in tools and gateways to move applications and data to and from the cloud and manage it securely. Users need to locate IT resources where their applications' needs are best met. This should be driven by factors like cost, performance, proximity to data, and regulation. (Note, hybrid cloud is not the same thing as multi-cloud; see endnote.⁸³)

Multiple new edge services are emerging. Cloud solutions such as AWS Outposts and Azure Stack can address many on-premises needs. These solutions are racks of hardware, software, and networking that are largely the same as what the cloud provider uses in their own data centers. They are physically installed in the customers' data centers, but the cloud services are extended to these on-premises racks. They are logically controlled by and integrate with the cloud's management tools. They are well suited to run workloads that need to be near-to large on-premises data stores or local industrial applications that are sensitive to delay. Alternatively, Google uses a service mesh to run containers on a customer's premises. Cloud providers are also offering new "mini" availability zones, such as AWS Local Zones and Microsoft Azure Edge Zones. These don't have the full capability of classic cloud regions but extend capabilities closer to users at the edge of the network. This is important for latency-sensitive workloads, and will push cloud infrastructure to cities and places that don't have the scale to justify the \$1 billion required for a full region.⁸⁴

For smaller form factors designed for disconnected or low-connectivity environments, AWS offers its Snow family of suitcase-sized products and Azure offers Data Box. Their original purpose was to transfer large data stores to the cloud. However, Snowball Edge has evolved into a ruggedized computing platform that extends cloud devices into field operations, including oil platforms, aircraft, ships, and military operations. It does not provide the full cloud experience with a select few services, but pushes the cloud even deeper into edge workloads. With the emergence of 5G

wireless, cloud providers are exploring ways of extending cloud services to swarms of mobile, connected devices. AWS has created Wavelength Zones and Microsoft has “Azure Edge Zones With Carrier” installed in wireless carriers’ data centers. 5G devices can reach application servers running in these Zones, offload compute to the cloud, and take advantage of familiar cloud services, tools, and APIs—without leaving the carriers’ network. Cloud providers are investing apace in new IoT services and devices to capitalize on 5G networks.

Autonomous vehicles as an edge example. A self-driving car generates 4,000 gigabytes of data per day and requires split-second response times.⁸⁵ It would be prohibitively expensive to process and store all this data locally in the car, in addition to the added weight and size. The central cloud is better for this. However, cars require millisecond decisions about when to stop, accelerate, or turn. The latency (delay) of moving and processing this data to and from the cloud core is unacceptable for moving vehicles, so some of the data will need to be stored and processed locally in the car. New software architectures will evolve that process and store some data locally, while sending other data to the cloud core. ML will likely play a role, training models (e.g., how to steer a car) on large datasets in the cloud core, while algorithms will run locally executing specific decisions (e.g., “the traffic light has changed to green, accelerate now”).

The explosion of data at the edge and new edge architectures. Local edge applications and new IoT services will create new demands to process, manage, and secure data locally. New software architectures from other vendors (e.g., telecoms, industrial software providers) may emerge as alternatives or, more likely, complements to the cloud. Cloud services are already moving closer to the edge, and it will be hard to match the scale, cost, and diversity of services in the cloud. Cloud architectures will adapt but are still emerging. For customers, the key is to make choices about where applications reside, considering such factors as cost, latency, user experience, and proximity to data and compute power. Customers should strive for the same control plane, tools, and APIs, and have clear security models. At a minimum, robust integration services will be needed. Otherwise, inconsistency will create complexity and cost, and become unmanageable.

VII. A POLICY AGENDA FOR CLOUD COMPUTING

The cloud has become an economy-scale platform that drives growth, productivity, and innovation. Its increasing pervasiveness and potent capabilities are drawing interest from government policymakers and administrators. This will only grow. We outline below five key steps policymakers should take to support the development and adoption of cloud computing. While some may want less government action and others may want more, these calibrated measures address priority policy concerns in a flexible, market-driven fashion. This will likely produce better outcomes than reactive “after the crisis” responses, or heavy-handed ex ante intervention.

1. Accelerating Adoption

Governments have taken first steps to achieve cost savings and program benefits in their own operations by moving to the cloud.⁸⁶ Yet, too many still rely on older legacy systems. Indeed, the need for cloud was vividly demonstrated when unemployment insurance websites on traditional infrastructure were overwhelmed with COVID-created demand. In contrast, cloud auto-scaling is expressly designed to handle surges without service disruption. These recommendations are framed in the U.S. context, but likely apply elsewhere.

Federal Cloud Modernization Moon Shot

Federal “Cloud First” and “Smart Cloud” policies helped federal agencies start the cloud migration process.⁸⁷ Ten years later, governments should build on these policies to aggressively move to the cloud in order to improve citizen services, strengthen IT security, and lower costs.

- **Modernize federal civilian IT in a decade.** The Office of Management and Budget (OMB), the federal CIO council, and the federal CTO should lead a “moon shot” initiative to modernize as practical all federal civilian workloads over a decade, targeting at least 10 percent of these workloads a year. This would ambitiously involve migrating thousands of applications and workloads per year, with cloud as the default option.⁸⁸ Private clouds for sensitive data and on-premises modernization would remain an option for limited cases, where specifically justified and approved. Federal CIOs would need “modernization factories” and to mobilize the federal IT partner community to scale this effort.⁸⁹ This would provide a proving ground for agile, scalable migration methodologies. Admittedly, new approaches will be needed to achieve this goal. The CIO council would need to set a baseline of systems to move, and then set, measurable, agency-specific outcome goals such as the number of target workloads, servers, and petabytes of data moved, cost savings, and priority programs supported. Federal CIOs should prioritize all major systems and then define a plan to move them in smaller stages, learning along the way. This would be supported by the U.S. Digital Service and the Technology Modernization Fund.

While this goal may not be fully achievable, the decade-long Federal Data Center Optimization Initiative targeted closing ~10 percent of federal data center square footage a year with goals and gains in cost savings, server virtualization and utilization, and energy efficiency.⁹⁰ Even so, a decade is arguably too long, given the age and security risks of much of federal IT. Yet, the diversity and scale of workloads will make this a big challenge. The CIO council and a modernization moon shot implementation office should provide biannual public progress reports to spur action across federal, state, and private-sector IT, as well as drive alignment with the large federal IT community. Based on these learnings, state legislators should also mandate that state CIOs do essentially the same.

- **Leverage the budget.** The U.S. federal government is the largest technology buyer, spending well over \$100 billion a year.⁹¹ Federal, state, and local governments should use their own spending to double-down on successful initiatives, especially the federal Technology Modernization Fund. The federal Technology Modernization Fund, which provides additional funding and expertise to upgrade and transform IT systems, should be increased to \$10 billion a year, or roughly 10 percent of the federal IT budget. This would fund the federal cloud modernization moon shot. In addition, the “re-payment” requirement should be aggressively lowered for moon shot projects, with more funding for more important systems, according to criteria developed by OMB. This is a big funding increase, but would place the government at the low end of the target share of IT spending dedicated to modernization. It’s not just money. Careful governance and program support are also needed. Yet, modernization funds can help get out of the trap where appropriations only support ongoing operations, leaving few funds to move to lower-cost modern systems. Spending should prioritize domains with public impact such as IT security, health care, education, and benefit payments and fraud.

- **Reform procurement policy.** Procurement rules should be implemented neutrally so new technologies such as cloud are not penalized. Contracting officials need training on the cloud model and how to align it with procurement best practices. OMB, General Services Administration (GSA), and the CIO council need to ensure the buying process supports pay-as-you-go pricing models, not just a fixed price, one-time CapEx purchase. Procurement should also accommodate multi-year service spending, despite annual budget cycles. GSA should also consider expanding use of government-wide contracting vehicles for civilian agencies, as has been done in defense (JEDI, DEOS contracts) and national security (C2S, C2E contracts). These issues have been addressed case by case, but need more predictability, scale, and simplicity.

Spurring Adoption by the Private Sector

Government can also use its multiple policy levers and programs to encourage broader adoption by the private sector.

- **Use the “bully pulpit.”** Government should build on its history of accelerating new markets and galvanize action by highlighting the cloud’s impact on improving mission and program outcomes. A series of high-level events (e.g., by the secretary of Commerce or secretary of Defense) showing examples of cloud impact will send a strong message to organizations that may still be hesitant to adopt cloud. Indeed, CIA’s early adoption of cloud computing in 2013 led many in the private sector to consider the cloud.⁹²
- **Use standards and goals to spur adoption.** Government can drive standards and goals in its own massive IT operations, including in the federal cloud modernization moon shot. For data center energy efficiency, OMB, NIST, and DOE’s Office of Energy Efficiency should work with industry (e.g., Uptime Institute) to update ambitious data center efficiency goals—similar to public auto fuel efficiency goals.⁹³ For IT operations productivity, OMB, NIST, and the CIO Council can build benchmarks and service development frameworks. These can build on ITIL concepts, ISO standards, or create new goals and benchmarks similar to a JD Power ranking. In addition, NIST should evaluate the effectiveness of multi-cloud management tools, including the need for any standards.
- **Leverage R&D partnerships.** Robust computing is key to addressing the largest public challenges, including health care research and climate change. The federal government is the largest funder of basic research.⁹⁴ Government should leverage cloud capabilities in public research initiatives across university consortia (e.g., CloudBank), industry labs, and government research agencies.⁹⁵ Twenty-two universities have proposed creating a national research cloud.⁹⁶ Cloud should also be included in federal supercomputing programs.⁹⁷ NSF has experience leading these efforts, including addressing funding policies such as overhead charges and “life of grant” rules that can discriminate against cloud. Many technical agencies and government labs, including at NASA and in the departments of Energy, Defense, and Commerce, are also using the cloud.⁹⁸ In addition to driving research programs, R&D partnerships can develop cloud performance test beds and demonstration projects.
- **Include cloud resources in NIST’s Manufacturing Extension Program,** especially to help small and medium businesses.⁹⁹ This program should offer training resources on using cloud computing in manufacturing settings and package best-practice examples of

manufacturing companies using cloud services. It could also include technical assistance from the cloud providers to run pilot workloads.

- **Prevent cloud discrimination globally through international trade agreements.** China has multiple restrictions on foreign cloud providers. The United States should engage China to allow majority foreign ownership and enable greater market access for foreign cloud providers.¹⁰⁰ Moreover, the United States Trade Representative (USTR) should continue to push back against foreign discriminatory efforts to build domestic cloud computing champions. For example, Europe's Gaia-X project should enable full participation, including on the Board of Directors, by European subsidiaries of international companies that meet the public-stated requirements.¹⁰¹

In addition, at least seven countries have adopted digital services taxes (DST) and more are considering them.¹⁰² DSTs can discriminate against the cloud by singling out digital services for taxation and in practice can apply almost exclusively to U.S. companies. USTR should continue leveraging the Organization for Economic Cooperation and Development (OECD) for a multilateral solution to DSTs, and supplement it with bilateral agreements as needed.¹⁰³

Data localization laws are also harmful to economic prosperity, and their goals are better achieved via other methods. The United States should work with OECD, G20, and the World Trade Organization on data sovereignty to protect user data, maintain trust in the cloud, and achieve policy goals transparently (see data governance recommendations).

- **Shared data demonstration project.** To showcase the benefits of cloud data sharing, the federal government should in less than one year create a large-scale data demonstration project to create a cloud data lake based on public data. This could focus on a priority public need such as a "life sciences-health care" data cloud as part of the COVID response, including public data from the National Institutes of Health, the Food and Drug Administration, the Department of Health and Human Services, the Centers for Medicare & Medicaid Services, and others. Alternatively, it could be a geospatial data cloud with the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration, the U.S. Department of Agriculture, and the U.S. Geological Survey to develop and test responses to climate change.

2. Improving Security and Resilience: Government-Industry Collaboration Program

Cloud computing improves IT security, but also creates new challenges. As the cloud grows and becomes a more critical infrastructure that is concentrated in a handful of cloud providers, government concerns about security, resilience, and systemic risk will grow.¹⁰⁴ Given many prior cybersecurity breaches in traditional IT environments, of which Solar Winds is just the latest, government-industry cloud security efforts should intensify to address this foundational priority. President Biden issued an executive order on May 12, 2021 to improve the country's cybersecurity and protect federal networks, which took several steps in this direction.¹⁰⁵ Further actions are needed, including combining it with the proposed modernization moon shot to actually fund and modernize code and systems, in addition to the order's process improvements.

- **Cloud collaboration program: data, dialogue, and action plans.** The federal government needs to deepen its relationship with cloud providers, coordinating its multiple conversations about cloud security and resilience across different agency touchpoints in a more regular

forum.¹⁰⁶ This is a mechanism for federal agencies and cloud providers to share data, dialogue, and develop actions in ongoing and technically informed consultations. This collaboration could potentially build on the Department of Homeland Security's (DHS) IT Sector Coordinating Council, for example, in a smaller, high-level, cloud working group.¹⁰⁷

For example, DHS's Cybersecurity and Infrastructure Security Agency (CISA) and the federal CIO council should regularly review cyberattack data, service outages, compliance program effectiveness, new technologies, and other industry developments. The May 12 order helpfully instructed CISA to develop a cloud-security technical reference architecture, a cloud-service governance framework, and risk-based data classification tiers. CISA will need the cloud providers technical input to construct this. And instead of one-off reports that often languish, CISA and the major cloud providers should engage in periodic implementation reviews and course-corrections. This program could also coordinate cloud provider input in other areas such as supply chain security, threat information sharing, and cyber-breach remediation. Security best practices formed by the program should be developed into an industry code or certification, likely building on NIST's Cybersecurity Framework. Government tech modernization funds can be prioritized to support these efforts. For the private sector, cybersecurity risks are covered in corporate risk disclosures to the Securities and Exchange Commission (SEC) but are unevenly implemented. The SEC, with participation by CISA, should evaluate the effectiveness of the 2018 SEC guidance on cybersecurity risks.¹⁰⁸

- **Simplify legal and compliance.** Dozens of security legal requirements, compliance certifications, and industry frameworks have proliferated. They take too long to get protections in place, absorb limited resources, and can create vulnerability “seams” between certifications. OMB, CISA, and the U.S. CIO council should rethink and rationalize certifications built on a select core (e.g., FedRAMP, DOD SRG, CJIS). This is a big task but could build on the Cloud Security Alliance's Cloud Controls Matrix and NIST's security control framework.¹⁰⁹ The May 12 order took steps in this direction, but we need actual rationalization to reduce the overlap of frameworks.¹¹⁰ This would reduce duplication, costs, and barriers for smaller businesses. It would also increase effectiveness and speed. Compliance rules also need to be implemented in technology-neutral ways so they don't discriminate against the cloud. NIST should review security standards and make recommendations to ease this.

The key FedRAMP certification also needs to be modernized, including more reciprocal acceptance across agencies. FedRAMP needs a more automated certification process across the lifecycle from initial authorization to monitoring (starting with FedRAMP Low). This will increase speed and reduce costs, since every service and major feature for each provider must be certified for each geographic region in which it operates. The May 12 order can help but GSA's FedRAMP Office needs to provide specific implementation steps to make this actionable. See prior ITIF research for more FedRAMP actions.¹¹¹

- **Smarter tooling and automation.** Greater *integration* of different security tool formats is needed to provide security professionals a more timely and end-to-end view of their security posture. Better *automation* of security processes is needed to help security teams detect and quickly respond to security threats. Cloud is well positioned to help, given common security models are built in. To encourage private sector implementation, NIST

should make this a best practice in its cybersecurity guidelines and CISA should include this in its advisories. Further, this needs to be prioritized in procurement policies. The Federal Acquisition Regulation Council should recommend changes to achieve this, including giving higher weight to these requirements in request for proposals.

3. Preserving Competition

Regulators are increasingly concerned about tech platforms, especially in the consumer Internet context.¹¹² This includes content regulation and social media (e.g., U.S. “section 230” debate), sales of illegal goods, and providers favoring their own first-party services over third parties in search, online advertising, mobile app stores, and retail e-commerce. In enterprise infrastructure, the context is different and the actors are typically organizations with significant technical and legal sophistication—not individual consumers. Yet, cloud computing is a part of this larger debate.

- **It’s hard to justify ex ante competitive guidelines.** Cloud adoption is early (<10 percent of global IT spending) and there is little evidence of a problem. Cloud expressly provides customers with multiple approaches to lower switching costs and increase portability. Moreover, cloud provides compelling customer benefits. The proposed EU Digital Markets Act addresses many consumer Internet protection issues, yet it has very broad reach and would likely be counterproductive for enterprise cloud computing.
- **Focus on conduct issues, not structure.** Cloud computing’s early years show competition and rapid innovation. Given the high capital intensity and economies of scale, cloud markets are unlikely to have dozens of IaaS providers. The cloud industry has relatively high concentration ratios and a handful of scale providers that compete aggressively. This should not be seen as a problem in and of itself, as long as cloud providers continue to compete fairly and invest robustly. However, acquisitions that would materially increase market concentration or capture sole-source capabilities should be carefully reviewed. Regulators should focus on potential conduct issues. These could include exclusionary behavior or contracts that require a customer to use a particular cloud provider or prevent the use of other clouds. Similarly, behavior or measures that require customers to use multiple services they would not otherwise use would be a concern. For example, requiring customers who want to use ML to also use analytics services would in the absence of other customer benefits cause concern. Since the major clouds are platforms, behavior or measures that unfairly discriminate against third parties on the platform or unfairly advantage their own first-party services over third parties also need to be reviewed carefully.
- **A customer bill of rights for workload portability.** The industry has several technical groups developing portability standards, including the IEEE Computer Society on cloud portability and interoperability (P2301/P2302), the International Standards Organization (ISO/IEC 19941), and others.¹¹³ The industry should advance these efforts. For example, principles could include customers owning their own data and their applications, having the right to move them, having performant tools to move data and workloads, receiving notice and transparency about portability restrictions, and having access to rapid dispute resolution. Portability efforts should focus on lower-level IaaS services such as compute and storage that are technically more mature and common to most workloads. Higher-level PaaS and SaaS services have many more providers and much greater product

differentiation. PaaS and SaaS also have much lower entry and exit costs because providers can build on top of IaaS platforms rather than build their own. One starting point is to identify any technical or contractual restraints that limit users' ability to move workloads and lack justifiable purposes. For example, license portability allows customers to "bring your own license" they've previously purchased to a cloud or on-premises provider of their choice. Where appropriate, best practices could be stipulated in government buying rules.

4. Strengthening Data Governance: Cross-Border Data Confidence-Building Measures

Governments are increasingly interested and assertive in data governance, including data residency (also called data localization) and data sovereignty (compelled access to data by another government).

Data residency: technical confidence-building measures not "Maginot lines." Governments around the world regulate whether government data and citizen data are stored, copied, or transit outside of their geographic boundaries.¹¹⁴ Cloud can enable cross-border data flows, which are increasingly integral to the global economy. Of 29 surveyed OECD countries, 11 have some kind of data localization rules that require data be stored in-country, as do major non-OECD countries such as China, Russia, and Indonesia.¹¹⁵ Yet, it's not economically viable for each country to have its own cloud. National clouds often become expensive, legacy "white elephant" investments. This is illustrated by Microsoft Cloud Deutschland, which didn't succeed in the world's fourth-largest economy (Germany). Microsoft Cloud Deutschland required a German data trustee and was isolated from Microsoft's global network; it's now being closed in favor of Microsoft's global Azure cloud.¹¹⁶ Because many businesses operate globally, rigid data-residency rules are likely to impose a large cost on economic growth.¹¹⁷

- **Governments don't need to localize data behind their physical borders to secure their data.** The cloud provides granular controls so governments can isolate where their data is stored and who has access to it, as well as enforce and audit the controls. As a starting point, governments should drive adoption of these geographic-location and access tools. To encourage use, periodic disclosures to regulators of adoption of these tools should be required. As part of these disclosures, customers should use data-classification tiers that have increasing security requirements for data with greater sensitivity and impact.
- **Governments and cloud providers should develop technical confidence-building measures** (CBMs) demonstrating that logical software controls can be just as effective, or more so, than residency rules to keep data secure.¹¹⁸ For example, encrypted data outside a government's territory is often more secure than unencrypted data stored in its territory. Moreover, many security concerns are independent of where the data is physically stored. Security breaches often come from compromised credentials or insider threats—that are independent of location. As part of the CBMs, a joint industry-government program should develop reference designs and technical artifacts demonstrating how cloud provides data security. CBM programs promise more actual security than data localization which, like the Maginot Line in WWII, may be bypassed and may provide a false sense of security. Government agencies and cloud providers can jointly review continuous monitoring data, IAM policies, and API logs in regular consultations. *This is about*

building trust not just technology. The CBM program is integral to providing technical expertise, data, and dialog to ensure governments trust and verify cloud controls.

Data Sovereignty. Data sovereignty concerns can arise when a government compels a cloud provider to provide access to data from another country or its citizens. Data in the cloud may be subject to the laws of another country, even if the data is not in the requesting country. This challenge is not specific to the cloud, but the global reach of the cloud and its architecture can make this a bigger issue. Data is subject to the laws of other countries through traditional (though often slower) methods such as mutual legal assistance treaties and Letters Rogatory, often used in criminal law enforcement. The March 2018 U.S. CLOUD Act responded to this challenge in part, but also raised concerns.¹¹⁹ The issue presents a thorny dilemma with no perfect answer. Mechanisms that can help include 1) disclosing government requests so they can be addressed in direct talks between the parties; 2) developing strong criteria that must be met before the cloud provider turns over data, including an appeal mechanism; 3) creating bilateral or multi-country data access agreements such as the new U.S.-U.K. data sharing agreement; and 4) using techniques such as encryption or tokenization to mitigate these concerns.

5. Promoting Skills and Inclusivity via a Public-Private Training Partnership

Cloud computing presents an opportunity to train a broader set of people with new skills so that the benefits of technology growth are more widely shared. Communities that are not well represented in technology need greater access and participation. These are imperatives from a policy perspective and will help achieve the full public benefits of the cloud. New technologies often create and destroy jobs, changing the pattern of employment. Public-private cloud training partnerships can and should do more to help address these needs. There is no broad-based national effort to address these needs. While cloud providers have employee training and public outreach programs, they are fragmented. New approaches and investment are required to make training more scalable, effective, and inclusive. At the same time, private education and training institutions are looking for new, more economically sustainable models and can play a greater role. The COVID-19 pandemic has made this even more urgent.

The time is ripe for a national initiative on a public-private technology training partnership, accompanied by a G-7 summit with funding commitments to training, re-skilling, and inclusivity. This should include state and local governments and incorporate community colleges and apprenticeships, not just four-year universities. Prior ITIF research makes the case for a strong “re-employment system.”¹²⁰ Cloud training could be combined with other priority areas such as ML and cybersecurity. Alternative IT certifications from institutions beyond four-year colleges could be part of this program. For example, Northern Virginia Community College offers a Cloud Computing degree, the Louisiana Community and Technical College System is creating an associate degree in cloud computing, and the Career Colleges Trust in the United Kingdom is developing a two-year course in cloud computing. The cloud providers can contribute funding, help develop curricula, and improve access to jobs. Training and re-skilling are integral to more broadly sharing the benefits of cloud computing and proactively including all communities in the cloud and technology industry.

Acknowledgments

Thank you to Rob Atkinson and ITIF for focusing on the importance of cloud computing and their leadership on technology issues. Thank you to my many colleagues who provided insights and feedback. Any errors or omissions are the author's responsibility.

About the Author

Bill Whyman is a leading expert on the dynamics of the technology sector. He is known for market-tested expertise on technology industry competition, business-model change, and technology policy. Whyman brings a distinctive perspective combining technology industry experience (Amazon AWS, McKinsey & Co), government policy (White House, State Department), and financial markets (Evercore ISI, Legg Mason).

About ITIF

The Information Technology and Innovation Foundation (ITIF) is an independent, nonprofit, nonpartisan research and educational institute focusing on the intersection of technological innovation and public policy. Recognized by its peers in the think tank community as the global center of excellence for science and technology policy, ITIF's mission is to formulate and promote policy solutions that accelerate innovation and boost productivity to spur growth, opportunity, and progress.

For more information, visit us at www.itif.org.

ENDNOTES

1. “Gartner Forecasts Worldwide Public Cloud End-User Spending to Grow 23 percent in 2021,” press release, April 21, 2021. <https://www.gartner.com/en/newsroom/press-releases/2021-04-21-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-grow-23-percent-in-2021>.
2. For start-up data, see VC firm Bessemer Capital’s review. Bessemer Venture Partners, State of the Cloud 2021, https://www.bvp.com/atlas/state-of-the-cloud-2021#page_top.
3. Author’s calculation from Gartner data. Cloud spending market share includes SaaS, PaaS, and IaaS, plus Gartner includes BPaaS or “Business Process as a Service.” Over time, many services (if not all) will fit this BPaaS definition as “software eats the world” and digitization of the economy increases. Further comments on BPaaS are in endnote 29. Total IT spending includes the >\$1 trillion a year spent on IT services. “Gartner Forecasts Worldwide Public Cloud End-User Spending to Grow 23 percent in 2021,” April 21, 2021; “Gartner Forecasts Worldwide IT Spending to Reach \$4 Trillion in 2021,” April 7, 2021.
4. The National Institute of Standards and Technology (NIST) is part of the U.S. Commerce Department, Peter Mell, “The NIST Definition of Cloud Computing,” NIST, September 2011, <https://csrc.nist.gov/publications/detail/sp/800-145/final>.
5. The table captures the vast majority of typical workloads. There are always nuances and border cases. For a primer on what is cloud, see, for example, the Atlantic Council. Simon Handler, Lily Liu, and Trey Herr, “Dude, where’s my cloud? A guide for wonks and users,” Atlantic Council, September 28, 2020, <https://www.atlanticcouncil.org/cloud-security-amp-policy/>.
6. For example, some customers will start by shifting individual applications to SaaS such as email and calendaring to Microsoft 365, Expense management to SAP Concur, or customer relationship software to Salesforce.com. Others will adopt cloud storage such as Box. Some businesses will keep their applications on premises but use the cloud IaaS for back-up and recovery. This reflects different customer segments that have different priorities, needs, and skill levels. We discuss this in the section on “multiple market segments.”
7. See Rolf Harms and Michael Yamartino, “The Economics of the Cloud,” white paper, Microsoft, November 2010, <https://news.microsoft.com/download/archived/presskits/cloud/docs/The-Economics-of-the-Cloud.pdf>. There are many variations of this chart showing the differences between IaaS, PaaS, and SaaS.
8. Private clouds use similar technologies such as virtualization and containers. However, use of private clouds is “fading” as users realize they often don’t have the scale, economics, innovation path, or skills of public clouds—and they can effectively create their own software-defined private clouds on a public cloud infrastructure.
9. PaaS is smaller in value terms than IaaS or SaaS. Yet, PaaS is technically important, helping enable and manage the application—but it is not the application itself. While technically distinct, PaaS is in many ways becoming part of the larger IaaS market as major cloud providers offer “managed services” that integrate both IaaS and PaaS into a single service.
10. This is a more complicated topic. For example, many SaaS providers such as Salesforce also offer PaaS platforms, such as Force.com. However, these were originally used to customize and support their own proprietary SaaS applications. Relative to other PaaS, they are typically not as broadly used by the developer community to support many applications. This paper does not focus on the SaaS model, given these major differences and the expansive issues specific to SaaS.
11. Business investment data is from the Department of Commerce, BEA. NIPA table 2.7. 2019 annual investment in private fixed assets. IT is defined as non-residential Information Processing Equipment plus software. The budgeted U.S. government IT spending is \$92 billion and this excludes billions of dollars in intelligence community spending and mission IT.

12. These technology architectures at a high level address similar problems and needs, but do so in new ways with lower costs and more capability. Over time, IT “encapsulates” the prior low-level functionality, and pushes it “down the stack.” I think of this as “technology sedimentation.” This shields the user from lower-level complexity that IT can itself manage more efficiently, and moves the technical “programming interfaces” (e.g., APIs) and economic value higher up the stack. See also David C. Moschella, *Waves of Power, The Dynamics of Global Technology Leadership, 1964-2010*, (AMACOM, February 26, 1997).
13. Carlota Perez, *Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages* (United Kingdom: Edward Elgar Publishing Ltd, 2010); Nicholas Carr, *The Big Switch: Rewiring the World from Edison to Google* (W.W. Norton, 2008). For a more pessimistic view, see Robert Gordon, “Is U.S. Economic Growth Over? Faltering Innovation Confronts The Six Headwinds,” NBER, Aug 2012, <https://www.nber.org/papers/w18315>. For a more optimistic view, see Lee Branstetter and Daniel Sichel, “The Case for an American Productivity Revival,” Peterson Institute for International Economics, June 2017, <https://www.piie.com/publications/policy-briefs/case-american-productivity-revival>.
14. IDC paper sponsored by AWS, “Fostering Business and Organizational Transformation to Generate Business Value with Amazon Web Services,” February 2018. It’s important to have an “apples to apples” cost comparison since the operating models are different. In the cloud model, costs such as data center space and electricity are included. These costs are sometimes excluded from traditional IT cost models if these costs are allocated to the “facilities” line item of a different department, in accounting systems. Economists have started to look at the economic impact of the cloud. See F. Etro, “The Economic Consequences of the Diffusion of Cloud Computing,” in S. Dutta and I. Mia (eds.), *The Global Information Technology Report 2009-2010*, World Economic Forum, Geneva. (2010).
15. David Byrne, Carol Corrado, and Daniel E. Sichel, “The Rise of Cloud Computing: Minding Your P’s, Q’s and K’s,” NBER Working Papers 25188, National Bureau of Economic Research, Inc., 2018.
16. IDC, “Fostering Business and Organizational Transformation to Generate Business Value with Amazon Web Services,” February 2018. See also Dave McCann, AWS Vice President, “re:Invent,” 2020, <https://reinvent.awsevents.com/>.
17. The average power usage effectiveness (PUE) of Google data centers in 2019 was 1.10 versus the industry average of 1.67. Urs Hölzle, “Data centers are more energy efficient than ever.” The Keyword, Google, <https://blog.google/outreach-initiatives/sustainability/data-centers-energy-efficient/>. A 451 Research Group study commissioned by AWS shows Amazon data centers are 3.6x times more efficient and perform the same operations with 88 percent less carbon footprint than median enterprise data centers. Daniel Bizo, “The Carbon Reduction Opportunity of Moving to Amazon Web Services,” AWS, <https://d39w7f4ix9f5s9.cloudfront.net/e3/79/42bf75c94c279c67d777f002051f/carbon-reduction-opportunity-of-moving-to-aws.pdf>. See Uptime Institute for more.
18. Colin Cunliff, “Beyond the Energy Techlash: The Real Climate Impacts of Information Technology” (ITIF, July 2020), <https://itif.org/publications/2020/07/06/beyond-energy-techlash-real-climate-impacts-information-technology>.
19. For example, Mary K. Pratt, “Why IT projects still fail,” *CIO*, March 3, 2021, <https://www.cio.com/article/3211485/why-it-projects-still-fail.html>; Michael Kringsman, “10 reasons for IT failure,” *ZDNet*, May 2, 2008, <https://www.zdnet.com/article/10-reasons-for-it-failure/>; Stan Maklan, Knox Simon, and Joe Peppard, “Why CRM Fails-and How to Fix It,” *Harvard Business Review*, July 1, 2011, <https://store.hbr.org/product/why-crm-fails-and-how-to-fix-it/smr392>.
20. This was an ERP application. These are mission-critical enterprise applications that manage systems of record. For HP, see “Learning from failure: 3 ERP disasters every adopter should understand,” Accent Software Inc, June 28, 2017, <https://accenterp.com/manufacturing/learning-from-failure-3->

- erp-disasters-every-adopter-should-understand/; Patrick Thibodeau and Don Tennant, “HP's CIO Points to Internal Issues in ERP Project Snafus,” *Computerworld*, September 27, 2004, <https://www.computerworld.com/article/2565928/hp-s-cio-points-to-internal-issues-in-erp-project-snafus.html>; Josh Fruhlinger, Thomas Wailgum and Peter Sayer, “16 famous ERP disasters, dustups and disappointments,” *CIO*, March 20, 2020, <https://www.cio.com/article/2429865/enterprise-resource-planning-10-famous-erp-disasters-dustups-and-disappointments.html>.
21. For example, Brianna Ehley, “Another Failed Gov’t Tech Project Cost \$1.1 Billion,” *The Fiscal Times*, November 26, 2013, <https://www.thefiscaltimes.com/Articles/2013/11/26/Another-Failed-Gov-t-Tech-Project-Cost-11-Billion>; Richard Sisk, “VA's \$16 Billion Electronic Health Records Modernization Plan Is Failing, IG Says,” *military.com*, April 28, 2020, <https://www.military.com/daily-news/2020/04/28/vas-16-billion-electronic-health-records-modernization-plan-failing-ig-says.html>.
 22. Specialization and division of labor are famously developed by no less than Adam Smith in the *Wealth of Nations* (1776) and David Ricardo in, *Principles of Political Economy and Taxation* (1817).
 23. Kim Nash, “J.P. Morgan Chase Names New CIO as Dana Deasy Exits,” *Wall Street Journal*, September 7, 2017, <https://www.wsj.com/articles/j-p-morgan-chase-names-new-cio-as-dana-deasy-exits-1504822667#:~:text=J.P.%20Morgan%20Chase%20%26%20Co.,by%20the%20Wall%20Street%20Journal>.
 24. Described by Nicholas Carr in *The Big Switch: Rewiring the World from Edison to Google* (W.W. Norton, 2008).
 25. Manufacturers used machines powered by electric motors and connected them to the electric grid. See also Paul A. David, “The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox,” *American Economic Review*, Vol. 80, No. 2, 355–361, 1990.
 26. F. Etro, “The Economic Consequences of the Diffusion of Cloud Computing,” in S. Dutta and I. Mia (eds.), *The Global Information Technology Report 2009-2010*, World Economic Forum, Geneva, (2010).
 27. As an example of the growing economic importance of the cloud, a study by insurer Lloyds of London and risk modeler AIR Worldwide estimated that a cyber incident that took a top-three cloud provider offline in the United States for three to six days would result in ground-up loss estimates between \$6.9 billion and \$14.7 billion. “Cloud Down. Impacts on the US Economy,” AIR Worldwide, 2018
 28. AWS re:Invent, Andy Jassy Keynote. December 11, 2020. For Gartner data, see footnote 2.
 29. “Annual Business Survey: Tables for Data Year 2017,” NSF, November 24, 2020, <https://nces.nsf.gov/pubs/nsf21303#general-notes&>. 24.3 percent said they had not purchased cloud, half said they didn’t know or collect this data, and 8.5 percent of businesses said they had purchased cloud for up to half their workload. Adoption of advanced technologies such as ML show 80 percent have no adoption and 1 percent have ML in production. Unfortunately, more recent data is not available from NSF.
 30. Gartner, “Forecasts Worldwide Public Cloud End-User Spending to Grow 23 percent in 2021,” press release, April 21, 2021, <https://www.gartner.com/en/newsroom/press-releases/2021-04-21-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-grow-23-percent-in-2021>. In the spending chart, management and security are included in the IaaS section. I have excluded Business Process as a Service (BPaaS). Business processes are specific to industry domains and industry-specific workflows. Over time, much of today’s manual workflows will be embedded in or assisted by software, as “software eats the world.” Hence, BPaaS will become a very expansive category that moves beyond IT to “business.” Many, many things will become BPaaS in the fullness of time. When technology fully succeeds it “disappears,” meaning, for example, that e-commerce just becomes

commerce, and e-payments just become payments. In a fully digital economy, BPaaS will simply become business processes.

31. Synergy Research Group, "The Decades Megatrends in Numbers – Part 2," January 7, 2020, <https://www.srgresearch.com/articles/the-decades-megatrends-in-numbers-part-2>.
32. I do not strictly include companies such as Facebook or Twitter for the purposes of this article in large part because they are consumer-focused companies. They operate their own cloud infrastructures and purchase cloud services from AWS, Azure, and Google. They are based on similar technologies, but to enable their own first party services. I consider them closer to SaaS companies, as they are providing an end-to-end application to the customer or end user.
33. The median number of AWS cloud services used by enterprises is roughly 10 percent of the more than 200 total available, according to a survey by cloud intelligence company Sumo Logic. "The Continuous Intelligence Report," 2020, <https://www.sumologic.com/brief/continuous-intelligence-report/>.
34. "Gartner Says Worldwide IaaS Public Cloud Service Market Grew 37.3% in 2019," August 10, 2020, <https://www.gartner.com/en/newsroom/press-releases/2020-08-10-gartner-says-worldwide-iaas-public-cloud-services-market-grew-37-point-3-percent-in-2019>; "Gartner Says Worldwide IaaS Public Cloud Services Market Grew 31.3% in 2018," July 29, 2019, <https://www.gartner.com/en/newsroom/press-releases/2019-07-29-gartner-says-worldwide-iaas-public-cloud-services-market-grew-31point3-percent-in-2018>; "Gartner Says Worldwide IaaS Public Cloud Services Market Grew 29.5 Percent in 2017," August 1, 2018, <https://www.gartner.com/en/newsroom/press-releases/2018-08-01-gartner-says-worldwide-iaas-public-cloud-services-market-grew-30-percent-in-2017>; "Gartner Says Worldwide IaaS Public Cloud Services Market Grew 31 Percent in 2016," September 27, 2017, <https://www.gartner.com/en/newsroom/press-releases/2017-09-27-gartner-says-worldwide-iaas-public-cloud-services-market-grew-31-percent-in-2016>.
35. IDC, "Spending on Public Cloud IT Infrastructure Surpasses Spending on Traditional IT infrastructure for the First Time in the Second Quarter of 2020," September 29, 2020, <https://www.idc.com/getdoc.jsp?containerId=prUS46895020#:~:text=29%20Sep%202020-,Spending%20on%20Public%20Cloud%20IT%20Infrastructure%20Surpasses%20Spending%20on%20Traditional,of%202020%2C%20According%20to%20IDC&text=Investments%20in%20traditional%2C%20non%2Dcloud,year%20over%20year%20in%202Q20..> While cloud providers have large fixed cost. Cloud computing converts them into variable costs for customers.
36. Microsoft and AWS, the two largest, are racing to build global infrastructure and global leadership. This is also a marketing battle and both claim to have the most infrastructure. They use different definitions which makes a like-for-like comparison difficult. AWS has 24 geographic regions with 77 availability zones, where each availability zone is 1 or more data centers and there can be many. <https://aws.amazon.com/about-aws/global-infrastructure/>; Microsoft has more than 60+ regions with 160 physical data centers. <https://azure.microsoft.com/en-us/global-infrastructure/>.
37. For example, Microsoft's new regions in Italy and Poland, and AWS's new region in Bahrain.
38. See IDC, "Worldwide Server Market Revenue Grew 2.2 percent Year Over Year in the Third Quarter of 2020," December 9, 2020. <https://www.idc.com/getdoc.jsp?containerId=prUS46895020>.
39. From Rolf Harms and Michael Yamartino "The Economics of the Cloud," Microsoft white paper, November 2010, <https://news.microsoft.com/download/archived/presskits/cloud/docs/The-Economics-of-the-Cloud.pdf>.
40. David Byrne, Carol Corrado, and Daniel E. Sichel, "The Rise of Cloud Computing: Minding Your P's, Q's and K's," NBER Working Papers 25188, National Bureau of Economic Research, Inc., 2018, <https://www.nber.org/papers/w25188>. See also forthcoming volume published by the University of Chicago Press: Carol Corrado et al., "Measuring and Accounting for Innovation in the Twenty-First

- Century,” NBER, April 2021, <https://www.nber.org/books-and-chapters/measuring-and-accounting-innovation-twenty-first-century>.
41. Calculation from author. Data from the companies.
 42. Quality-adjusted price indices of public AWS prices from Byrne, Corrado, and Sichel. Author’s presentation.
 43. Synergy market share release. For Q4 2020, see February 2, 2021, “Cloud Market Ends 2020 on a High while Microsoft Continues to Gain Ground on Amazon”. Canalys uses a similar definition as Synergy and their data corroborates Synergy. See February 2, 2021, “Global cloud services market surges by US\$10 billion in Q4 2020.” Gartner data measuring only the IaaS market shows a similar picture by ranking and change, but has more concentration with the top three at 72 percent share, driven by a larger share for Amazon. Drawing the market definition more tightly raises concentration and the HHI. See “Gartner says Worldwide IaaS Public Cloud Market grew 37.3 percent in 2019.” These market share statistics are static views that don’t capture forward change.
 44. DOJ HHI guidelines, “HERFINDAHL-HIRSCHMAN INDEX,” Department of Justice, <https://www.justice.gov/atr/herfindahl-hirschman-index>. Author’s estimate based on publicly available market share data.
 45. Gartner magic quadrant at Laura Shiff, “The 2020 Gartner Magic Quadrant for Cloud Infrastructure and Platform Services,” September 17, 2020, <https://www.bmc.com/blogs/gartner-magic-quadrant-cips-cloud-infrastructure-platform-services/>.
 46. Europe Gaia project, <https://www.data-infrastructure.eu/GAIA/Navigation/EN/Home/home.html>.
 47. Synergy Research Group, “Amazon and Microsoft Maintain their Grip on the Market but Others are also Growing Rapidly,” press release, April 29, 2021, <https://www.srgresearch.com/articles/amazon-and-microsoft-maintain-their-grip-market-others-are-also-growing-rapidly>.
 48. “Accenture Cloud First Launches with \$3 Billion Investment to Accelerate Clients’ Move to Cloud and Digital Transformation,” press release, September 17, 2020, <https://newsroom.accenture.com/news/accenture-cloud-first-launches-with-3-billion-investment-to-accelerate-clients-move-to-cloud-and-digital-transformation.htm>.
 49. Doug Yeum, AWS, Head of Global Partner Organization, and Dave McCann AWS VP, re:Invent Partner Keynote, December 18, 2020.
 50. Byron Deeter, “State of the Cloud 2020,” Bessemer Venture Partners, April 22, 2020, <https://www.bvp.com/atlas/state-of-the-cloud-2020>. The updated 2021 Bessemer report says “\$186 billion (is) going into private cloud companies in 2020 alone,” and there was, “record breaking cloud M&A activity in 2020.” See <https://www.bvp.com/atlas/state-of-the-cloud-2021>
 51. Cloud Native Computing Foundation, <https://www.cncf.io/>.
 52. Operating margins for the five largest traditional software companies, Microsoft, Oracle, IBM, SAP, and Adobe, average 30 percent in their recent fiscal year (2019 or 2020). Operating margins for AWS in Calendar 2020 were 29.8 percent.
 53. Commercial Cloud business is Azure plus Microsoft’s SaaS products. This assumes that Microsoft’s Commercial Cloud business has the same operating cost structure (R&D, sales and marketing, general & administrative) as the rest of Microsoft.
 54. Google recently separated out its Google Cloud business for the first time in its February 2, 2021, earnings release. It showed losses with -42.9 percent operating margin in 2020 and -32.4 percent in Q4 2020, but has improved over the last three years. Losses reflect that Google has not yet achieved the scale for profitability and is making huge investments as the number-three global player to compete with AWS and Microsoft Azure.
 55. Dorion Carroll, “Zynga’s Journey (Back) to the AWS Cloud,” presentation delivered at AWS re:Invent 2015 (October 12, 2015), <https://www.youtube.com/watch?v=BZfhhGvY3c0>.

56. IDC paper sponsored by AWS, Larry Carvalho and Matthew Marden, “Fostering Business and Organizational Transformation to Generate Business Value with Amazon Web Services,” IDC, February 2018, <https://pages.awscloud.com/rs/112-TZM-766/images/AWS-BV%20IDC%202018.pdf>.
57. “Making the Cloud Pay: How Industrial Companies Can Accelerate Impact From The Cloud,” McKinsey & Company, October, 2020, <https://www.mckinsey.com/industries/advanced-electronics/our-insights/making-the-cloud-pay-how-industrial-companies-can-accelerate-impact-from-the-cloud#>.
58. Cloud Customer public references for AWS, Microsoft, Google, IBM, Oracle:
AWS: <https://aws.amazon.com/solutions/case-studies/?customer-references-cards.sort-by=item.additionalFields.publishedDate&customer-references-cards.sort-order=desc>;
Microsoft Azure: <https://azure.microsoft.com/en-us/case-studies/>; Google Cloud: <https://cloud.google.com/customers>; IBM cloud: <https://www.ibm.com/cloud/case-studies/>
Oracle Cloud: <https://www.oracle.com/customers/>.
59. Carvalho and Marden, “Fostering Business and Organizational Transformation.”
60. AWS architecture here. <https://aws.amazon.com/about-aws/global-infrastructure/>. Microsoft has also started to adopt this availability zone architecture.
61. AWS, Microsoft Azure, and Google all provider global coverage of the most of the global economy.
Google Cloud: <https://cloud.google.com/about/locations/>.
AWS: <https://docs.aws.amazon.com/whitepapers/latest/aws-overview/global-infrastructure.html>
Microsoft Azure: latest disclosed number of countries is “over 140” in October 23, 2018. With new regions, the number is now even larger and likely comparable to AWS and Google.
<https://azure.microsoft.com/en-us/blog/azure-availability-zones-expand-with-new-services-and-to-new-regions-in-europe-and-united-states/>.
62. See for example, “Sharing Data on AWS,” the section on “Putting Data To Work: Uses Cases and Insights.” Undated.
63. On How APIs increase firm performance and innovation. Seth G. Benzell, Guillermo Lagarda, and Marshall Van Alstyne, “The Impact of APIs in Firm Performance,” May 21, 2017, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2843326. Perhaps, even more fascinating is the reported history of APIs and service architectures at Amazon driven by Jeff Bezos. “Stevey’s Google Platforms Rant,” GitHub Gist, <https://gist.github.com/chitchcock/1281611>.
64. Cloud providers typically charge “egress” costs for data leaving the cloud but not for “ingress” or moving data into the cloud. The egress costs can be managed in multiple ways. First, customers typically save costs over the life of an application (31 percent operating cost savings) that more than pay for the costs to move. In addition, “gateway” appliances can be used to minimize the data sent to the cloud and keep local copies on premises. In addition, there are multiple other tools and services to move data and applications from the cloud and track and manage costs. Some customers retain a “gold copy” of their data on premises and delete the encryption keys of their encrypted data in the cloud so they don’t actually need to move their data to exit and avoid the cost.
65. See, for example, online learning company’s “Desire to Learn” view on this. They calculate the lower costs of the cloud versus the costs of switching and use the cloud when there is a positive return on investment. They also established a rule not to invest in their own IT capabilities and to instead use the cloud when the function is not related to their customer or differentiating. Video is here. <https://www.youtube.com/watch?v=kh8TxYbIEik>.
66. “Flexera 2021 State of the Cloud Report,” Flexera, <https://info.flexera.com/CM-REPORT-State-of-the-Cloud>.
67. All the major clouds have publicly available pricing calculators. See “AWS Pricing Calculator Estimate the cost for your architecture solution,” AWS, <https://calculator.aws/#/>.

68. Andy Jassy, AWS CEO, re:Invent keynote. December 1, 2020.
69. The AWS Well Architected Framework and Tool is a good example of a response to these issues. Yet, it still requires substantial cloud technical knowledge and becomes a large corpus of content. It could also be a starting point for a simpler and cross-industry approach. See “AWS Well-Architected,” AWS, <https://aws.amazon.com/architecture/well-architected/?ref=wellarchitected-wp&wa-lens-whitepapers.sort-by=item.additionalFields.sortDate&wa-lens-whitepapers.sort-order=desc>
70. According to CIA CIO John Edwards, on moving to their own private cloud that "it's the best decision we've ever made ... It's the most innovative thing we've ever done," June 14, 2017, public speech. Troy K. Schneider, “CIA CIO: Private cloud 'the best decision we've ever made'," *FCW*, June 14, 2017, <https://fcw.com/articles/2017/06/14/cia-cloud-aws.aspx>. Also see Kasey Panetta, “Is the Cloud Secure?” Smarter With Gartner, October 10, 2019, <https://www.gartner.com/smarterwithgartner/is-the-cloud-secure/>.
71. “Azure compliance,” Microsoft Azure, <https://azure.microsoft.com/en-us/overview/trusted-cloud/compliance/>.
72. The shared security model highlights the “supply chain” security issues in the cloud. This is not just the hardware supply chain security of servers, storage, and networking components, but requirements of “outsourcers” that hold and process customer data on behalf of another. On the hardware supply chain, see Trey Herr, “Four Myths of Cloud: The Geopolitics of Cloud Computing,” The Atlantic Council, August 21, 2020, <https://www.atlanticcouncil.org/wp-content/uploads/2020/08/CLOUD-MYTHS.pdf>.
73. “Flexera 2021 State of the Cloud Report.”
74. Ibid.
75. RTO is Recovery Time Objective, meaning how long it takes to restore a system. RPO is Recovery Point Objective, meaning the maximum time that a system was last updated or, said differently, the time from the loss event to the most recent preceding back-up.
76. 451 Research, “Cloud Trends in 2020: The Year of Complexity and its Management,” <https://go.451research.com/2020-mi-cloud-trends-year-of-complexity-and-its-management.html>; Melanie Posey, “Cloud Maturity and Cloud Skills: Moving Up the Learning,” S&P Global Market Intelligence, March 2, 2021, <https://www.spglobal.com/marketintelligence/en/news-insights/blog/cloud-maturity-and-cloud-skills-moving-up-the-learning-curve>.
77. Brad Smith, “Microsoft launches initiative to help 25 million people worldwide acquire the digital skills needed in a COVID-19 economy,” Microsoft, June 20, 2020, <https://blogs.microsoft.com/blog/2020/06/30/microsoft-launches-initiative-to-help-25-million-people-worldwide-acquire-the-digital-skills-needed-in-a-covid-19-economy/>; Teresa Carlson, “By 2025 Amazon Web Services (AWS) will help 29 million people globally grow their technical skills with free cloud computing skills training,” Amazon.com, December 10, 2020, <https://www.aboutamazon.com/news/workplace/amazon-to-help-29-million-people-around-the-world-grow-their-tech-skills-with-free-cloud-computing-skills-training-by-2025>.
78. For example, see economist Laurence Katz and Robert Margo, “Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective,” NBER, October 2014, <https://www.nber.org/papers/w18752>.
79. See Robert D. Atkinson, “How to Reform Worker-Training and Adjustment policies for an Era of Technical Change,” ITIF, February 2018. <https://itif.org/publications/2018/02/20/technological-innovation-employment-and-workforce-adjustment-policies>.
80. “Flexera 2021 State of the Cloud Report.”
81. To take full advantage of the cloud, customers need to know their usage patterns and performance requirements—how much of what IT resource (e.g., database, storage) is used by what applications,

where, and when. They can then optimize their cloud spending by “right sizing” compute-instance types, shifting to discounted “reserved instance” pricing, and moving infrequently accessed data to lower-cost performance tiers. However, many customers do not know this information initially. This can disadvantage the cloud in the purchasing decision, compared with a typical on-premises capital expenditure wherein customers can simply buy for peak-demand growth plus a margin of error. This CapEx purchasing model can be simpler and easier for the buyer, but leads to inefficiency and extra cost. In addition, when comparing cloud and on-premises costs, customers need to include on-premises costs for data center space, electricity, power and cooling, and data center manpower, which may not be in their budget even though these costs are paid for by another unit’s budget in their organization.

82. The history of IT shows that customer owned and operated IT creates strong pressures to customize IT, since IT leaders work for the business and IT’s purpose is to support the business. And that has led to expensive, brittle, complex custom deployments that are hard to manage and modernize. More fundamentally, customization aside, on-premises deployments lack the cloud’s scale, automation, and economics, and draw the business back into areas that are not their core competency. See section V in “Where Cloud Computing Excels” for more.
83. Hybrid cloud overlaps with, but is not the same as the multi-cloud. There are no formally agreed-upon definitions, but multi-cloud typically refers to using multiple public cloud services such as Azure, AWS, and Google Cloud. Hybrid typically refers to using both traditional on-premises IT and public cloud providers such as Azure. This distinction may blur over time as the multi-cloud management tools also aspire to provide visibility and management to on-premises IT resources.
84. In his 2019 re:Invent keynote, AWS CEO Andy Jassy noted that while AWS had built many regions (composed of multiple data centers) with many more coming, they were very expensive and he did not necessarily envision 200–300 more in light of the cost.
85. Intel CEO Brian Krzanich, “Data is the New Oil in the Future of Automated Driving,” Intel, November 15, 2016, <https://newsroom.intel.com/editorials/krzanich-the-future-of-automated-driving/#gs.y1zazu>.
86. GAO, “Cloud Computing: Agencies Have Increased Usage and Realized Benefits, but Cost and Savings Data Need to Be Better Tracked,” April 2019, <https://www.gao.gov/assets/700/698236.pdf>.
87. “From Cloud First to Cloud Smart,” cloud.cio.gov, <https://cloud.cio.gov/strategy/>. They also contain helpful steps for IT security and federal IT workforce training.
88. The modernization campaign would support the spectrum of migration options including “lifting and shifting” workloads to the cloud as well as more complicated and effective modernization approaches of re-architecting and re-factoring applications. While the default migration path would be to the public cloud, the CIO council should develop criteria that allow for private clouds (e.g., sensitive workloads) and on-premises deployments where specifically justified and approved. The 10 percent goal would be an annual average, as modernization would likely be slower in the first years as the program is set up and best practices are captured. Like the federal Data Center Optimization initiative, the goals, implementation approaches, and success metrics are likely to evolve with experience.
89. Modernization factories are techniques and methodologies for migrating large numbers of applications in more standardized repeatable ways. See for example, “The rise of cloud application migration factories,” DXC Technology, <https://www.dxc.technology/cloud/insights/147246-the-rise-of-cloud-application-migration-factories> or <https://techbeacon.com/enterprise-it/how-build-your-cloud-app-migration-factory>.
90. The initiative targeted 31 percent reduction over a ~three-year period. See August 1, 2016 “Memorandum for Heads of Executive Departments and Agencies” from Federal CIO Tony Scott on Data Center Optimization Initiative; and June 25, 2019, update Memorandum from Federal CIO

Suzette Kent. The Federal Data Center Optimization Initiative evolved with learnings and spanned multiple administrations.

91. Budgeted spending is \$92 billion, which excludes the many billions of dollars spent by the Intelligence community. This also does not fully capture the very large IT spending embedded in “mission” spending. <https://www.govinfo.gov/content/pkg/BUDGET-2021-PER/pdf/BUDGET-2021-PER.pdf>
92. Frank R. Konkell, “Amazon is building a cloud for the intelligence community under a plan to upend the status quo.” Nextgov, July 9, 2014, <https://www.nextgov.com/it-modernization/2014/07/daring-deal/100872/>.
93. OMB and the CIO council already set data center efficiency goals for Federal data centers. See “Data Center Optimization Initiative,” datacenters.cio.gov, June 25, 2019, <https://datacenters.cio.gov/policy/>.
94. See National Science Board, Science and Engineering Indicators, 2020. <https://nces.nsf.gov/pubs/nsb20201/global-science-and-technology-capabilities>.
95. For example: “NSF Cloud and Autonomic Computing (CAC) Industry-University Cooperative Research Center,” CAC, <https://nsfcac.org/> and <https://iucrc.nsf.gov/> on partnerships; “Enabling Access to Cloud Computing Resources for CISE Research and Education (Cloud Access),” NSF, https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505591. Government funding and contracting issues (CapEx versus OpEx, life of research grants rules that limit funding data stored in the cloud after the grant period ends, overhead charges, etc.) need to ensure they don’t discriminate against cloud.
96. “National Research Cloud Call To Action,” HAI, Stanford University, <https://hai.stanford.edu/national-research-cloud-joint-letter>.
97. See ITIF’s prior work on expanding access to HPC resources, including from the cloud: Hodan Omaar, “How the United States Can Increase Access to Supercomputing,” Center for Data Innovation, December 7, 2020, <https://datainnovation.org/2020/12/how-the-united-states-can-increase-access-to-supercomputing/>.
98. For examples: “Leadership in Cloud Computing,” DOE, November 2012, <https://www.nrel.gov/docs/fy13osti/56143.pdf>; “Cloud Computing,” NASA, <https://www.nccs.nasa.gov/services/cloud-computing>; NOAA website, [https://nrc.noaa.gov/Portals/0/NOAA percent20Cloud percent20Strategy percent20Fact percent20Sheet.pdf?ver=2020-02-26-145239-740](https://nrc.noaa.gov/Portals/0/NOAA%20percent20Cloud%20Strategy%20Fact%20Sheet.pdf?ver=2020-02-26-145239-740).
99. Hollings Manufacturing Extension Partnership Programs. See “Manufacturing Extension Partnership (MEP),” NIST, <https://www.nist.gov/mep> and <https://fas.org/sgp/crs/misc/R44308.pdf>.
100. Chinese support for its cloud providers and their hardware and software supply chain is a large and important subject that should be addressed in its own right. The key point is that it is a SYSTEMATIC effort encompassing dozens of policies such as taxes, subsidies, training, IPR, standards, government procurement, product licensing, and foreign access to Chinese markets. In short, it is woven into the very sinews of the marketplace making it difficult to address, especially in a piecemeal fashion. See Bill Whyman presentation at the “4th Annual NYU Conference on China Capital Markets,” <https://www.youtube.com/watch?v=38Etr7sdSgw>.
101. See for example “Response to the European Commission’s Consultation on the European Strategy for Data,” Center for Data Innovation, <https://www2.datainnovation.org/2020-eu-data-strategy.pdf>. The GAIA-X website states, “Members of GAIA-X AISBL are required to uphold the European values of enhanced data privacy, transparency, security and respect for data rights. In the interest of guaranteeing a European direction and destiny for the project, the association’s Board of Directors will consist entirely of Europe-based companies.” It’s not fully clear what qualifies as a Europe-based company for GAIA-X, and it needs to be more clearly defined in public. Similarly, “European

values” needs to be more clearly defined. European subsidiaries of international companies that are fully bound by EU law should be able to fully participate.

102. “Fact Sheet on the 2020 National Trade Estimate: Strong Binding Rules to Advance Digital Trade,” USTR, <https://ustr.gov/about-us/policy-offices/press-office/fact-sheets/2020/march/fact-sheet-2020-national-trade-estimate-strong-binding-rules-advance-digital-trade;>
103. On March 26, 2021, USTR announced next steps in its Section 301 investigation into Digital Services Taxes. See “<https://ustr.gov/about-us/policy-offices/press-office/press-releases/2021/march/ustr-announces-next-steps-section-301-digital-services-taxes-investigations>”
104. Note, this analysis does not incorporate acts of war by nation states or terrorism. For example, the geographic concentration of data centers and Internet exchange points are vulnerable to military action or terrorists use of force. In the context of security and resilience, the concentration risk problem is not three business providers, which is addressed in the competition policy section. Rather, it’s that three major clouds host an increasing share of mission critical applications. The good news is that cloud providers have strong business incentives to prevent outages and have designed independent regions and availability zones. However, some services and management functions are inherently global.
105. On May 12, 2021, President Biden signed a sweeping executive order to improve cybersecurity. There have been many prior executive orders to improve security, but this order is much more specific. See: <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/>.
106. State and local governments should be included where appropriate and leverage existing forums and channels. Senior officials who are responsible to lead these efforts could include the director of DHS’s Cybersecurity and Critical Infrastructure Security Agency or similar.
107. The IT Sector Coordinating Council (IT-SCC) and its government counterpart, the IT Sector Coordinating Council, could be a way to implement this with an existing organization. However, we propose something smaller and more informal, enabling candid conversations. It should also be more technical; for example, members could include the corporate Chief Information Security Officers (CISOs). A working group of the IT-SCC could be created, composed of the CISOs of the top 7 cloud providers. For more, see <https://www.it-scc.org/>. There are additional government efforts spread across multiple agencies and regulatory requirements that need to be better aligned and rationalized. This could build on the DHS’s voluntary Critical Infrastructure Cyber Community Program, see “Critical Infrastructure Cyber Community C³ Voluntary Program,” Cybersecurity & Infrastructure Security Agency, last updated January 17, 2020, [https://www.cisa.gov/ccubedvp](https://www.cisa.gov/ccubedvp;); NIST’s Cybersecurity Framework, <https://www.nist.gov/cyberframework>.
108. See Securities and Exchange Commission, 17 CFR Parts 229 and 249, [Release Nos. 33-10459; 34-82746] “Commission Statement and Guidance on Public Company Cybersecurity Disclosures,” <https://www.sec.gov/rules/interp/2018/33-10459.pdf>.
109. See the Cloud Security Alliance website, <https://cloudsecurityalliance.org/research/cloud-controls-matrix/>, or security control mapping for NIST Assurance Frameworks, <https://docs.aws.amazon.com/quickstart/latest/compliance-nist/welcome.html>. For DOD IT infrastructure the Cybersecurity Maturity Model Certification is the relevant certification.
110. The May 12, 2021 cybersecurity order instructs OMB and GSA to map compliance frameworks onto FedRAMP requirements and allows those frameworks to be used as a substitute for the relevant portion of the FedRAMP authorization process. This is a good first step, but it needs to go further and rationalize and consolidate the frameworks, reducing them, not just allowing for substitution.
111. ITIF has previously analyzed the effectiveness of FedRAMP and made multiple suggestions for improving the FedRAMP process. See Michael McLoughlin, “Reforming FedRAMP: A Guide to Improving the Federal Procurement and Risk Management of Cloud Services” (ITIF, June 2020),

<https://itif.org/publications/2020/06/15/reforming-fedramp-guide-improving-federal-procurement-and-risk-management>.

112. The EU has shown the most interest in regulating online platforms introducing two “ambitious” and “comprehensive” laws on December 15, 2020. See “Shaping Europe’s digital future,” European Commission, <https://ec.europa.eu/digital-single-market/en/online-platforms>; “Europe fit for the Digital Age: Commission proposes new rules for digital platforms,” press release, December 15, 2020, https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2347. In addition, see the U.S. House Judiciary Committee’s Subcommittee on Antitrust, “Investigation of Competition in Digital Markets,” 2020, https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf. The appointments of leading academic Tim Wu to a *new* White House role on “Technology and Competition” and the nomination of Lena Khan to the FTC are also notable. Academic research includes: Kenji E. Kushida, Jonathan Murray, and John Zysman, “Diffusing the Cloud: Cloud Computing and Implications for Public Policy,” *J Ind Compet Trade* (2011) 11:209–237, <https://link.springer.com/content/pdf/10.1007/s10842-011-0106-5.pdf>; Martin Kenny and John Zysman, “The Rise of the Platform Economy,” *Issues in Science and Technology*. Vol XXXII, No 3, Spring 2016, <https://issues.org/rise-platform-economy-big-data-work/>.
113. See IEEE Cloud computing standards. <https://www.computer.org/volunteering/boards-and-committees/standards-activities/committees/cloud>. Other groups include the Distributed Management Task Force at DMTF.org and the Object Management Group at OMG.org
114. See, for example, “Data Sovereignty and the Public Cloud,” Canadian government white paper, <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/cloud-services/gc-white-paper-data-sovereignty-public-cloud.html>; “The OECD Privacy Framework,” OECD, 2013, http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf; “OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data,” OECD, 2013, <https://www.oecd.org/sti/ieconomy/oecdguidelinesontheProtectionofPrivacyandTransborderFlowsOfPersonalData.htm>. ITIF has testified to the U.S. Congress on these issues. see Daniel Castro, “Testimony Before the House Committee on the Judiciary on Cloud Computing” (ITIF, July 25, 2012), <https://itif.org/publications/2012/07/25/cloud-computing-overview-technology-and-issues-facing-american-innovators>.
115. 2020 OECD Digital Economy Outlook. For non-OECD data localization rules, see “Strong Binding Rules to Advance Digital Trade,” in “Fact Sheet on the 2020 National Trade Estimate.”
116. If a private cloud in the world’s fourth largest economy were not successful, it would suggest that there are major problems with this approach. Esat Dedezade, “Microsoft to deliver cloud services from new datacentres in Germany in 2019 to meet evolving customer needs,” Microsoft, August 31, 2018, <https://news.microsoft.com/europe/2018/08/31/microsoft-to-deliver-cloud-services-from-new-datacentres-in-germany-in-2019-to-meet-evolving-customer-needs/>; “Microsoft Cloud Deutschland transition,” Microsoft, <https://www.microsoft.com/en-us/cloud-platform/germany-cloud-regions>.
117. A February 2016 report by McKinsey Global Institute estimates that cross-border data flows contributed nearly \$2.8 trillion to the global economy in 2014 through its enablement of the flow of goods, services, and other resources. James Manyika et al., “Digital Globalization: The new era of global flows,” McKinsey & Company, February 24, 2016. Further, the number of requests is also relatively small. For example, the latest AWS H2 2020 biannual report on Information Requests shows 523 requests for information (court orders, subpoenas).
118. Confidence-building measures have been used in other domains, including governmental national security treaties such as arms control. These technical programs promise more actual security than walls that can be breached and provide false-security. The Maginot line refers to a fixed series of French fortifications that the German army simply bypassed and went around in World War II by invading through Belgium.

119. U.S. Department of Justice, “Promoting Public Safety, Privacy, and the Rule of Law Around the World: The Purpose and Impact of the CLOUD Act,” White Paper, April 10, 2019, <https://www.justice.gov/opa/pr/justice-department-announces-publication-white-paper-cloud-act>.
120. Robert D. Atkinson, “How to Reform Worker-Training and Adjustment policies for an Era of Technical Change” (ITIF, February 20, 2018), <https://itif.org/publications/2018/02/20/technological-innovation-employment-and-workforce-adjustment-policies>.