CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

WORKING PAPER

AUGUST 2020

Cloud Security: A Primer for Policymakers

Tim Maurer and Garrett Hinck

Cloud Security: A Primer for Policymakers

Tim Maurer and Garrett Hinck

Carnegie Endowment for International Peace
Publications Department
1779 Massachusetts Avenue NW
Washington, DC 20036
P: + 1 202 483 7600
F: + 1 202 483 1840
CarnegieEndowment.org

This publication can be downloaded at no cost at CarnegieEndowment.org.

CONTENTS

About the Authors

Tim Maurer is co-director of the Cyber Policy Initiative and a senior fellow in Carnegie's Technology and International Affairs program. He works on the geopolitical implications of the Internet and cybersecurity, with a focus on the global financial system, influence operations, and other areas of importance as actors exploit the gray space between war and peace. In 2018, Cambridge University Press published his *Cyber Mercenaries: The State, Hackers, and Power*, a comprehensive analysis examining proxy relationships between states and hackers.

As part of his policy engagement, he regularly engages with governments and industry and participates in U.S. track 1.5 dialogues. He served as a member of the Freedom Online Coalition's working group "An Internet Free and Secure," the Research Advisory Network of the Global Commission on Internet Governance, and co-chaired the Advisory Board of the Global Conference on CyberSpace. Maurer's research over the past decade covers cyber conflict, strategy and norms relating to cyberspace, as well as Internet governance, in addition to assessing the human rights implications of export controls and sanctions. His work has been published by the *Washington Post*, *Foreign Policy*, CNN, *Slate*, *Lawfare*, *Jane's Intelligence Review*, *TIME*, and he has appeared on BBC World Service, Al Jazeera, and Bloomberg.

Prior to joining Carnegie, Maurer was the director of the Global Cybersecurity Norms and Resilience Project at New America and head of research of New America's Cybersecurity Initiative. He also spent several years working with refugees and in the humanitarian field, including with the United Nations in Rwanda, Geneva, and New York. He is a mentor for first generation students through Harvard University's First Generation Mentorship Program.

Garrett Hinck is a PhD student at Columbia University. Previously, he was a research assistant with the Cyber Policy Initiative at the Carnegie Endowment for International Peace.

Acknowledgments

The authors wish to acknowledge the invaluable contributions of their current and former colleagues at Carnegie's Cyber Policy Initiative, including Ariel Levite, Arthur Nelson, Evan Burke, George Perkovich, Jon Bateman, Natalie Thompson, Ronit Langer, and Wyatt Hoffman. Additionally, gratitude goes to Trey Herr and several other experts in industry as well as government and regulatory agencies who wish to remain anonymous for their outstanding insights, suggestions, and comments on drafts of this paper. The paper's content is the sole responsibility of the authors and does not necessarily represent the views of any other individuals or institutions.

Summary

The growth of the cloud has been truly astonishing. In less than fifteen years, it has become part of everyday life and casual conversations about moving photos and other data into the cloud. Omnipresent advertisements at airports, on buses, and on websites further embed the term in society's collective consciousness. Tech companies report multiple billions of dollars in revenues, increasingly driven by their cloud businesses. Even the Pentagon is betting on the cloud with its $10 billion Joint Enterprise Defense Infrastructure (JEDI) contract.[1] By 2020, the overall cloud services market is expected to be $266.4 billion, a 17 percent increase compared to 2019.[2]

The coronavirus pandemic has revealed how important the cloud is for bolstering societal resilience. According to a March 2020 *Business Insider* article, one expert projected that more than half (55 percent) of workloads would be migrated to the cloud by 2022 compared to 33 percent now; he claimed that these projections "now look conservative as these targets could be reached a full year ahead of expectations given [the current] pace."[3] In the wake of the pandemic's initial outbreak and the accompanying move to telework, previously cautious executives started seeing migration to the cloud as an urgent necessity.

As businesses increasingly rely on cloud services, the role of the huge cloud service providers (CSPs) has received greater scrutiny. Calls for regulating CSPs have been growing amid concerns about the systemic risk of businesses' move to the cloud. For example, a 2018 report estimates that a three-to-six-day outage of a major CSP would cause economic losses up to $15 billion.[4]

However, the debate about cloud security remains vague and the public policy implications poorly understood. This starts with the question: what is the cloud? Most of the debate is about the public cloud,[5] and the short answer is "cloud computing is really just a fancy name for someone else's computer," as Rob Joyce, then chief of the Tailored Access Operations at the U.S. National Security Agency, explained in 2016.[6]

Thinking through the public policy implications, the image of a cloud obscures as much as it explains. A more nuanced picture emerges when the cloud is considered in terms of its layers—from the physical data centers and network cabling that form its foundation to the virtual software environments and applications that everyday users interact with. Yet a more technical understanding will only go so far. An appreciation of the multibillion-dollar marketplace for cloud services is also required.

What makes the public cloud interesting is that the thousands of "someone else's computers" that compose it are concentrated in the hands of a few CSPs. Amazon Web Services (AWS), Microsoft Azure, and Google Cloud are known as hyperscale CSPs with firms like Alibaba Cloud and Tencent playing a similar role in China. As cloud services have grown, a few vast enterprises built on the backs of these tech giants, in their U.S. and Chinese variants, have secured most of this lucrative market.

Protecting this new, highly complex infrastructure is a herculean task, one enabled by the size and accumulated talent of the major cloud providers but also potentially imperiled by their growing importance for critical industries. When thinking about cloud security from a public policy perspective, the need to address an existing public policy problem must be further differentiated from the need to address an emerging public policy problem. The existing problem is the rising cost of cyber attacks and the reality that most organizations—governments and companies—cannot effectively protect themselves. Very few organizations can rival the security teams of the major CSPs and are therefore better off entrusting their security to these external firms' security teams. The emerging problem is the systemic risk associated with a centralized approach.

Overall, cloud security is a nascent policy area, particularly for policymakers concerned about potential systemic risk. As policymakers consider risks associated with the cloud, it will be important for them to connect threats to impacts. This is a difficult task due to the variance in potential impact depending on the data and services at risk. Furthermore, any potential regulations will have to balance other public policy interest areas such as data governance, geopolitics, and antitrust policy.

This primer provides an overview of the cloud and its security dimensions covering the basics of some of the most pressing questions for policymakers and technologists today. In many cases, this paper is only a starting point highlighting the need for further study. To avoid recreating an insecure cyberspace in the cloud, further study on such topics should be an urgent priority. This primer specifically adds value by offering (1) a conceptualization of the layers of the cloud services' architecture in table 1 on page 9, (2) an overview of the evolution of the cloud marketplace in chapter 2 starting on page 10, (3) a timeline of key cloud security incidents in table 5 on page 25, (4) a mapping of potential cloud security threat vectors to their impacts from a technical perspective in figure 5 on page 30, and (5) a framework for assessing the severity of cloud security incidents based on their impact in table 6 on page 37.

Introduction

Relying on someone else, such as a cloud service provider (CSP), to store and process data requires trust and a willingness to give up control. There are different reasons why people are sometimes willing to do so. Often, someone else has more expertise to do something, so people are willing to let them do it. Sometimes, someone else can do a task cheaper or faster, so others are willing to hand the task off. At other times, people may just do something because apparently everybody else is doing it. Governments, companies, and individuals alike have been increasingly relying on the public or hybrid cloud, run by CSPs, for these various reasons.

What makes the cloud particularly interesting is that the thousands of "someone else's computers" that compose it are concentrated in the hands of a few CSPs. Unlike the Organization of the Petroleum Exporting Countries (OPEC), this oligopoly is not determined by the geographic location of resources; the cloud's market structure is a combination of historic path dependence, access to large markets, and, importantly, the network effect.[7] Amazon Web Services (AWS), Microsoft Azure, and Google Cloud are known as hyperscale CSPs, with firms like Alibaba and Tencent playing a similar role in China.[8]

With this explosive growth, it is no surprise that policymakers are increasingly turning their attention to the cloud. Cloud computing and storage affects (and is affected by) policymakers in numerous areas such as data governance, technological development, geopolitical influence, and antitrust legislation, making the security of the cloud a particularly salient topic for public policy. As the public cloud infrastructure has become more consolidated within the cloud infrastructure market, concerns have risen about potential systemic risks. For example, a 2018 report estimates that a three-to-six day outage of a major CSP would cause economic losses of up to $15 billion.[9] However, the debate about cloud security remains vague, and the public policy implications are poorly understood. This paper provides an overview of the different policy dimensions that must be considered so as to inform a more nuanced and robust debate.

Concerns about the cloud's systemic risks have become common. However, it is important not to lose sight of the full picture and not to conflate two important dimensions of the public policy concerns surrounding cloud security: the need to address an existing public policy problem and the need to address an emerging one.

The existing public policy problem is cybersecurity. In 2017, cyber crime cost the global economy as much as $600 billion.[10] In 2019, Accenture estimated that a total value of $5.2 trillion will be at risk due to cyber crime globally over the next five years.[11] Of course, these estimates are just that—approximations that vary widely. The precise figures matter less than the scale of potential costs. In short, despite various efforts to contain these risks over the past twenty-five years, the costs of cyber attacks continue to increase, not decrease, and most organizations—governments and companies—struggle to effectively protect themselves.

The move to the public or hybrid cloud is one of the most promising options for better protecting organizations from cyber attacks. Very few organizations can rival the security teams of the large CSPs and are therefore better off entrusting their security to these external teams. This does not mean that the cloud is secure, but it is more secure relative to the security measures most organizations could otherwise achieve. That is part of the reason why companies like Capital One continue to pursue their "cloud first" strategy despite the massive reputational costs that followed its 2019 data breach and why governments have been putting in place "cloud first" policies.[12]

In short, the move to the cloud is the "Fort Knox" solution to the existing and growing problem of cyber insecurity. According to Harvard professor Jonathan Zittrain, "Fort Knox represents the ideal of security through centralization: gunships, tanks, and 30,000 soldiers surround a vault containing over $700 billion in American government gold."[13] Private CSPs represent the same idea for the protection of digital assets and processes.[14] From Alex Stamos, who served as Facebook's former chief information security officer (CISO), to the CISO of a large financial institution, technical experts agree that the security provided by major CSPs is significantly better than most organizations can achieve themselves.[15]

The emerging public policy problem is the new forms of systemic risk that cloud services may create. As highlighted by the villain Goldfinger in the 1964 eponymous James Bond movie, the downside of creating a Fort Knox is that it becomes an inspiration and dream for criminals to target and that, once breached, the impact could be devastating and widespread. This is an important public policy problem to consider but one that remains hypothetical to date.

Yet cloud security is not an all-or-nothing affair. It is simply less well conceptualized than existing cybersecurity. Potential risks range from the cascading effects of temporary disruptions to the exploitation of vulnerabilities in the underlying hardware and software that run the cloud. And, of course, while the cloud can be seen as "someone else's computer," the basics of cybersecurity still apply, and customers may expose themselves by not fulfilling their end of the shared responsibility for security. Understanding the new potential risks associated with the cloud and what their impacts might be is a crucial task for policymakers to undertake now.

This paper is a first step to building that understanding. As a primer for policymakers on the cloud, this study outlines how to conceptualize the cloud and describes the evolution of the cloud market. It then discusses cloud security in detail, using a timeline of past incidents together with in-depth case studies of the most significant incidents that are publicly known. Together, these serve as a foundation for developing a comprehensive framework for mapping the various risks and a severity schema to prioritize them. The paper then briefly outlines additional public policy issues to take into account while considering cloud security. Finally, it sums up and discusses the implications for public policy, while listing promising areas for future security-related research.

Chapter 1: What Is the Cloud?

At its most basic level, the cloud is simply someone else's more powerful computer that does work for others. There is no one single cloud—so while it might be accurate to say that data crosses the internet, it is not correct to say that such data is stored in an ephemeral form, hovering somewhere in the sky. In fact, the cloud stores and transports data across a global infrastructure of data centers and networks. A more accurate description of the cloud is that cloud services are an abstraction of a parallel system of computers, data centers, cables, infrastructure, and networks that provides the power to run modern enterprises' and organizations' digital operations and to store their data. Building the necessary infrastructure for cloud services on a truly global scale has been one of the most significant architectural achievements of the past decade—and it mostly exists behind the scenes, out of common knowledge. With that said, as chapter 2 highlights, the cloud marketplace has evolved significantly over the years, as has the cloud itself.

To make sense of the transformative impact of cloud services, first consider how computing, for example, worked prior to widespread cloud adoption. In the past few decades, for every computational task that a company or individual needed to do, they had to have their own computers, servers, and even data centers. For instance, Capital One, a major company in the financial services sector, announced in 2015 that it would move all of its apps to the AWS cloud, meaning that it subsequently did not have to build and buy data center storage as it rolled out new apps.[16] For smaller businesses, the costs of information technology (IT) procurement—that is, buying all the necessary computers and setting up the necessary networking for inhouse data storage and processing capabilities—were prohibitive to rapid growth.

When companies like Amazon, Google, and Microsoft began to offer storage and computing power as services in the late 2000s, they changed this paradigm. These massive IT giants could manage networks of data centers, servers, and networking at global scales—meaning they could take

advantage of economies of scale to offer computing as a service—at prices that would beat internal costs for most companies and still make them a profit, especially after significant price drops starting in 2014.

Amazon, Google, and Microsoft particularly focus on providing the basic elements of IT infrastructure—server space and computing power—that are highly scalable, custom-configurable, and capable of being rapidly deployed and shifted. However, cloud computing encompasses a wide range of service types in which different firms predominate (see chapter 2), and the services provided include the basic infrastructure to build digital platforms on top of ready-made applications delivered over the internet. These various services can be grouped into the three principal types of cloud services. In practice, the major CSPs offer different services spanning all three of these categories.[17]

- Infrastructure as a service (IaaS): CSPs provide basic access to storage, networking, servers, or other computing resources.

- Platform as a service (PaaS): CSPs provide an environment—a platform—for customers to build and deliver applications.

- Software as a service (SaaS): CSPs build, run, and host applications delivered over the internet, which customers pay to access.[18]

Defining Cloud Computing

Given the particular importance of cloud computing as a service, it is worth considering a 2011 definition of cloud computing by the U.S. Department of Commerce's National Institute of Standards and Technology (NIST)—the agency that sets technology standards. According to NIST, "cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (. . . networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[19] Essentially, this definition touches on five key characteristics of cloud computing: 1) on-demand self-service, 2) rapid elasticity, 3) measured service, 4) broad network access, and 5) resource pooling.[20]

The first, on-demand self-service, means that customers can use capabilities only when needed and don't have to pay more. They are also siloed from each other even while making use of the same resources. Customers can select computing capabilities automatically—without needing any human support from the CSPs they use. And on the CSP side, on-demand self-service means that any customer request can be handled automatically. No technician has to go configure a server when a customer selects additional computing capacity. Automatic digital systems handle the allocation, provisioning, and deployment of the needed infrastructure, platforms, and services.

Second, rapid elasticity means that the amount of resources dedicated to any one customer can at any time quickly increase or decrease depending on the needs of the customer. Because the resource capacity of a CSP is exponentially larger than the likely needs of any one customer, customers can scale their operations rapidly without taxing the CSP.

Third, measured service captures how CSPs manage and price their services. CSPs, like AWS and Microsoft Azure, charge customers for the resources they use on a unit-per-time basis— these units are an abstraction of the resources used. For instance, AWS's Elastic Compute Cloud measures its service in units that AWS defines in terms of standard central processing unit (CPU) integer processing power.[21]

Fourth, broad network access means that customers access these services over the network, including potentially the public internet. This is a straightforward but highly important point from a security point of view. No longer are computing resources solely part of a firm's internal network—instead, in many cases, the core operating systems rely on connections that could be open to the entire global internet. In this respect, both the capacity and security of these network connections are critical. Even private cloud solutions, where the cloud servers are accessed over a private connection, would require significant bandwidth.

Fifth, resource pooling means that a CSP combines its resources such that each customer shares the same infrastructure with other customers in a dynamic fashion, to be apportioned and reapportioned as necessary. This feature is what makes cloud computing a more efficient model than separate computing resources for each firm. CSPs can take advantage of economies of scale to build infrastructure and platforms at mass scale—and share these resources among multiple customers at the same time to save costs and unneeded capacity.

Underlying Technologies

While modern CSPs rely on highly complex systems for allocating, managing, and deploying resources among millions of customers, three key technologies are essential for understanding how cloud services work at scale: virtualization, hypervisors, and containerization.

Virtualization allows for abstraction between physical hardware and individual computers. Essentially, virtualization allows for multiple computers, referred to as virtual machines, to exist on the same physical server. Beyond computational tasks and storage, entire networks can be built through virtualization.

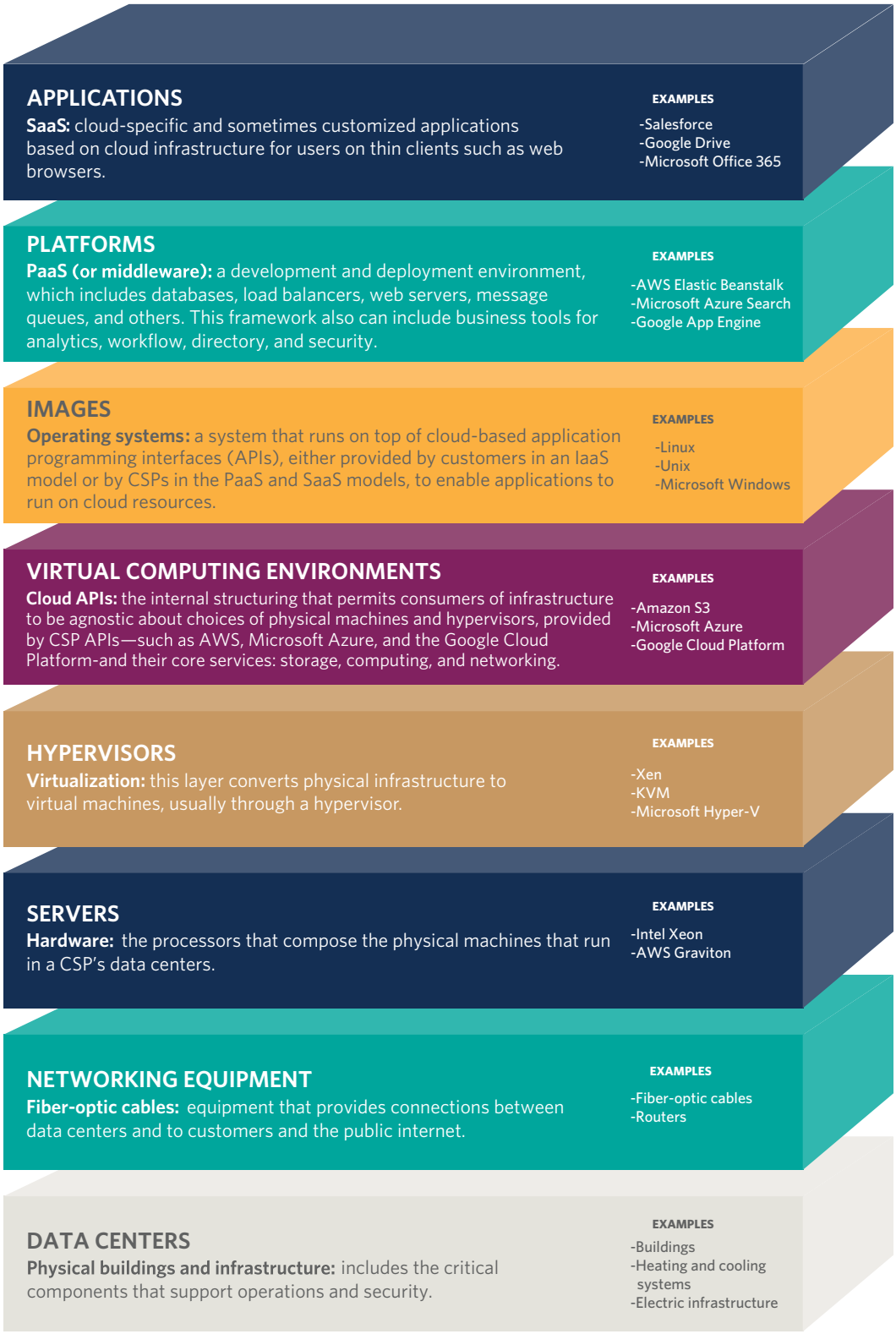
Hypervisors are programs that manage virtual machines, servers, the connection between those virtual machines and servers, and the allocation of resources to the virtual machines. Thus, it becomes possible to have a whole bank of physical servers, each running a hypervisor, and then create virtual machines across this bank of servers. CSPs refer to the virtual machines they create for their customers as instances, since they only last as long as needed and, once they are no longer needed, they are spun down to free up capacity.

Containerization is a refinement of virtualization that works by running discrete containers within the same operating system, basically moving up the abstraction provided by virtualization by one level. Containerization caught on around 2014 with the introduction of a new tool called Docker, which made it much more convenient and efficient to implement containerization for business uses. While a virtual machine includes an entirely virtual operating system, a container is an isolated environment within one single operating system. In terms of layers, while a hypervisor lies between the hardware and virtual machines, each with their own operating system inside them, a container sits on top of an operating system that is on top of the container engine, and then the hardware is below.

In table 1, the infrastructure for the cloud is visualized according to a layers model, similar to the Open Systems Interconnection model for the internet itself.[22] The table presents a hierarchy of layers from the physical data centers at the bottom to the entirely virtual application layer at the top, allowing the various parts of the cloud to be simplified. The descriptions for each layer present the key technologies operating at that level, along with several examples.

TABLE 1

Visualizing Cloud Architecture



Cloud Deployment Models

CSPs also offer their services in three main deployment models: the public cloud, a private cloud, or a hybrid cloud.

In the **public cloud**, customers share the same infrastructure available to be rented out to the public. A CSP manages the infrastructure and allows prospective customers to purchase resources. The customer uses the same CSP-allocated infrastructure as other customers.

A **private cloud** arrangement is designed for a single customer, essentially housing resources on premises or off premises but still isolated from other customers. An organization may set up its own private cloud, or it may contract with a CSP to do this.

Some organizations choose a **hybrid cloud**, in which they combine a public cloud service from a CSP with either a private cloud setup or a more traditional data center so they can communicate and share data and applications.[23] Some organizations opt for this arrangement because they have sensitive data that they consider too risky to store in a public-cloud-only environment, but they still want to take advantage of the public cloud's computing power to run applications.[24]

Customers may choose a deployment model based on their particular needs, including whether the data and services they are migrating to the cloud are especially critical or sensitive. For instance, a customer with sensitive personal data to manage (like a healthcare provider) might choose a private cloud to minimize the risks that their data may be exposed to their CSPs' other clients. Increasingly, hybrid cloud options are predominating, as customers realize that different types of data require different levels of security.

Chapter 2: The Origins and Evolution of the Cloud and Its Market

The concept of cloud services predates the internet itself. The history of these services illustrates how computing has moved from the core to the edge of a network and back, as technology evolves. An early example is the mainframe computer model produced by IBM in the 1950s and 1960s, which were the most powerful machines at the time. Computing was organized around the core of these mainframe computers used by large government and industry leaders. Access to these computers was organized on a "time-share" model allowing multiple users to share the resources of one computer.[25]

Starting in the 1980s, IBM's mainframe computers became obsolete as computing capabilities drastically improved, making it possible for desktop machines to carry out most ordinary work. Computing became decentralized and moved from the core—IBM mainframe computers—to the edge, the devices that provide users entry points into the network, which increasingly became individuals' desktop computers. In the 2000s, the trend swung back in the other direction with computing power becoming centralized again around the core of data centers, partly due to the development of virtualization and hypervisors. Several years earlier, in 1996, business executives and academics had first coined the term "cloud computing" to refer to the delivery of computing services over the quickly expanding yet still nascent commercial internet.[26]

The development of cloud services was enabled by a series of technological advances in computing hardware—chips and networking equipment—as well as the growth of the internet in the 1990s (see table 2 below). It was the growing level of interconnections made possible by linking multitudes of individuals and organizations together that enabled cloud computing to become economically attractive. Bandwidth was also a critical factor in enabling the delivery of cloud-based applications over the internet. Later, innovations in software, like containerization, would further enhance the possibilities of cloud storage and computing.

The origin of the cloud as it is known today was, if not accidental, a byproduct of a business strategy very much focused on tangible rather than virtual goods: In 2003, Amazon needed a solution to scale and to manage the backend computing resources required for its quickly expanding online marketplace for goods. Amazon developed a strategy that was "completely standardized, completely automated, and relied extensively on web services for things like storage" to solve this problem.[27] Amazon then realized that this did not only solve its own operational problem but could be a viable business itself: selling virtual storage as a service.[28] Other businesses needed their computing infrastructure to scale over time as their customer base grew, but the cost of procuring and setting up servers was a major barrier.

By 2006, Amazon had launched two services, Elastic Compute Cloud and Simple Storage Service, which provided computing and storage resources in the first major commercial cloud offering. These were the first two major components of AWS. During a speech in 2006 shortly after these announcements, Amazon Chief Executive Officer Jeff Bezos called the services "the guts of Amazon," drawing attention to their role in supporting the company's web store.[29] Two years later, in 2008, Google launched its App Engine, which would go on to become a central part of its cloud services offerings.[30] Rather than providing an IaaS offering like AWS, the Google App Engine was a PaaS offering, essentially providing developers with an integrated environment for building out new products and deploying them.

TABLE 2

Key Developments in Cloud Services

1959	Computer scientist John McCarthy proposes a time-sharing system for an IBM computer to be acquired by Massachusetts Institute of Technology research scientists, inspiring a number of time-sharing systems in the 1960s.[31]
1969	The U.S. Department of Defense's Advanced Research Projects Agency connects the first two computers on its network, known as ARPANET, which would go on to become the predecessor to the internet.[32]
1996	The term cloud computing is first used by engineers at Compaq.
2003	An internal paper at Amazon outlines a concept for standardized, automated computing infrastructure that relies on web storage.
2006	AWS offers the first wide-scale commercial cloud services: Simple Storage Service and Elastic Compute Cloud.
2008	Google's App Engine is the first major competitor to Amazon in the cloud services market; initially, Google focuses on a PaaS model.
2009	Chinese e-commerce giant Alibaba launches its cloud service, Aliyun, later rebranded as Alibaba Cloud.
2010	Microsoft launches its Azure cloud service.
2011	NIST, a part of the U.S. Department of Commerce, publishes the first government definition of cloud computing. Then president Barack Obama's administration outlines a "cloud first" strategy for the federal government.[33] Microsoft launches its Office 365 service, a SaaS offering.
2014	A new tool called Docker enables the widespread adoption of containerization in many cloud solutions, increasing efficiency. AWS and Azure both bring their services to China in partnership with local firms.
2015	A price-cutting war between major CSPs forces many other firms that attempted to build cloud offerings, such as HP and Oracle, to leave the market. According to market research, the size of the global cloud market is $100 billion in revenue.[34]
2018	AWS, Azure, and Google Cloud continue rapid expansions of their global footprint, building out regions in Europe and Asia and moving into emerging markets in the Middle East and Africa.

As other companies started their own cloud offerings, AWS was expanding, as illustrated by Netflix's 2008 announcement that it would move all of its data to AWS. By the end of 2008, both Elastic Compute Cloud and Simple Storage Service were available in Europe, and, in 2009, AWS announced a second U.S. region on the West Coast.[35] IBM entered the cloud market in 2009 with a private cloud service for business storage and debuted its first public cloud service in 2011.[36] Microsoft announced its cloud service, Azure, in 2008 and launched service in early 2010.[37]

Each cloud platform grew out of the various companies' business models and strengths. For instance, Microsoft Azure built on the firm's other suite of Windows software to offer a PaaS solution initially.[38] IBM drew on its strengths in supplying enterprises by initially providing a private cloud product. Meanwhile, Google's App Engine was designed to appeal to application developers to build apps to operate on top of Google's other products.

The period of 2009 to 2011 was one of significant public and government interest in cloud services. In 2009, NIST published an initial draft definition of cloud computing, which went through a number of rounds of public comment before being finalized in 2011.[39] The same year, the U.S. government's chief information officer published a cloud strategy for the federal government, outlining a "cloud first" approach that envisioned moving 25 percent of the entire federal IT budget to cloud solutions.[40]

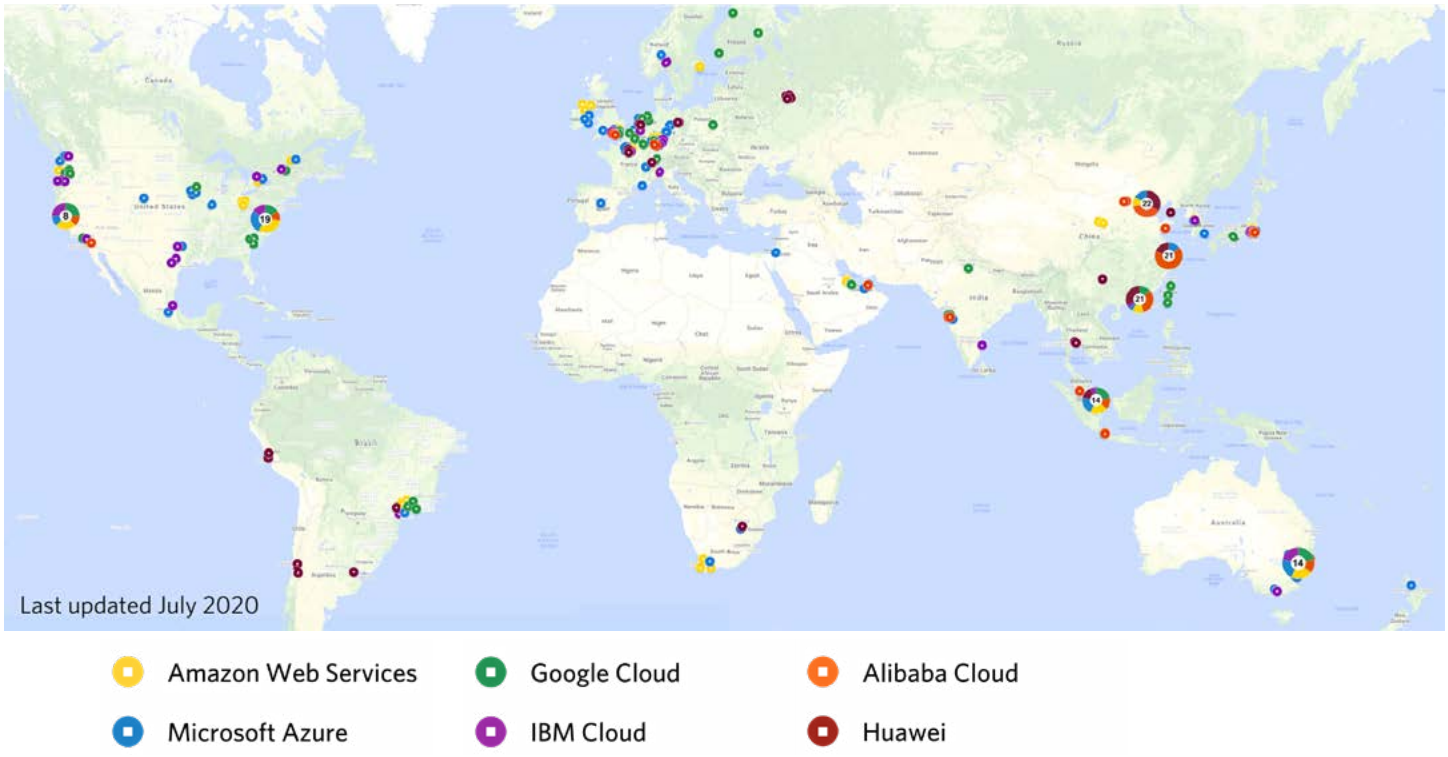
Global Expansion

The first global expansions undertaken by the major U.S. CSPs (AWS, Microsoft Azure, and Google Cloud) in the early 2010s were all similarly focused on a few key locations: Ireland, Singapore, and Tokyo—each of which provided low latency to large, developed markets. Since mid-2015, the CSPs have engaged in aggressive campaigns to expand their global presence, with AWS and Azure often entering the same region around the same time, attempting to keep up with each other. For instance, both launched new regions in London in 2016 and Paris in 2017.[41] But, in some cases, one CSP had a first-mover advantage: AWS opened a region in São Paulo, Brazil, in 2011, while Azure did not follow until 2014.[42]

Each CSP has a slightly different model for distributing its computing capacity and resources—which they typically divide into regions and zones. A region, in cloud computing parlance, is a number of data centers in an independent geographic area, logically separate from other regions managed by the same CSP, consisting of one or more zones. A zone is an isolated set of one or more data centers.[43] Typically within each region, a low-latency network connects the data centers. AWS has a smaller number of regions, but each region has at least three availability zones (AZs). Microsoft Azure has a larger number of regions than AWS but fewer AZs. Google Cloud's and Alibaba Cloud's models are closer to that of AWS—each of their regions has a number of AZs as well. An open question is whether some CSPs are globally more resilient than others depending on the location and distribution of their zones and data centers.

FIGURE 1

The Global Footprints of the Major CSPs

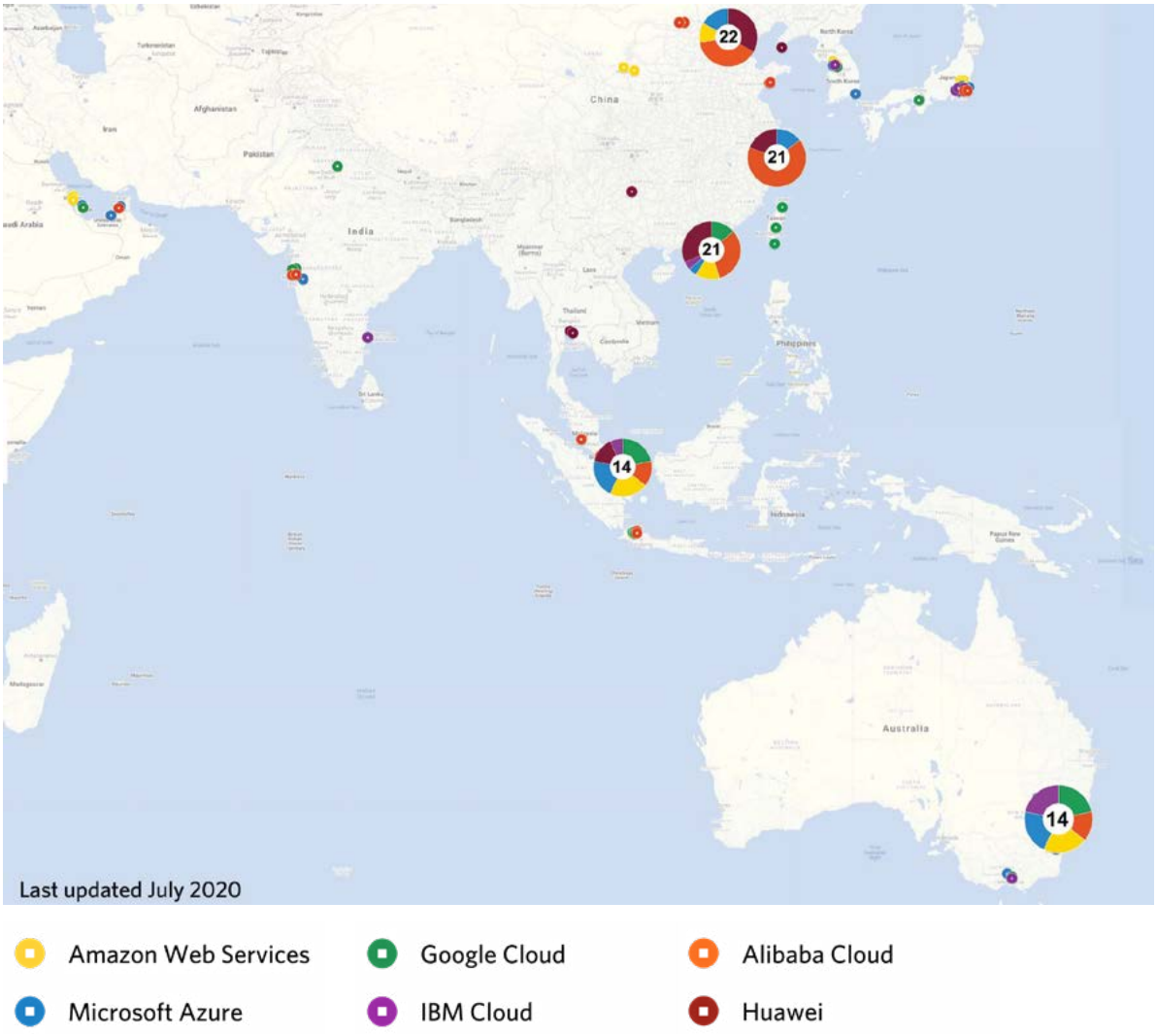


In expanding, AWS, Microsoft Azure, and Google Cloud have built their respective presences in nearly all large economies, a trend that has occurred quite rapidly. Figure 1 shows the expansion of the major CSPs across the world, and figure 2 shows the competition between U.S.-based and China-based CSPs for markets in Asia.

In Asia, Alibaba Cloud has built out from its main presence in China. It was also the first CSP in the Middle East, opening a region in Dubai in 2016. But Alibaba Cloud has also made forays into the West—it opened two regions in the United States in 2015 and has two regions in Europe too.[44] IBM expanded its global network earlier than some of the other players, with a $1.2 billion investment in 2014 to build data centers around the world.[45] While IBM has opened data centers in many of the same locations as its competitors, many of these are single data centers. In contrast, an AWS region, for instance, is composed of multiple AZs, with multiple data centers in each AZ. IBM may have a global presence, but it does not have the worldwide capacity of some of its larger peers.

FIGURE 2

The U.S.-China Competition for Market Share in Cloud Services in Asia



Market Trends (2011–2018)

Comparing revenue estimates for cloud services over time illuminates the astonishing growth of the market for such services over the last eight years. With that said, comparing different CSPs is more difficult because of several factors, including how the companies do not report their revenues by the various types of cloud services (IaaS, PaaS, and SaaS) and how many of these services straddle the lines between these three differing models.

Reporting on the market share of a given CSP requires a specific methodology that makes decisions about what counts as a cloud service and what is instead considered a virtualization service, so the numerous private companies that provide analyses of market share in the cloud often differ widely on their estimates. These reports therefore provide some insights but are not definitive. Moreover, there are many more specialized cloud companies that aim to provide cloud-based services for specific sectors. Given the overall focus of this primer on general, public CSPs, this section focuses on dynamics between the major CSPs, not the more specialized firms.

In 2011, the cloud market had just gotten off the ground, and several major players (Amazon, Google, and Microsoft) had recently entered it. According to one estimate, the global public cloud market grew almost sixfold in size, from $25.5 billion in revenue in 2011 to $145.3 billion in 2017.[46] These large totals must be taken with a grain of salt. They are estimates of the entire public cloud market and include other services, like business processes and advertising, based on underlying cloud services and platforms, including IaaS or PaaS.[47]

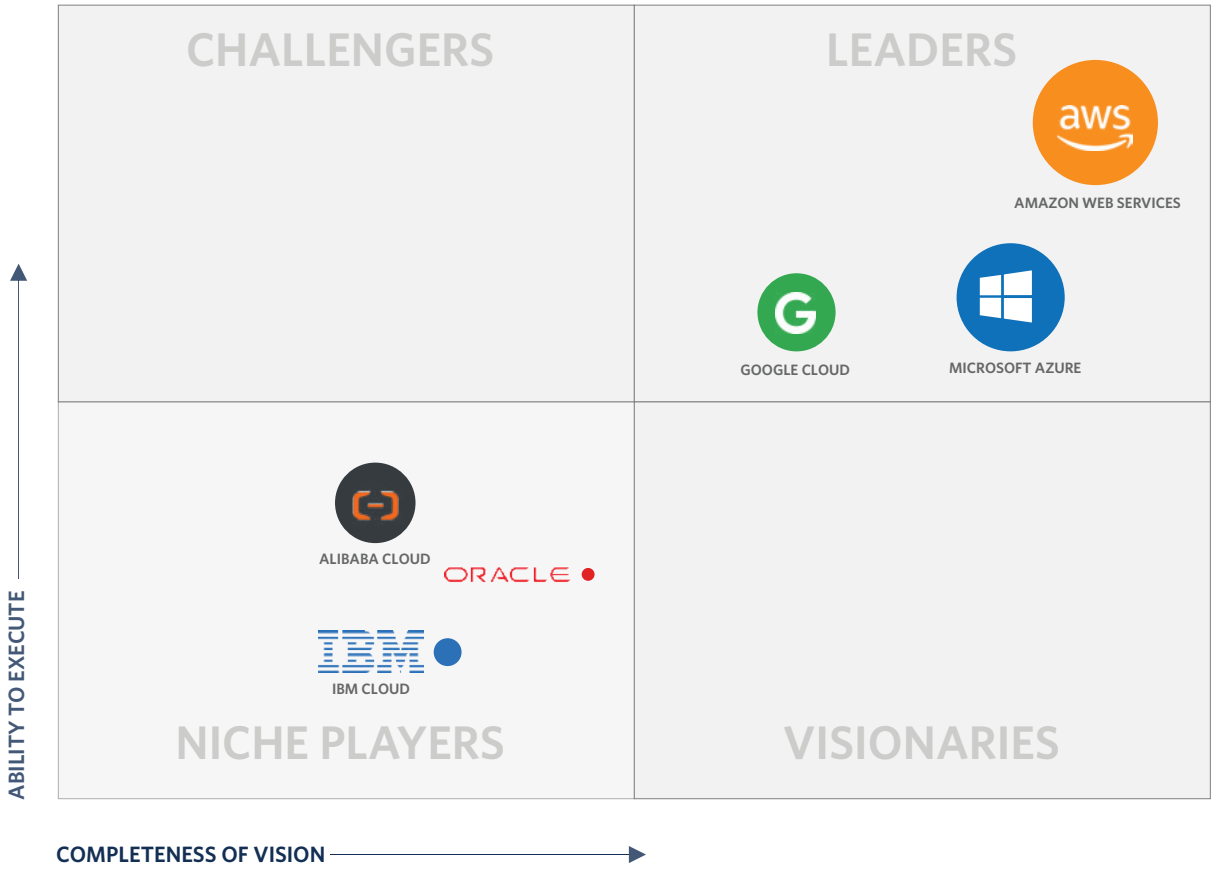
In 2011, market competition among CSPs started intensifying (see figures 3 and 4). As the first major IaaS offering, AWS dominated, but there were other major players, including the hosting company Rackspace—which also built an open-source cloud computing platform called OpenStack in collaboration with the National Aeronautics and Space Administration (NASA) in 2010.[48] In 2012, a market research report put AWS's share of the global IaaS market at around 35 percent.[49] At this time, other competitors, such as IBM and Telecom, had shares of less than 10 percent.[50] In the global PaaS market, in 2012, the same report put Salesforce as the leader, with a share of around 20 percent, closely followed by AWS and Microsoft Azure.[51]

By 2013, the overall cloud market was growing quickly—AWS's worldwide revenues in the third quarter that year were 55 percent more than the previous one, and other competitors were growing as well: the overall IaaS/PaaS market grew 46 percent compared to 2012.[52] This massive growth pointed to a transformation of business IT—a 2015 report found that 88 percent of U.S. businesses were using public cloud computing in some capacity and 63 percent were also using private cloud solutions.[53] However impressive these numbers are, it is difficult to estimate what this means—there is a vast difference between using a cloud service for email or data storage versus migrating core business functions.

One key development around 2014 was the improvement in containerization, discussed in chapter 1. Prior to this point, nearly all cloud computing services exclusively made use of virtual machines that emulate an entire computer—from applications down to hardware—in virtual form. In 2014,

FIGURE 3

Industry Leaders in Cloud Services



SOURCE: Gartner, "Magic Quadrant For Cloud Infrastructure as a Service," July 2019

NOTE: The sizes of the circles indicates each firm's relative 2018 market share (see table 3).

the new containerization tool Docker enabled a different type of emulation with containers, which provide isolated computing environments but run on a shared operating system. The Docker engine lays on top of a shared operating system and then generates isolated environments—containers—that provide the same separation of a virtual machine.

Efficiency comes from the shared operating system, meaning that, in contrast to virtual machines, containers do not have to take the space and computing power to replicate an operating system that will be discarded once the virtual machine is no longer needed. This allows much more efficient computing.[54] Many CSPs quickly integrated Docker into their products.

Price Wars

As competition in the market for infrastructure and platform services was heating up, the major CSPs dramatically dropped their prices. The year 2014, in particular, marked an inflection point after which increased competition coincided with more aggressive price drops. For instance, AWS's database storage rates decreased at an annual rate of 3.3 percent between 2010 and 2013, before dropping by 22.6 percent between 2014 and 2016.[55] This price cutting has continued since 2014, with the major providers moving from cutting rates for virtual machines to cutting prices for object storage.[56]

Price cutting led to consolidation in the market, as other companies that had hoped to build cloud arms off their existing businesses started to fold. In August 2014, Rackspace said it would discontinue its IaaS offering, instead shifting to a "managed cloud" model.[57] In October 2015, HP announced it would shutter its public cloud service, an announcement foreshadowed by one executive's explanation that "it makes no sense for us to go head-to-head" against AWS, Microsoft Azure, and Google Cloud.[58] Over the next two years, Verizon, AT&T, and Cisco also bowed out of cloud services.[59] A 2017 Gartner report found that AWS and Microsoft Azure were hosting 70 percent of IaaS workloads and predicted that, by 2019, their duopoly would force 90 percent of other IaaS providers out of the global market.[60]

A Global Market and Hyperscale Clouds

In 2018, estimates put the size of the overall cloud market at $160 billion.[61] The big players were Google Cloud, Microsoft Azure, and, biggest of all, AWS—each of these titans known in industry parlance as hyperscale CSPs (see tables 3 and 4). According to a widely cited definition from Cisco, to be hyperscale, a CSP must either take in more than $1 billion in annual revenues from IaaS, PaaS, or infrastructure hosting or receive $2 billion annually from SaaS.[62] Based on Cisco's definitions, which apply to hyperscale data centers for non–cloud service provider companies as well, there are twenty-four companies that have hyperscale data center operations.[63]

TABLE 3

Worldwide Market Share in IaaS Public Cloud Services (2017–2018)

Company	2018 Revenue (in billions $)	2018 Market Share (%)	2017–2018 Revenue Growth (%)
Amazon	$15.5 billion	47.8	26.8
Microsoft	$5.0 billion	15.5	60.9
Alibaba	$2.5 billion	7.7	92.6
Google	$1.3 billion	4.0	60.2
IBM	$0.6 billion	1.8	24.7
Others	$7.5 billion	23.2	11.1
Total	**$32.4 billion**	**100.0**	**31.3**

SOURCE: "Gartner Says Worldwide IaaS Public Cloud Services Market Grew 31.3% in 2018," Gartner, press release, 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.

These operators came to dominate global cloud markets by using economies of scale to drive down costs and amass market share at the expense of smaller providers.[64] In the last three years, Microsoft Azure has increased its market share in IaaS and PaaS, although AWS has retained its dominant position. Other hyperscale CSPs, in particular Google Cloud and Alibaba Cloud, have also increased their growth much faster than the cloud market's growth overall.[65]

TABLE 4

Hyperscale Companies in Cloud Services

	Hyperscale	Nonhyperscale
Major CSPs	AWS, Google Cloud, Microsoft Azure, Alibaba Cloud, Rackspace, Salesforce, Oracle	IBM, NTT, OVH, SAP, Dropbox
Other companies	Yahoo, Apple, eBay, Baidu, Tencent, Facebook	All other companies.

Hyperscale cloud computing has several important implications not just for cloud computing but for the design and functioning of the internet itself. The first is the sheer scale of hyperscale cloud computing. A mere handful of companies are now responsible for managing infrastructure on the scale of dozens of massive data centers, millions of computing nodes, millions of kilometers of network fiber connecting data centers, and a backbone that could be considered almost parallel to the internet, but which is not publicly connected (see table 4 above).[66] Second, hyperscale cloud computing changes the way that data travels between users and servers. Previously, users communicated directly with servers to retrieve their data in simple point-to-point communications. In hyperscale, with the amassing of resources inside the cloud, a user request will prompt internal communications and networking within the data center as the cloud provides not just raw data but services. This trend implies that complex computing and smart networking will increasingly become important for delivering internet services—as opposed to an internet that is primarily (or merely) a network of so-called dumb nodes designed for communication.

China and the Cloud

Increasingly, competitors from China are contesting U.S.-based CSPs for market share. The biggest of these competitors is Alibaba, the Chinese e-commerce company. Alibaba launched a cloud service, Aliyun, in 2009 and expanded it in 2015 with a billion-dollar investment that also rebranded it as Alibaba Cloud.[67] A major factor in the rise of Chinese CSPs has been their ability to dominate the Chinese market because of restrictions on foreign CSPs. For instance, a July 2018 McKinsey report found that 64 percent of public cloud spending in China went to Chinese CSPs and only 22 percent went to multinational vendors, even though, overall, Chinese companies have invested less in cloud services as part of their IT budgets compared to U.S. companies.[68] One report found that Alibaba Cloud had a market share of 45.5 percent in China in IaaS, followed by Tencent Cloud at 10.3 percent and China Telecom at 7.6 percent—with Azure and AWS having only around 5 percent each.[69]

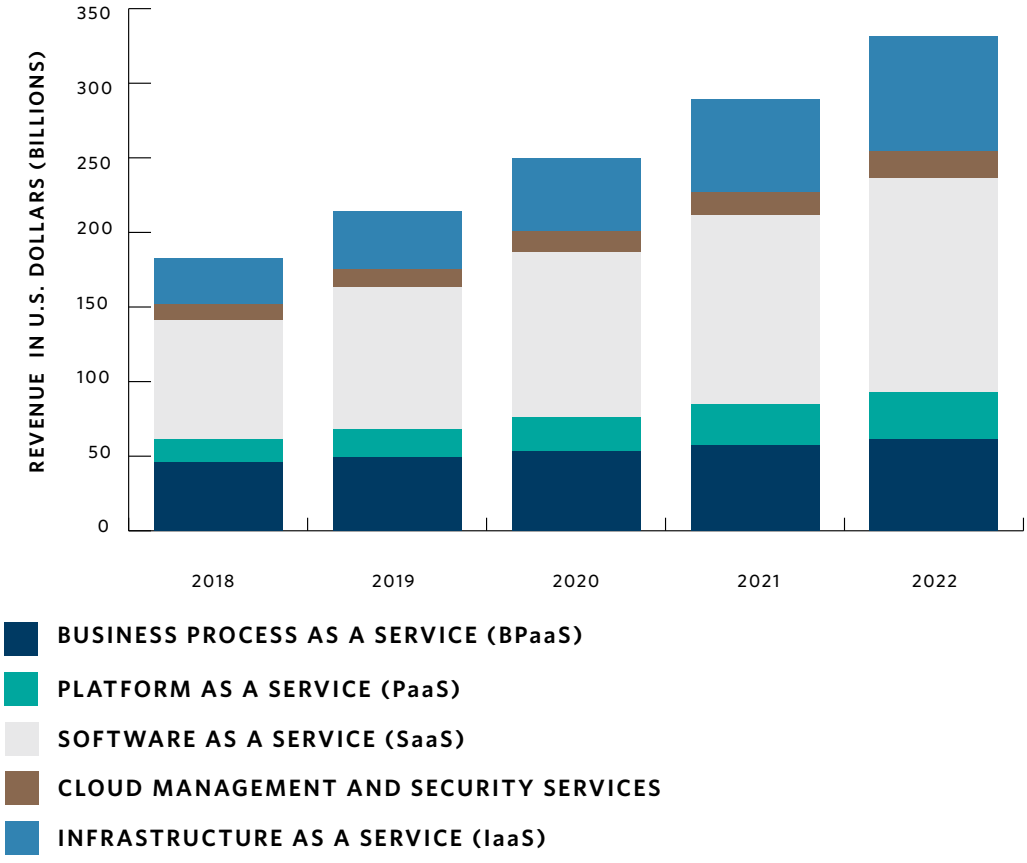
Beijing has imposed strict regulations on foreign CSPs that make it difficult to operate in China. For instance, foreign companies cannot directly set up data centers in mainland China and must instead contract with local Chinese companies and provide services through them. Azure was the first to develop such a partnership, with a firm called 21Vianet, and brought its services online in China in 2014. AWS followed shortly thereafter, partnering with Beijing Sinnet.[70] In 2017, new permit requirements from China's internet regulator and a new cybersecurity law made it even more challenging for foreign CSPs in China, particularly because of the law's data localization requirements.[71] These regulations forced AWS to sell a significant amount of its infrastructure in China to its local partner.[72]

The Growing Global Cloud

Overall, the cloud services market has grown massively since its nascent stage in 2011 in terms of revenue, business maturity, and global footprint. AWS has continued to be the global leader throughout this period, maintaining a steady market share. Competition has consolidated to include only a few other large players such as Azure, Google Cloud, and IBM. One area where these CSPs have failed to grow is China—instead, the rise of domestic Chinese CSPs, most notably Alibaba Cloud with its nearly majority market share, has allowed a new source of regional competition, particularly for Asian and emerging markets. Projections indicate that the global cloud market, particularly for IaaS and PaaS, will continue to exhibit strong growth (see figure 4 above), and Gartner has predicted the whole cloud market will be worth $280 billion in 2021.[73] As cloud services make up an increasing proportion of the profit margins of mega-multinationals like Amazon and Microsoft, competition for growing revenues will only intensify.

FIGURE 4

A Revenue Forecast for Worldwide Public Cloud Services



SOURCE: "Gartner Forecasts Worldwide Public Cloud Revenue to Grow 17.5 Percent in 2020," Gartner, press release, April 2, 2019, https://www.gartner.com/en/newsroom/press-releases/2019-04-02-gartner-forecasts-worldwide-public-cloud-revenue-to-g.

The general trend toward market consolidation, particularly for general public CSPs like AWS and Microsoft Azure, is a key development that policymakers should note. The extent to which business functions of so much of the modern global economy now depend on the operations of a handful of firms is staggering. Like electricity in the early twentieth century, the cloud infrastructure is now becoming an important pillar of modern life. As policymakers turn to evaluate the cloud market, as in any industry consolidation, these developments may provoke antitrust concerns, as discussed briefly in chapter 4. Yet this consolidation also has important security implications, discussed further below in chapter 3.

Chapter 3: Cloud Security

To discuss security with respect to the cloud, it is important to first highlight the broader cybersecurity landscape and benefits that migrating to the public or a hybrid cloud model provide. For most organizations, a migration to the cloud can significantly improve their security.

The State of (Continued) Cyber Insecurity

Despite years of investment and focus on cyber risks, the costs of cyber incidents are rising. Cyber attacks now count among the global risks with the highest likelihood and the greatest impact, according to the World Economic Forum's (WEF) 2019 Global Risks Report.[74] A WEF survey of experts in government, business, and civil society found that over 80 percent of respondents expected the risks of disruptive cyber attacks and thefts to increase further in 2019.[75] Data breaches are growing in size and severity. Between January 2017 to March 2018 alone, close to 1.9 billion records of personal or other sensitive data were compromised in data breaches.[76] A survey by the insurer Lloyd's of London found that 92 percent of European businesses surveyed had suffered a data breach, a statistic that indicates the ubiquity of cyber threats.[77]

The average costs of cyber incidents are also growing. Accenture's 2019 report on the cost of cyber crime found that, for the organizations surveyed, the average cost of cyber crime in 2018 was $13 million, a 12 percent increase over the previous year and a 67 percent increase over the previous five years.[78] The same report estimated that the value at risk from both the direct and indirect costs of cyber crime from 2019 to 2023 could be as high as $5.2 trillion.[79]

In response, organizations are increasing their cybersecurity spending but still often fall short. A report by Ernst and Young (now known as EY) found that only 8 percent of business professionals surveyed said that the information security function fully met their organization's needs, despite increasing budgets for cybersecurity.[80] In particular, smaller firms have fewer resources to allocate to

address all the vulnerabilities they face. The resource challenges affecting smaller organizations have broader consequences as well, since attackers often attempt to compromise the weakest link and leapfrog to other organizations.

As Alex Stamos, former chief information security officer at Yahoo and Facebook, has stated, "For most organizations, you are ten times more secure with your data stored and defended by a major cloud service provider than if you'd try to protect it yourself."[81] The CISO of a major financial institution echoed those remarks in a private roundtable discussion.[82] This rule of thumb applies to the vast majority of organizations that do not have the size and resources of a major CSP such as AWS or Google Cloud to bring together—and pay the salaries—of the world's leading security experts.

However, cloud services have their own associated security issues. They also do not eliminate all existing cybersecurity risks for those operating with on-premises systems. This section examines known security incidents and threats to the availability, confidentiality, and integrity of data in the cloud. Based on a series of case studies of key security incidents, it inductively develops a risk framework for categorizing threats to cloud security and the security of those operating in the cloud. Finally, it discusses how to connect these risks to potential impacts.

The Security of the Cloud

As previously mentioned, it is a misnomer to speak of the cloud as a single, cohesive entity. Therefore, the security of the cloud is much like security of the internet writ large—a broad swath of various potential threats, vulnerabilities, and risks affecting thousands, if not millions, of different types of services provided to actors ranging from single individuals to entire governments. There will never be an incident where the cloud goes down, so to speak, as if it could be switched off like a lamp. Instead, security issues in the cloud cover a spectrum ranging from complete failure and unavailability to limited performance or effects limited to subsets of data and services.

Cloud security risks are different from other types of cybersecurity risks because cloud security is networked, concentrated, and shared. First, cloud-based computing is inherently networked. Whether relying on an internet connection for the public cloud or specialized networking for a private cloud, anything in the cloud to some extent relies on networking. This means that any cloud infrastructure will involve network security and contending with potential threats that could take advantage of this.

Second, the cloud infrastructure is concentrated. As the discussion of the evolution of the cloud market showed, there is a remarkably small set of major CSPs, and though there is a diverse

ecosystem especially for PaaS and SaaS, many of those services rely on the Big Three CSPs—AWS, Azure, and Google Cloud. One positive effect of this concentrated market is that each CSP now has developed incredibly well-resourced security protocols and high concentrations of skilled personnel. Given the rarity of skilled personnel, this concentration effect is quite substantial in improving the security of the cloud. However, a downside of concentration is that incidents disrupting such CSP services have a much broader effect than narrower potential outages affecting one of their clients might. If an outage affected a key AWS service, as it did in 2017, platforms across the internet could suffer widespread effects.

Third, responsibility for risk in the cloud is inherently shared between customers and CSPs. The amount of responsibility and its delineation is dependent on what model of cloud services is being used, as is discussed later in this section. But, in general, there is no escaping that security is dependent on the efforts of two or more parties.

None of these conditions are new in cybersecurity. Yet it is important to remember when evaluating cloud security issues that all three apply for every cloud arrangement. These conditions may exacerbate existing security issues, as is often the case when customers' failures to put security protections in place exposes data on the public internet. Or they may combine to create new potential risks, such as the risk of a cloud provider being compromised and thereby exposing the data of its customers to theft.

Given the rapidly evolving nature of cloud services, it is difficult to map all security issues that may affect the cloud in the future or what the impact of those threats might be. There are many unresolved questions about the potential impacts of threats to cloud security that are relevant for public policy. For instance, to what extent does migration to the cloud reduce current cybersecurity threats? Are threats to the confidentiality, availability, and integrity of data mitigated or exacerbated in the cloud? Do cloud-specific security threats increase the potential magnitude of cybersecurity incidents?

An inductive approach based on past security incidents can begin to provide some form of an answer to such questions. The series of incidents over the past five years listed in table 5 and the in-depth case studies in the appendix reveal that CSPs have experienced disruptions affecting the confidentiality, availability, and integrity of data as well as processes and their reliability. While like much of this paper, these incidents revolve around the major CSPs, there are, of course, many other security issues involving more specialized varieties of cloud services.

So far, incidents involving CSPs have not yet caused long-term, widespread disruptions and to the extent that they have affected significant numbers of people, it is through the indirect effects of website downtime resulting from cloud outages. The list of incidents in table 5 shows that some

problems, such as environmental ones affecting data centers, occur repeatedly, while other risks, like the potential threats associated with the Meltdown and Spectre vulnerabilities, emerged unexpectedly and have yet to lead to any form of disaster.

TABLE 5

Key Cloud Security Incidents (2014–2019)

Sept. 2014	Sesponding to a vulnerability in the Xen hypervisor, AWS and Rackspace initiated reboots of their computing instances around the globe to implement security patches.83
Nov. 2014	A system update rolled out globally for Microsoft Azure's storage services caused failures in its virtual machines. Some customers experienced disruptions for up to eleven hours.[84]
Aug. 2015	Lighting strikes that hit the power grid in Belgium caused a failure in Google Compute Engine services for the local region. Some customers lost data because storage systems experienced repeated power drain.[85]
Sept. 2015	A failure in AWS's internal metadata service caused a cascading set of disruptions that triggered outages for many customers relying on AWS services in the US-EAST-1 region, including Airbnb, IMDb, and Netflix.[86]
Jan. 2016	A faulty update prevented Microsoft Office 365 users from accessing some email messages, a delay lasting up to five days for some users[87]
May 2016	A Salesforce U.S. data center went down for about one day, tracing its cause to a failed circuit breaker in another data center that caused a flood of traffic leading to disruptions for many customers.[88]
Feb. 2017	Fn AWS outage in US-EAST-1 region caused failures in many online platforms and Organizations, including Airbnb, Signal, Slack, and the U.S. Securities and Exchange Commission over a five-hour period. One firm later estimated that the downtime caused a loss of $150 million for the S&P 500 companies affected.[89]
Mar. 2017	A power failure leading to software errors caused issues with Azure's storage service, especially for customers in its U.S. East region. Azure restored full service in eight hours.[90]
Jun. 2017	Security researchers warned chip manufacturer Intel and others about the Spectre/Meltdown speculative exploitation vulnerabilities. They were kept secret for six months with the chip manufacturers working with major tech companies to implement a solution.[91]
Jan. 2018	The Spectre and Meltdown vulnerabilities became public, with CSPs working to implement software and hardware fixes.[92]
Jun. 2018	Azure customers in Northern Europe experienced a five-hour outage due to hot summer temperatures in a data center, leading to automated infrastructure shutdowns.[93]
Sept. 2018	Lightning strikes caused failure at an Azure data center in Texas, affecting customers using storage in the local region as well as some Azure services globally. The local region was offline for about four hours.[94]
Nov. 2018	The misconfiguration of an internet routing protocol by a Nigerian internet service provider caused failures for traffic to Google Cloud after traffic was mistakenly sent through China.[95]
Jan. 2019	Issues with an external domain name service provider caused errors in internal Microsoft Azure systems that lead to the accidental dropping of customer databases, which were later recovered.[96]

In this sense, cloud security thus far is a series of potential catastrophes narrowly averted, perhaps a positive indicator of CSPs' mitigation efforts and the general resilience of cloud service architectures. On the other hand, without an incident where a CSP truly failed, say one that is comparable to the WannaCry or NotPetya cyber attacks, the level of risk associated with environmental, operational, and adversarial threats is still unquantifiable.

Case Studies of Major Incidents

The eight case studies detailed in the appendix provide in-depth assessments of various risk factors affecting security in and of the cloud, as illustrated by some of the most significant publicly known incidents. They illuminate not just specific threats to cloud security but also how cloud security affects customers and the public writ large, as well as how CSPs have performed to limit their effects.

One main finding is that many cloud incidents are not caused by malicious adversaries but rather by human error as well as natural phenomena. In the first case, a mistaken command entry—a typo—by an AWS engineer set off a chain of inadvertent changes to a core AWS service in a region with heavy traffic that resulted in a nearly five-hour outage of key AWS tools. These tools supported many websites, leading to downtime affecting not just AWS but many of its customers worldwide.[97]

A similar chain reaction affecting globally distributed customers occurred in the second case, but this time because of a lightning strike resulting in a power surge to a Microsoft Azure data center. This took the data center offline and consequently caused a failure in legacy data management and authentication services that had data critical for their operations stored in the affected region. Complexities in the nested set of cloud-provided services thus made CSPs vulnerable to outages in critical regions, despite extensive efforts to build global networks and CSPs' insistence that each region is logically separate and independent from all others.

The cloud is also vulnerable to internet routing issues. In the third case, Google Cloud suffered disruptions because an external internet service provider in Nigeria mistakenly listed its address so Google Cloud traffic directed itself through that internet service provider, overwhelming it and causing connectivity issues worldwide. This incident underscores that the cloud remains vulnerable to familiar internet security problems, and perhaps even more so because of the sheer amount of data it sends over the public web.

Another key takeaway is that the complex structure of the cloud can cause cascading effects. For instance, in the fourth case, a number of critical datasets in Azure's SQL databases were accidentally deleted because of a complex cascade. It began with an outside domain name system (DNS) provider

error affecting Azure's authentication systems. Although Azure was able to provide customers with a copy of recovered data from five minutes prior to deletion, this incident illustrates how structural cloud incidents occur: minor technical issues can expose vulnerabilities in a CSP's internal operations that can have unpredictable effects. A similar process played out in the fifth case, when a deliberate tweak to Facebook's servers caused outages lasting up to fourteen hours for core services like Instagram and WhatsApp. Although less public information is available for this case, it appears similar to the first case. Facebook is not a CSP offering cloud services to third parties, but it operates very similarly to one because of its size and massive global network.

The vulnerability that is arguably the most significant thus far for cloud security is the one that has not yet resulted in any disruption whatsoever to cloud services: the discovery of the Meltdown and Spectre chip hardware vulnerabilities. As detailed in the sixth case, in 2017, security researchers discovered major vulnerabilities that would allow attackers to surreptitiously read data from many of the chips used in CSPs' data centers. Fixing these flaws required a massive effort not just to implement software patches but also to replace the affected hardware. While CSPs' responses to the Meltdown/Spectre vulnerabilities have been lauded as highly collaborative and proactive, these vulnerabilities still represent a significant what if scenario. If the vulnerabilities had been discovered not by security researchers but after a successful attack stealing data from the cloud, such a theft could have undermined confidence in CSPs and called into question the security of the cloud more fundamentally.

Adversarial attacks also have affected cloud security in high-profile incidents, notably the Capital One data breach. As detailed in the seventh case, the attacker, a former AWS employee, stole social security numbers and information about people who had applied for Capital One credit credits through an exploit accessing data stored by Capital One in the AWS public cloud. Like many adversarial attacks, this attacker made use of a misconfiguration by a customer to gain access, underscoring how security in the cloud is jointly managed by customers and CSPs. This event appears to be one of the most significant incidents of this kind, especially because Capital One was among the leaders in the financial sector in moving all its services to the cloud. In the eighth case, state-linked hackers conducted one of the most far-reaching intrusions against cloud companies and their clients, stealing valuable intellectual property and trade secrets from corporations around the world. Operation Cloud Hopper underscores that security in the cloud reflects the Fort Knox problem—while it can be much more secure, hackers like the Chinese actors suspected in this case were able to make off with the keys to a far bigger store of data than if they had just targeted one firm directly.

Lastly, thus far, these incidents do not appear to have had physical consequences beyond disruptions to internet of things (IoT)-connected services like smart lightbulbs. However, it is difficult to forecast the potential impacts of a cloud outage for a critical, real-time service like financial transactions or a complicated IoT system like a self-driving car network.

Mapping Risk Vectors More Systematically

This list of past incidents is the starting point for producing a more systematic mapping of risk associated with CSPs. Past cloud security incidents provide the basis for developing a list of risk vectors, which can be classified as accidental, adversarial, environmental, and structural. While the first three are self-explanatory, the fourth (structural) refers to incidents where the complexity of the architecture of the cloud itself produces risks. For instance, as some of the cases show, an automated error process can lead to inadvertent shutdowns of cloud services because of a complex interrelationship among the various internal elements of a CSP. In some cases, structural risks intersect with human error or environmental effects, but the mark of a structural risk vector is the exacerbation of a risk through internal failures and the unpredictable effects on cloud services that can result.

The framework outlined in figure 5 below applies the well-known cybersecurity triad of confidentiality, integrity, and availability to these risk vectors and includes rough guesstimates of the probabilities that various incidents will occur, ranging from more common incidents to potential black swan events. These probabilities are not intended as predictors but rather as a starting point for a discussion of how different risks could be classified that will hopefully be tested and improved with feedback from other experts over time. The notional probabilities were based on the authors' assessment of the frequency of past occurrences, with events that have not yet occurred being assigned lower probabilities.[98]

By dividing up risk vectors according to their potential effects on confidentiality, integrity, and availability, this framework connects risks to specific impacts while also underscoring that multiple risks could combine to create simultaneous impacts with unpredictable effects. This method further identifies how each risk impacts each element of the triad. For instance, a customer misconfiguration of data might lead to external data leakage, while a hardware vulnerability similar to Meltdown/Spectre could lead to customers maliciously accessing other customers' data—two outcomes which are both effects on different dimensions of confidentiality. Malicious data theft is the third way in which the vectors can affect confidentiality in the cloud.

As for availability, risks can be divided between those that cause temporary or permanent losses of availability, with environmental vectors responsible for a significant fraction of the former. Structural risks can also lead to second-order risks of availability such as outages in key regions causing failures of global services dependent on that region's availability.

Finally, threats to the integrity of data in the cloud segment into the deletion of data, the manipulation of data, and data asynchrony. The last of these terms refers to a potential situation where a dataset could exist in two or more different locations with different values for data that is supposed to be identical. Such asynchrony could result from accidents in a cloud incident that lead to inadvertent changes to data stored across different regions.

This risk framework points to a potential scaling of threat vectors based on their probabilities, which allows some assessment of the types of impacts that might be more or less common.

In the most likely category are environmental events that could temporarily disrupt availability of data or cloud services and accidental exposures of customer data to the public, due to misconfiguration. The case studies suggest that, although temporary unavailability can result from complex failures and potentially affect wide swaths of customers, CSPs are relatively adept at restoring services quickly in response to environmental or accidental incidents. For exposures of customer data, these are some of the most commonly reported cloud security incidents, but they may have a low impact, because even though the data were exposed, it is not commonly reported to be exploited by malicious actors.

In the more medium-to-low range of probabilities are adversarial threats to the confidentiality of customers' data in the cloud. These include incidents where the CSP itself may be compromised, like in the Cloudhopper campaign. Structural causes of unavailability also fall in this range of probabilities, these being the causes of what is noted as second-order effects of unavailability which may exacerbate temporary outages or lead to other effects. These may be incidents where accidental or environmental incidents at a local level end up having global effects. As a consequence, their impact is difficult to predict. There are also several medium-probability accidental threats that could lead to data deletion. Notably, all potential threats affecting data integrity are either medium or low probability. In other words, incidents resulting in the deletion or alteration of data in the cloud will not be common. The case studies support this tentative finding.

In the least likely set of threats are adversarial attempts by customers to steal data from each other: such scenarios include the risks that one customer may infringe on the confidentiality of another's data, threats to permanently disrupt availability (whether adversarial or environmental), and adversarial attempts to delete or manipulate data integrity. This observation is not to suggest that the cloud is immune to adversarial attacks. Rather, those attacks will have a low likelihood of success—especially those targeting the integrity of data as opposed to its confidentiality—because of the highly sophisticated security and resiliency built into cloud architecture by major CSPs to avoid exactly this scenario. This finding parallels broader industry trends: in cybersecurity circles, the theft of data is much more common than destructive attacks.

FIGURE 5

Mapping the Impact of Cloud Security Risk Vectors

EFFECTS ON		VECTORS	PROBABILITY
CONFIDENTIALITY	**External:** Unintentional data leakage	Customer misconfigures or does not enable security keys for stored data	Very High
		CSP misconfigures or does not enable security keys for stored data	Medium
		Vulnerability discovered in security protocols making data accessible to third parties or the public internet	Low
	External: Malicious data theft	Malicious actor steals credentials from customer to steal data hosted in the cloud	Medium
		Threat actor compromises CSP to steal security keys to access customer accounts	Low
		Insider threat at customer or CSP permits theft of data	Low
		Manipulated domain name system (DNS) or Border Gateway Protocol (BGP) routing information allows malicious actors to redirect cloud-customer traffic	Medium
		Installation of fake hypervisor through server compromise to exfiltrate data	Very Low
	Internal: From one customer to another	Misconfiguration of hypervisor or containers permits customers to access data of other customers	Low
		Exploitation of hypervisor or container vulnerability permits virtual machine escape	Low
		Chip or hardware vulnerability allows virtual machine/container escape	Low
INTEGRITY	Data deletion	Misconfiguration of automated process leads to deletion of virtual machines and stored data	Medium
		Automated process deletes datasets because of internal errors or unavailability	Medium
		Automated process or human error causes overwriting of datasets, losing information	Medium
		Adversary compromises CSP system managers and deletes large swaths of customer data	Low
		CSP internal systems infected with wiper malware or ransomware	Low
	Data manipulation	Error in automated process or human error causes alterations to replicated data	Medium
		Malicious insider within enterprise customer alters data for personal gain	Low
		Malicious insider at CSP alters data of single or many customers	Low
		Hacker steals credentials from user, gains access to data in the cloud, and alters it	Medium
		Compromise of CSP permits hackers to alter data across many customer accounts	Medium
		Man-in-the-middle threat between CSP and customer substitutes altered data to be stored in the cloud	Low
	Data asynchrony	Failures in availability lead to different copies of data in different CSP regions because of asynchronous geo-replication	Medium

KEY: Accidental Structural Adversarial Environmental

EFFECTS ON		VECTORS	PROBABILITY
AVAILABILITY	Temporary unavailability	Lightning strike on data center	Very High
		Flooding of data center	High
		Earthquake near data center	High
		Damage to power lines leading to failures of backups (from natural disasters)	High
		Accidental cutting of undersea or local fiber-optic cables	Medium
		Unintentional rebooting of all servers within an AZ or availability region	Medium
		Accidental deletion of a large number of virtual machines	High
		Use of incorrect configuration settings during routine upgrades leads to loss of availability	Medium
		Insufficient capacity of backup servers during routine maintenance	Low
		Internal automated or human errors during routine maintenance lead to internal traffic flood, causing denial of service	Medium
		Expiration of HTTPS certificates leads to authentication unavailability	Medium
		Misconfiguration of BGP or DNS information by outside providers for CSPs leads external networks to drop traffic	Medium
		Compromise of customer accounts to conduct cryptocurrency mining operations (cryptojacking)	Medium
		Distributed denial of service attack on CSP	High
		Intentional deletion of virtual machine or stoppage of services by insider	Low
	Permanent unavailability	Nuclear meltdown or accident renders data center inoperable	Very Low
		Bombing or other attack on data centers by terrorists or state actors	Very Low
		Intentional destruction of power grids leads to data center failure	Low
		Cutting of multiple undersea cables degrades international internet connectivity	Low
	Second-order effects of unavailability	Automated failure detection systems mask errors, leading to catastrophic failure and large-scale downtime	Medium
		Unavailability of core systems or other components delays efforts to restore system	Medium
		Unavailability of core systems leads to unintentional activities by automated systems, resulting in either deletion of virtual machines, dropping of databases, or other services going offline	Medium
		Unavailability in a key region leads to widespread downtime of key platforms that rely on services based in that region	Medium

Discussions of cloud security tend to either focus on common events (like the inadvertent exposure of customer data, primarily) or the most significant, destructive attacks. The focus on the latter often leads to a broader conversation about systemic risk and the cloud. When discussing this topic, it is important to be clear about which system is being referred to. First, there could be potential systemic risk to major CSPs if their entire systems could be disrupted worldwide. So far, however, there has never been a case of this happening, and it appears extremely unlikely. Second, there could also be systemic risk to critical industries if critical data or services using the cloud were to be disrupted or destroyed. As the framework in figure 5 illustrates, it would not be necessary for an entire CSP to be disrupted for such an incident to occur.

Regarding a second form of systemic risk, drawing the pathway between a specific threat vector and a consequence and then assigning it a probability—the definition of risk—is not possible without reference to the function of whatever data or service in question is being held in the cloud. That is, while threats to the cloud may be data agnostic, their impacts are not. In fact, a similar threat might have vastly different consequences across incidents depending on which customers and which specific data or services are affected.

This insight leads to the potential conclusion that a sector-specific approach would be better for assessing systemic cloud risk to critical industries. One example of this is a 2018 Bank for International Settlements Report on cloud risks for the insurance sector, which discussed risks in terms of specific systems and data, as in the following: "Business continuity and operational resilience may be compromised if . . . insurers and their supervisory authorities are restricted from accessing, auditing, and making on-site examinations to cloud providers located in different jurisdictions."[99]

Regarding systemic risk to the cloud, one starting point for that assessment is to examine which points of vulnerability are shared across the cloud ecosystem and where some of the potential threats could have a widespread impact. The hardware (server) and hypervisor layers are obvious candidates. Indeed, the Spectre and Meltdown vulnerabilities are prime examples of this type of risk, one that was thankfully averted before it was exploited (at least according to publicly known accounts). Given the diversity of approaches higher up the stack among various CSPs, different customers, and various applications of the cloud, it is much harder to imagine potential systemic vulnerabilities manifesting in a similar manner. Like the cloud writ large, this diversity of approaches then makes it more imperative to focus defense on hardware and hypervisors, especially on potential supply chain risks.

Hypervisor Vulnerabilities

Hypervisors are especially important to the security of CSPs.[100] Whereas other impact vectors such as lightning strikes or ransomware affect many other infrastructures as well, hypervisors are of critical importance for the cloud, managing the connection between its virtual machines and the allocation of resources to them. CSPs must be able to trust their hypervisors to the same extent that they trust the underlying machines, if not even more so, as hypervisors manage the isolation of guest virtual machines. Therefore, vulnerabilities in hypervisors are a crucial issue for CSPs to address. The bad news is that, like all software, hypervisors have vulnerabilities.[101]

For instance, winners of a 2019 white-hat hacking competition found several vulnerabilities in virtualization software from VMware and Oracle—including one that allowed code execution in the hypervisor, which could essentially allow a threat actor to escape the confines of a virtual machine to run the overall server.[102] Generally, such techniques are referred to as virtual machine escape. One possibility is that an attacker could install a fake hypervisor to take control of an entire server system—a practice called "hyperjacking."[103]

The good news is that there are no known instances of attacks having exploited hypervisors. What is known is that the Spectre vulnerability and another reported vulnerability, one called virtualized environment neglected operations manipulation (VENOM), could have enabled such exploits.[104] And at least one research article[105] discusses hypervisor vulnerabilities in cloud service environments, setting up a demonstration environment and carrying out a number of classic attack types against it.[106]

The biggest blind spots in the current debate are medium-to-low probability risks. These include structural risks that result from highly complex, tightly coupled systems that could be significantly damaged by accidents or environmental threats, as well as highly sophisticated adversary threats against data confidentiality. The Cloudhopper campaign is a potent example of the potential impact and significance that the compromise of a CSP could have on its customers. As more and more data that is important but perhaps less than critical is located in the cloud, threat actors will have greater incentives to engage in more of these types of campaigns, especially capable government-linked hacking organizations.

Of course, mitigating these potential risks is highly challenging. States are able to employ advanced persistent approaches to find their ways into even the most secure of targets. Criminal organizations have also demonstrated sophisticated methods in the theft of data. And in many cases, the complexity of the cloud works against its security. Neither a CSP nor a customer may have full visibility into the joint architecture of their systems and thus neither might be able to rectify potential structural vulnerabilities before they cause failures.

Moreover, the level of risk depends on what is in the cloud. Identifying potential impacts is simply not possible on a general basis. In broad strokes, as more important data and services are located in the cloud, and as certain organizations make their entire operations cloud-based, these low-to-medium-probability risks will grow in consequence.

Whose Burden Is It? Shared Responsibility Models for Security

As organizations transition parts of their functions to CSPs, their security teams have to update their roles and responsibilities and prioritize accordingly. Some responsibilities will migrate to CSPs to free up the bandwidth of customer IT teams, while other responsibilities will emerge with respect to managing the relationship between a given client and its CSP. That is why the major CSPs have developed shared responsibility models as an educational tool to provide guidance to clients (see figure 6). It is this interface between a given CSP and its customer that presents a new technical and human vulnerability, partly because of the technology's novelty. For example, the Capital One data leak was the result of a misconfiguration that fell into this gray space of shared responsibility.

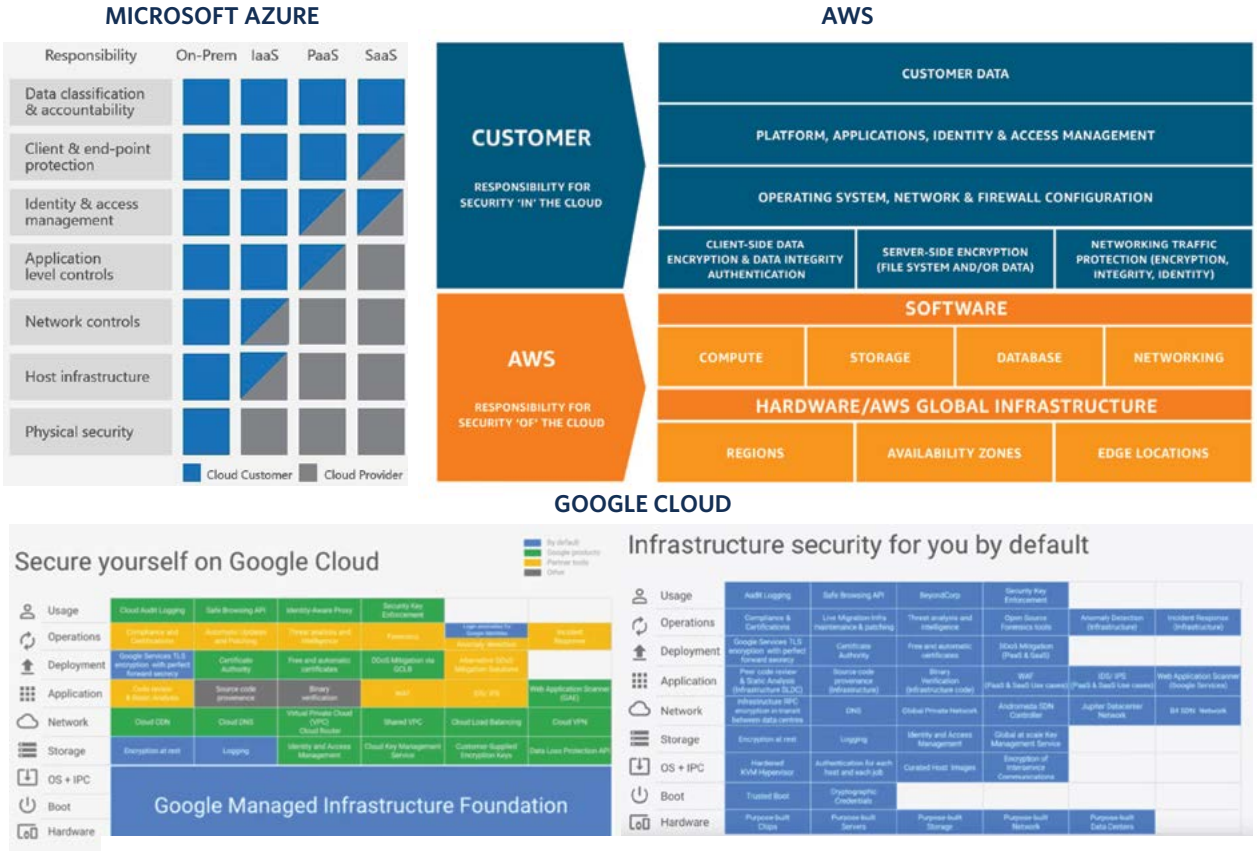
There are a lot of commonalities between the various shared responsibility models that major CSPs have formulated to delineate how they break down various aspects of security responsibilities between them and their clients. Figure 6 shows how they draw similar delineations between CSPs and customers. Variations tend to be small. For example, the AWS model does not display the different deployment model (IaaS, PaaS, SaaS) compared to the Azure illustration. Azure takes a layered approach, defining responsibilities in terms of functions rather than systems. The AWS model more closely tracks the Azure IaaS column, although it does not have the shared functions between the CSP and its customer, as Azure shows. In the AWS approach, there is some CSP responsibility for similar functions, but they are shown separately. Google Cloud has produced a slightly different conceptualization that provides good insights into specific security activities but does not delineate CSP and customer responsibilities as clearly, while the logic remains the same.

This shared responsibility cannot be avoided given that migration to the cloud requires that customers involve and trust CSPs as a third party. That is why a clear understanding of the roles and responsibilities of each side is paramount for avoiding accidents and preventing the interface from

becoming a gateway for nefarious actors. The Capital One incident has illustrated the reputational cost for CSPs and their clients that accompany any shortcomings in this area. It is therefore in the self-interest of CSPs to provide assistance to their customers to minimize these risks. This is also an area where government agencies might study how to ensure that such assistance is provided and that in the unequal relationship between most customers and their giant CSPs, the burden (and potential blame if something does happen) does not shift incrementally or otherwise to the customers.

FIGURE 6

Shared Responsibility Models of Major CSPs



Cloud Cyber Incident Severity Framework

The comprehensive framework outlined above in figure 5 details a variety of incidents that could affect CSPs as well as rough estimates for the probabilities associated with them. Yet, as discussed above, it is impossible to directly connect potential threat vectors to impacts. With the variety of data and services held in the cloud, threats could potentially have next to no impact or be some of the most damaging cybersecurity incidents to ever occur. Indeed, similar to other complex systems, policymakers may only be able to assess the source of a systemic cloud incident after the fact, a reality that underscores the need to focus on response and recovery rather than prevention.

Intended to be of more utility to policymakers concerned about the impact of incidents, table 6 goes beyond the technically informed framework outlined in figure 5 by detailing a framework to assess the severity of a cloud-based incident. This framework is based on the general cyber incident severity schema developed by White House staff in 2016 for Presidential Policy Directive 41.[107]

In the lower categories (white through yellow), only limited threats to either cloud confidentiality or availability occur, similar to past public incidents. Incidents rise in severity with their impact depending on which specific types of data and cloud customers are affected. For instance, a hypothetical misconfiguration that were to expose the data of a record label may be embarrassing and costly from a reputational standpoint, but it would pale in comparison to a data leak involving official government data for a vital purpose such as security clearance reviews. Of course, this all depends on the extent to which such services, especially those critical for immediate human needs like healthcare providers, depend on the cloud. Greater societal dependence on the cloud raises the potential for risks to rise up this severity scale.

Lastly, the example incidents underscore that threats to data integrity as well as permanent losses of data availability may have the largest potential impacts. Such risks have greater potential to trigger the types of physical consequences and loss of confidence that the most dire analyses of cloud incidents warn of.

The utility of this framework will depend on CSPs and their customers. If it is used by both to assess potential threats to critical data and services, it could be a way for both parties to understand the impacts of various incident scenarios. It also could be a jumping-off point for future efforts to better specify which types of cloud security issues could have the most significant impacts and then how to mitigate those threats.

TABLE 6

A Severity Framework for Cloud-based Cyber Incidents

LEVEL 5 - EMERGENCY Poses an imminent threat to the provision of wide-scale critical infrastructure services, national government stability, or to the lives of people	● A cloud-based network used by hospitals and healthcare workers for patient data and management suffers a critical outage, leading to an inability to care for patients effectively. ● Cloud infrastructure supporting core internet routing is disrupted, causing internet-wide outages and preventing both private and public critical infrastructure from functioning.
LEVEL 4 - SEVERE Likely to result in a significant impact on public health or safety, national security, economic security, foreign relations, or civil liberties	● Malicious insiders release security keys for a large store of classified data stored in a government cloud service, disclosing information about intelligence and military personnel and operations. ● A major cloud-based service used by the financial sector is compromised and leads to the corruption of the integrity of critical financial data, causing a worldwide crisis of confidence. ● An electricity management system that relies on cloud infrastructure is disrupted because of attacks on the cloud, leading to its destruction and widespread power outages.
LEVEL 3 - HIGH Likely to result in a demonstrable impact on public health or safety, national security, economic security, foreign relations, civil liberties, or public confidence	● Misconfigured security controls for census data stored in the cloud allow attackers to alter and delete census records, throwing the accuracy of the census count into question. ● A cloud-based database and logistics system for a global shipping firm is destroyed by a malicious attack that deletes critical data, causing major disruptions to international shipping. ● Critical backups of financial records for a major insurance company stored in the cloud are deleted, causing losses and instability in global markets.
LEVEL 2 - MEDIUM May impact public health or safety, national security, economic security, foreign relations, civil liberties, or public confidence	● An automobile manufacturer suffers business disruptions because of a lengthy (24+ hours) cloud outage, impacting production. ● A cloud service disruption causes interruptions to nonessential government services, such as social programs and public access to government databases. ● A lengthy cloud outage disrupts most major online communication platforms, causing public confusion.
LEVEL 1 - LOW Unlikely to impact public health or safety, national security, economic security, foreign relations, civil liberties, or public confidence	● Temporary cloud outages in key regions cause short disruptions or slower service on internet platforms like Google's search engine or Slack. ● Misconfigurations in internal cloud security protocols expose data from some customers to others in the same cloud service, revealing some business information.
LEVEL 0 - BASELINE Unsubstantiated or inconsequential event	● An attempted distributed denial-of-service attack (DDoS)a gainst a major CSP that does not affect service to customers. ● Data center partially disrupted after lightning strike but CSPs are able to shift resources and mitigate its effects.

NOTE: The authors created this table, drawing on a previously published White House document. See "Cyber Incident Serverity Schema," White House, July 26, 2016, https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Cyber%2BIncident%2BSeverity%2BSchema.pdf.

Chapter 4: Additional Public Policy Issues to Consider

The focus of this paper is on the public policy implications of cloud security. There are other important, relevant public policy issues for which extensive treatments are beyond the scope of this paper, namely 1) data governance, 2) the connection of technology industries and geopolitical influence, and 3) antitrust regulation. Some of these issues have direct relevance to the suite of security policy issues this paper discusses, including policy debates about data localization and critical infrastructure, while others like antitrust considerations are indirectly related.

Data Governance

The field of data governance encompasses a large set of issues involving the control, access, protection, and regulation of personal and commercial data by both technology companies and governments. As huge amounts of both individual and business data have started being stored and processed in the cloud, this transition has raised questions about the access that governments can exercise to data in the cloud, particularly for CSPs that store data overseas.

The nature of cloud services is such that data belonging to one country's citizens or government may be stored in different countries and jurisdictions. This state of affairs then poses data governance issues: Who makes the rules regulating that data? Should it be the country where the data is stored (keeping in mind that cloud data may in fact be sharded, or split up and simultaneously stored in different data centers) or the country where the data's owners are located? [109] The latter option raises problems due to the application of law extraterritorially, particularly given the deficiencies in mutual legal assistance frameworks for digital information. Yet, if that were not done, there would be serious challenges to legitimate lawful access requests in the United States and other countries with strong rule of law. On the other hand, expanding lawful access could enable countries with weaker human rights records to conduct illegitimate surveillance and exploit the sovereignty of other countries where privacy values may differ.

In March 2018, the United States passed legislation to clarify that the Stored Communications Act, which allows law enforcement to issue warrants or court order to access data held by service providers, applies to data held overseas. This legislation, the Clarifying Lawful Overseas Use of Data (CLOUD) Act, also enabled the United States to conclude agreements with foreign countries that meet a set of privacy protections to allow their law enforcement agencies to request data from U.S. service providers.[110] While the CLOUD Act did not solve all issues associated with cross-border data access, it did provide some greater clarity for CSPs and remediated some potential situations where, to comply with the law of one jurisdiction, they might violate the law of a different jurisdiction.

Similarly, in February 2019, the United Kingdom (UK) passed legislation empowering law enforcement to issue "overseas production orders" to providers to compel them to provide electronic evidence, regardless of location.[111] Both pieces of legislation are significant in that they extend the reach of legal authorities beyond national boundaries and attempt to adapt electronic evidence law to the new reality of global CSPs with data distributed in many jurisdictions.

The wide reach of foreign governments has long been a major data governance issue and a key concern for many countries. Revelations about the surveillance capabilities of U.S. intelligence agencies and their close cooperation with U.S. tech companies, as disclosed by former intelligence contractor Edward Snowden in 2013, led to profound mistrust of U.S. tech companies overseas. Since the major CSPs (AWS, Google Cloud, and Microsoft Azure) and other tech companies with major cloud offerings (Salesforce, Dropbox, and IBM) are all based in the United States, other countries have sought to put protections in place to limit how those companies use the data of their citizens. Concerns about foreign surveillance contributed to the European Union's (EU) General Data Protection Regulation (GDPR) in 2018, which applied EU data protection standards to any data related to EU citizens, even if located outside the EU. Depending on the specifics of their arrangements, many CSPs qualify as "data processors" for their European customers, meaning they must comply with GDPR restrictions.[112]

Concerns over access to content and the implications for human rights also affect CSPs entering markets in regions that include authoritarian governments. Some governments are widely known to conduct surveillance of human rights activists and dissidents and to seek to use technology companies to enforce censorship and take down politically controversial content.[113] Nevertheless, cloud firms appear to be moving forward with expanding their presence in authoritarian and authoritarian-leaning countries.[114] In 2017, AWS announced it would open its first region in the Middle East, in Bahrain, to come online in early 2019, while also saying it would have an edge network location in the United Arab Emirates (UAE).[115] Microsoft Azure followed suit with regions in Abu Dhabi and Dubai.[116] A month later, Google's quarterly earnings report confirmed[117] that it would be building a region in Saudi Arabia, bringing its global total to twenty. However, as of June 2020, this region is not yet online, though there have been reports that Google is still pursuing it.[118]

Reports about AWS investment in Saudi Arabia are also dated. It is unclear to what extent the murder of Saudi journalist Jamal Khashoggi and a high-profile dispute between Bezos and Saudi Arabia's Crown Prince Mohammed bin Salman have influenced this state of affairs. Major Western cloud firms appear to be attempting to balance the human rights concerns of doing business in the Middle East with the potential profits from partnering with wealthy governments in a growing market.

Data Localization

Data localization requirements have also been on the rise around the globe. From domestic to foreign companies, some governments are increasingly requiring CSPs to store data on citizens of a jurisdiction in that same jurisdiction. For instance, India has either passed or proposed a number of data localization requirements, including a rule that all payment data on financial transactions in India be stored in the country, a development that has sparked pushback from major financial firms and the U.S. government.[119] South Korea and Brazil have also passed measures requiring data localization in specific circumstances.

The most significant localization requirements, however, are in China. The Chinese government requires that "important data" on Chinese citizens be located in China, as well as data related to critical "information infrastructure."[120] These vague terms allow Chinese authorities to impose requirements for localization on many large foreign companies by mandating, for instance, that they work with local partners to use Chinese data centers.[121] Western CSPs have by and large complied with these requirements (like AWS and Microsoft Azure did) or have left the Chinese market (as Google did).

Data localization has significantly impacted how CSPs do business in China, but the push for data localization has also led to different models for cloud services elsewhere. One of these models is Microsoft Azure's cloud offering in Germany, which employs a trustee model to meet concerns about U.S. spying. In 2015, in the immediate aftermath of the Snowden revelations, Microsoft announced it would open a cloud offering in Germany that would be run through a data trustee, T-Systems. T-Systems, a subsidiary of Deutsche Telekom, controlled the data stored in Germany, and Microsoft would only be able to access the data with the permission of both the customer and T-Systems.

This arrangement appeared to be a solution to some of the European concerns about U.S. CSPs, but it did not last. In September 2018, Microsoft Azure announced it would not be accepting any more new customers to Azure Germany and said it would open two regions that would be part of its global cloud network in Germany.[122] According to reports, Azure Germany was not successful because of higher costs and the segmentation from international business.[123] The failure of this Germany-based cloud offering points to some of the advantages that made cloud services a global business: their low prices and ability to connect and serve customers across borders, which localization requirements cut against.

Tech Companies and Geopolitical Influence

The landscape of major CSPs divides along geopolitical lines. U.S.-based firms (particularly AWS, Microsoft Azure, and Google Cloud) dominate the global cloud market. The only major alternatives are Chinese CSPs (Alibaba and Tencent). Such a divide has geopolitical ramifications in the larger context of U.S.-China technology competition, especially in other countries where U.S. CSPs are competing with Chinese CSPs for market share. Because cloud computing can provide services crucial to developing a technology industry, the decisions about where CSPs make investments and where they open data centers, AZs, and service regions are often intertwined with geopolitical issues.

The burdens that China has placed on foreign CSPs effectively prohibited them from exerting the same degree of market dominance earlier this decade as they did elsewhere—in Europe, for example. As an official from the U.S. Information Technology Industry Council stated in 2019, U.S.-based CSPs "face written and unwritten requirements that do not allow foreign companies to obtain licenses to operate without a Chinese partner; force U.S. CSPs to surrender use of their brand names; and require companies to hand over operation and control of their businesses to Chinese companies in order to do business in the Chinese market."[124] These restrictions likely enabled China's CSPs to dominate the local market and to become large enough to compete with U.S. CSPs, as they are doing in regions such as East and Southeast Asia and the Middle East.[125]

China's influential role also extends to the undersea cable ecosystem. Huawei was previously a 51 percent stakeholder in Huawei Marine, one of the Big Four cable-laying companies (the others being the U.S.-based SubCom, the Finnish-owned Alcatel, and the Japanese-owned NEC Corporation).[126] Huawei Marine has grown its share of major cable projects and is working on a Europe–Middle East–Africa cable, although Huawei has not worked on a U.S. cable project since 2013. It is no surprise that Western officials have become concerned about China's growing influence in this area. In 2017, Australia intervened successfully to convince the Solomon Islands not to allow Huawei to build a cable from the island nation to Sydney.[127] Yet, in 2018, Australia, Japan, and the United States failed to convince Papua New Guinea to back away from a deal with the company.[128] In 2019, Huawei sold its stake in Huawei Marine to another Chinese company, after Huawei's role in constructing undersea cables received similar scrutiny as efforts to ban Huawei from 5G infrastructure in various countries.[129]

This competition has broader implications as countries face a choice of whether to 1) embrace U.S.-based CSPs and accept the U.S. government's potential access to the data, 2) embrace Chinese CSPs and face the risk posed by the Chinese government having access to that data, or 3) go without the security and economic benefits of cloud services at all. These geopolitical issues related to cloud investments are therefore not confined to U.S.-China competition and merit further research.

Antitrust Considerations

Antitrust regulation has emerged as a major issue in technology policy as well. The EU has fined major U.S. tech companies for violating rules to protect competition, including levying huge fines against Google for its advertising practices.[130] While European regulators have been active in fining large technology companies for violating competition law, U.S. regulators have not yet brought any actions against big tech firms. However, there are numerous calls for the Department of Justice and the Federal Trade Commission to be more aggressive in their antitrust oversight of the Big Four tech companies—Apple, Amazon, Facebook, and Google. U.S. Senator and former presidential candidate Elizabeth Warren went further, proposing to "break up Big Tech" by designating large tech companies as "platform utilities" and separating their core services from their other businesses.[131]

Most antitrust discussions focus on activities by large tech companies not related to cloud services, such as Amazon's strategy of undercutting competitors by selling products at a loss on its online marketplace. EU regulatory activities have also thus far not targeted U.S. firms for their cloud businesses. However, proposals to break up large firms would affect their cloud businesses, especially as cloud services contribute a large fraction of profits for Amazon and Microsoft. The cloud services market itself does have significant competition between the large CSPs as well as smaller firms. In many cases, one potential concern is that, because so many businesses rely on cloud services, competitors of the leading CSP firms in other lines of business are their cloud customers, creating a conflict of interest.[132] Thus, some have called specifically to separate AWS from Amazon's online marketplace and other offerings.[133]

However, there may be strategic benefits to preserving cloud businesses within larger conglomerates like Google and Amazon. Particularly since governments now seek to partner with commercial cloud providers to develop their own clouds, such as the U.S. Department of Defense's JEDI project, only large corporations with the resources to develop their own global infrastructure may be sufficiently able to meet national security requirements. As some have argued, the singular demands of the U.S. government may require singularly large corporate partners.[134]

In conclusion, while this paper focuses on security, these other three issues cannot be ignored and are intertwined with each other. For example, questions about the ability of the U.S. intelligence community to access data of CSPs are a major security concern to other countries. And competition between U.S. and Chinese CSPs parallels security concerns about U.S.-China technology competition. Public policy decisionmakers grappling with the implications of migrating to the cloud for security reasons must therefore also bear in mind these other dimensions to make informed, comprehensive, risk-based decisions.

The Cloud Is in Need of Protection

The rise of cloud computing and cloud storage has transformed the cybersecurity landscape. Migrating to the cloud is a solution to many persistent problems of cyber insecurity for organizations that previously managed their IT systems alone, yet it also simultaneously presents a new set of highly complex, globe-spanning security issues. As this ecosystem evolves, policymakers need to improve their own grasp of the critical policy issues affecting (and affected by) the cloud; security is chief among them.

This primer attempts to outline some key concepts that bridge the technical and policy worlds. For all the appeals of analogies, the cloud as a term obscures more than it explains. The cloud is neither unitary nor inaccessible to malicious actors. The diversity of cloud services, deployment modes, and platforms and their omnipresence in modern life should underscore the importance of a nuanced understanding of this ecosystem.

In place of a vague evocation of an ethereal place where data goes and is dispensed from on high, it is helpful to disaggregate parts of the cloud and speak of more granular systems. Figure 1 in chapter 1 offers one approach, categorizing various elements of a cloud architecture according to a layer-based model, so that one might speak of the physical layer (data centers) as opposed to the hypervisor or cloud API layers. Such disaggregation helps break down potential incidents as well as the regulatory impacts.

But a technical understanding will not suffice for policymakers looking to assess the impacts of cloud services. A broader understanding of the market dynamics described in chapter 2 is necessary. The most salient fact that policymakers should note is that the market for public cloud infrastructure has become remarkably concentrated in a handful of major technology firms. While this may change in

the medium- to long-term, as a result of technological innovation or increasing competition between U.S. and Chinese CSPs for emerging markets, to cite a couple of examples, it will remain a dominant feature of the ecosystem and risk landscape at least in the short term.

Clearest among the findings of the section dedicated to security is not a deficiency in the security of cloud services, but a positive one for the broader IT ecosystem: a migration to the cloud will help a majority of organizations address the existing problem of cyber insecurity. With that said, certain factors set cloud services apart from other security challenges—they are inherently networked, concentrated, and shared. While there have been many different types of security incidents in the cloud, conceptualizing security risks in the cloud more comprehensively remains an underdeveloped area of study.

As policymakers further assess the security of the cloud, they should be wary of simplistic analyses that discount security threats to CSPs and those that warn of disaster should a CSP supposedly fail. As is the case with most cybersecurity problems, the truth is somewhere in between. This does not mean that cloud security issues should not be prioritized. Indeed, the fact that little is known about the potential impacts highlights an urgent need to study this problem and to identify specific problem sets where further action, either from the private or public sector, would improve security.

To better conceptualize cloud security, figure 5 in chapter 3 presents a technical approach, centered on the CIA framework, which highlights the confidentiality, integrity, and availability of cloud data and services. It highlights how risks could be differentiated based on their probabilities and their likely technical impacts. One key takeaway is that, while there are relatively common threats like attempts to breach the confidentiality of cloud data, what remains understudied and underdiscussed are those medium-to-low probability events that could cause significant but not catastrophic costs. More work is needed to understand the potential impact of these types of events and what interventions might be more successful at preventing them.

For policymakers focusing not only on the current cloud security environment but considering potential future scenarios, the severity scale presented in table 6 of chapter 3 is designed to provide a starting point for assessing cloud incidents based on their impacts. One way to connect the technical approach of figure 5 with the impact-based approach of table 6 is to assess how incidents might affect the various layers of the cloud architecture as well as key markets and critical sectors. Again, further research is needed, particularly to assess systemic risks by layer, but the work of conceptualizing cloud risks for specific sectors has already begun. The aforementioned 2018 Bank for International Settlements report that looks at the cloud and the insurance sector is a notable example.[135]

Ultimately, the key public policy challenge remains: How do practitioners and policymakers maximize the cybersecurity benefits provided by CSPs and how do they minimize their risks?[136] The former consideration is clearly driven by security and market incentives to migrate to the cloud. The latter is one where there will likely be additional regulation. This paper has remained agnostic about particular regulatory solutions to some of the issues it has outlined. This is a topic for future detailed study, particularly when it comes to understanding how existing regulatory frameworks already apply to cloud services. Of course, any future regulation would also have to take into consideration the other public policy issues that chapter 4 outlined in brief. Striking a balance between various equities—such as utility, flexibility, privacy, and security—will be necessary.

This paper's discussion of cloud security highlights three trade-offs that policymakers will have to address when considering cloud security: everyday versus systemic risk, the benefits and costs of sector concentration, and general versus sector-specific approaches.

Everyday versus systemic risks: Cloud companies' security teams deal with attempted intrusions on a daily basis. Their customers also configure their cloud-based data and services to prevent them from being inadvertently exposed to common criminals. Dealing with these and other daily challenges likely consumes most of the focus on security at CSPs. However, the potential remains for a future security incident to have a catastrophic impact on the entire cloud-based ecosystem. Planning for how to deal with these different types of risks will require different strategies and competing prioritization of resources, time, and focus.

Benefits and costs of concentration: The concentration of market power among major CSPs confers important security benefits similar to the logic behind Fort Knox. One particular benefit worth highlighting is their unrivalled ability to attract talented security personnel who are knowledgeable about securing data and processes across threat vectors and in very rare supply. On the other hand, failures or security incidents affecting a cloud provider could create systemic consequences. Importantly, the benefits of CSPs' size are felt every day, while the potential risks have not yet manifested.

General versus sector-specific approaches: As this primer has shown, the impact of a cloud security incident usually depends on what type of data or service is affected. Thus, the most suitable potential regulatory requirements with respect to security may differ across sectors that deal with different types of data—from the highly sensitive, fast-moving data common in the financial sector to the more privacy-sensitive personal data used by medical service providers. However, crafting regulation on a sector by sector basis would likely create conflicting requirements and incomplete standards.

These are not the only tradeoffs policymakers face with respect to the cloud. As this conclusion has emphasized, more research is needed to answer many critical questions about cloud security. To that end, following this section is a list of several research topics that should be explored in future research. These are a starting place for future efforts to build on the frameworks sketched out in chapter 3.

As the task of securing cloud increasingly becomes synonymous with cybersecurity writ large, policymakers must better understand and consider the impact of the cloud. A failure to pay attention could lead to a repeat of the cyber insecurity problem on an even bigger scale, with more severe consequences than seen to date.

Promising Areas for Future Security-Related Research

In that spirit, there are a few areas for future security-related research that should be highlighted.

Transparency on the use of CSPs: A key challenge for policymakers remains the lack of transparency about which companies are relying on CSPs and for what purpose. As a Reuters article points out, "Businesses generally are not required to disclose their cloud vendors."[137] The article points out that the reasons for disclosing the use of a CSP is usually a "passing mention" followed by "corporate risk." Whether or not governments should require greater transparency about companies' use of CSPs is a question that requires further research. For example, a risk associated with such a transparency requirement is that it would increase security risk in a scenario where a CSP is intentionally targeted to get at a customer. However, a transparency requirement will likely also improve risk mitigation policies.

Protecting undersea cables: Undersea cables have become critical for ensuring access to data on an ongoing basis.[138] Yet the undersea cable infrastructure that CSPs rely on and its vulnerabilities are a neglected international policy challenge.[139] This host of interrelated issues includes reliance on a single cable or a few cables to provide access to individual countries and the concentration of cable landing sites within single countries.[140] While companies that operate submarine cables often conclude agreements with each other to reroute data in event of failures, it is unclear if this system would be effective for a significant fraction of cables.

Various ideas have been proposed to enhance the security of undersea cables, including a subsidized insurance model for CSPs that meet certain government-defined security guidelines.[141] If they meet such guidelines, based on the idea of redundancy and models for threats to undersea cables, CSPs could get a certification as a recommended CSP. In return, the companies could pay for government-subsidized disruption insurance. The government would cover damages resulting from catastrophic disruptions like natural disasters and terrorist attacks, but CSPs would still have to pay for damages

from ordinary causes like human error. Further research on the potential downsides of such an arrangement is necessary, such as its potential effects on the cloud market. Other ideas include governments identifying a single point of contact for undersea cable issues and for the United States to ratify the United Nations Convention for the Law of the Sea (UNCLOS) to bring its legal position in line with those of its allies.

Government in the cloud: Another area that merits further dedicated research is how governments use commercial CSPs. A growing number of governments have embarked on efforts to move the operations of some government agencies to the cloud. The U.S. government, for example, adopted its cloud first policy in 2010, and the UK government did so in 2013.[142] In 2019, the U.S. government renewed its commitment with its updated "cloud smart" policy.[143] A 2019 report from the Government Accountability Office (GAO) found that federal agencies had made progress in cloud adoption and that fifteen of sixteen agencies examined had identified significant benefits from acquiring cloud services.[144] This transition raises some issues unique to government data and operations, such as how to protect confidential or classified data or how best to comply with legislative procurement requirements.

The U.S. Federal Risk and Authorization Management Plan (FedRAMP) is a program for certifying CSPs to serve federal agencies.[145] It defines baseline security authorization requirements for CSPs to become certified by either a Joint Authorization Board (composed of experts from the Departments of Defense and Homeland Security and the General Services Administration) or specific agencies. The UK government has developed a similar forum called the G-Cloud framework.[146] Other countries like Australia, Germany, and Singapore have all migrated some operations to cloud services.[147]

In the United States, Congress has also encouraged cloud adoption by federal agencies through legislation, such as the 2014 Federal Information Technology Acquisition Reform Act, which explicitly aimed at migration to the cloud. There is now more focus on oversight of federal cloud adoption, including a March 2020 Congressional Research Service report that identified several further options for Congress.[148] These included holding hearings focused on the Office of Management and Budget, which oversees the management of the Cloud Smart policy; reviewing individual agencies' cloud plans and implementation assessments; and tasking the GAO with assessing agencies' progress.

CSPs have responded to government interest by offering specialized services fitted to particular government needs. In the United States, both AWS and Microsoft Azure offer the Government Cloud for federal, state, and local agencies, separate from their commercial services. These offerings have even adapted to meet requirements for national security–related data and operations. For instance, in 2013, the Central Intelligence Agency (CIA) contracted with AWS to provide services to the U.S. intelligence community through a private cloud built by AWS.[149]

Further research is needed to assess the tradeoffs and policy goals of employing cloud services for government agencies and data. For the United States, assessing the efficacy of the FedRAMP program and any potential additional legislative requirements for cloud security is an important area for further study.

The JEDI contract: major trend in cloud security in recent years has been a move toward having multiple cloud vendors for different functions as opposed to a single vendor fulfilling all of an organization's cloud services. In a May 2019 survey of business tech professionals, almost 70 percent said they already are or are planning to use multiple cloud providers.[150] Another survey, the *Flexera 2020 State of the Cloud* report, found that 93 percent of respondents had a multicloud strategy.[151] Proponents argue that this approach has benefits such as avoiding vendor lock-in, increasing resilience to outages, and taking advantage of competitive pricing.[152] However, each of these points is contested—for instance, some might argue that ensuring a CSP's infrastructure architecture is secure is a much better way to enhance resilience than replicating workloads across multiple CSPs. Nevertheless, the appeal of multicloud arrangements is clearly growing.

However, using multiple cloud providers can create greater complexity for organizations in terms of managing their cloud usage and creating more potential points of vulnerability, as the Cloud Security Alliance discussed in a May 2019 report.[153] This report also noted that, in addition to using multiple cloud providers, many organizations are also using a mix of private and public cloud offerings, adding to the complexity of their respective cloud environments.

The controversy over the Pentagon's cloud contract for JEDI provides an illustrative case study about the benefits and potential downsides of single versus multicloud strategies. This project intended to kickstart the Pentagon's migration to public cloud services to link and support across the U.S. military's IT environments. The contract for this project was widely anticipated because of the Defense Department's massive size. In July 2018, the Pentagon announced that the contract would have a ceiling of $10 billion over ten years and would be offered to only a single vendor.[154]

A Pentagon official mentioned the following as a key reason why the contract was written for a single vendor: "The lack of standardization and interoperability today creates pretty significant barriers to accessing our data where and when it is needed, especially at the tactical edge on the battlefield. . . . We believe that multiple award cloud would exponentially increase the overall complexity."[155]

Amid disputes between the CSPs over the contract, competitors of the presumed favorite, AWS, argued against a single cloud decision. Oracle even went so far as to file a lawsuit that in part relied on arguing that a single cloud would not be in the best interests of the Defense Department.[156] Experts closely tied to defense contractors and Pentagon leadership argued in response that adopting a multicloud strategy for JEDI would simply be too complex for a bureaucracy as large as the Pentagon's to handle.[157] In the future, U.S. government cloud contracts will likely trend toward multicloud, as this is the direction many larger organizations are moving in. The CIA said in April 2019 that it is planning to follow up its initial intelligence cloud contract with a new procurement for multicloud services that will be in the "tens of billions of dollars."[158]

The status of the JEDI contract is currently in doubt because of lawsuits and concerns over political motivations, amid perceptions that U.S. President Donald Trump may have impacted the government's decision on which vendor to use: Microsoft Azure or AWS. While this delay may adversely impact national security, it provides additional time to assess the costs and benefits of single versus multicloud approaches. This question is particularly acute for governments considering moving key agencies and data to the cloud, as it touches on many of the issues about security and resiliency raised by the JEDI example. Further study could address potential ways for government procurement processes to be modified to better capture the risks and benefits of deciding on a single versus a multicloud approach, how organizations decide how to divide among different cloud providers and how they decide how many cloud providers to work with, and whether the trend toward multicloud approaches could encourage more interoperability and standardization between cloud providers.

Appendix A: Notable Cloud-Related Incidents

The following appendix contains brief accounts of eight major cloud-related incidents and their security implications.

Case 1: February 2017 Amazon Web Services East Coast Outage

On the morning of February 28, 2017, an engineer with Amazon Web Services (AWS) entered a command to remove some storage capacity from the billing system for AWS's Simple Storage Service. A typo inadvertently removed more storage capacity than intended.[159] The servers that were removed provided support to two critical Simple Storage Service internal systems: an indexing system and a system that allocated new storage when clients executed a request to store a new object. This error caused the whole Simple Storage Service system in this region, US-EAST-1—one of AWS's largest and most important regions—to become unavailable to process requests.[160]

In turn, the Simple Storage Service outage affected other AWS services, such as new instances in its computing service, Elastic Compute Cloud, Elastic Block Store volumes, and AWS Lambda. It took AWS engineers over four hours to restore Simple Storage Service to full service. However, during this interim, another Amazon service that went down was its AWS Service Health Dashboard, which is how service status updates are usually provided.[161]

Affected customers included a long list of companies and organizations, including Adobe, Airbnb, Coursera, Docker, Expedia, GitHub, Imgur, Medium, Pinterest, Quora, Signal, Slack, Trello, and the U.S. Securities and Exchange Commission (SEC).[162] It can thus be inferred that all of these companies and organizations had some of their systems relying on AWS services that were solely based in the US-EAST-1 region (located in northern Virginia). In contrast, the *Guardian* noted that it did not suffer an outage because it relies on AWS's Ireland-based region.[163] Even firms that maintain IoT devices (like home lighting systems) suffered outages, which meant that some customers could not turn on the lights in their homes because of the disruptions. Cyence, a firm that models cyber risk, issued an estimate that S&P 500 companies lost $150 million, a figure that doesn't include the second-order impacts from the outage.[164]

Simple Storage Service is a core AWS service on which many customers depend—one firm estimates that more than 250,000 (as of August 2020) unique domains rely on it.[165] AWS promises an availability of 99.99 percent for Simple Storage Service, which translates to fifty-three minutes of downtime every year (authors' calculations) and which this outage alone exceeded by about three hours.[166]

Case 2: September 2018 Microsoft Azure Outage

On September 4, 2018, early morning storms in San Antonio, Texas, led to lightning strikes that caused voltage sags and spikes on utility supplies for a Microsoft Azure data center in its South Central U.S. region. The data center switched its power supply over to a generator. The voltage spikes also caused the data center's cooling systems to shut down. Shortly thereafter, an automatic shutdown of the devices in the data center initiated in response to rising temperatures. Although this shutdown mechanism is designed to preserve the integrity of the Microsoft Azure servers and the data on them, some machines were damaged because temperatures rose so rapidly that they could not be shut down in time.[167]

As described in Azure's post-incident report, Microsoft engineers first sought to recover one critical component: the Azure software load balancers for its storage scale units. These components manage the allocation of new storage and internal networking, as well as the routing of customer data. The next part of the recovery involved replacing failed components in the data center. Microsoft engineers decided to try to recover all the customer data in the affected data center and to not "fail over," which entails resorting to backed-up data in another data center; they avoided this option because doing so would have resulted in some data being lost because data was replicated asynchronously, per the report.[168] Because replication of data across different geographic areas occurs at set time intervals, the data accumulated in the South Central data center after the previous replication would have been lost.

The failure at this data center affected Microsoft Azure services first in the South Central U.S. region—and almost all Azure services for customers who located data there were offline. However, the failure had a global impact because of its effects on a legacy management system for older types of Azure resources, called the Azure Service Manager. This system stored its metadata primarily in the South Central U.S. region, and the failure there caused it to experience delays until the South Central servers came back online in the early afternoon on September 5.

In response, Microsoft said it would review dependencies associated with the Azure Service Manager and move those services to a newer system, the Azure Resource Manager, which stores data in every Azure region for global resiliency. Another key affected system was Azure Active Directory, which is an authentication service critical to many other Azure services. Press reports indicated that, despite Microsoft's claims, there were problems with the company's cloud services globally.[169] Microsoft said in a blog post that one reason the incident had global impacts was because it affected "global" services like Azure Active Directory that in turn affected other systems.[170]

One of the key issues this incident highlighted was the distribution of data within regions. At the time of the incident, Microsoft Azure had only introduced availability zones (AZs) to three of its fifty-four regions and had not introduced any in the South Central U.S. region.[171] AZs are different data centers within the same region that provide resiliency for data in the region. Microsoft Azure suffered failures due to environmental factors in several other incidents, notably because of humidity in Ireland and, relatedly, the accidental release of fire suppression systems in a data center in North Europe in 2017.[172]

Case 3: November 2018 Google Cloud Outage

On the evening of November 12, 2018, Google Cloud traffic suffered connectivity problems, and its services were unavailable for many customers.[173] Network analysis pointed to an internet service provider in Nigeria that, at the time the incident began, updated its network addresses for the protocol that determines how internet traffic is routed, the Border Gateway Protocol (BGP), to say that its network was the best path to reach a set of Google-owned IP addresses. Then China Telecom, a major internet service provider in China that recently beforehand had been said to reroute U.S. internet traffic through China, improperly accepted the route, causing other network providers to list the route as correct.[174] The mistake was reportedly due to a configuration error with the Nigerian internet service provider's BGP filtering.[175] Google also said the "root cause of the issue was external to Google."[176]

This incident highlighted that major CSPs depend on broad availability of internet routing just like other major internet services, and these failures can happen as part of broader networking issues. However, because major CSPs have a concentration of services and infrastructure located in a specific set of IP addresses, failures that affect this set of addresses have an outsize impact on broader internet resilience. BGP in particular is a long-standing point of vulnerability in the internet's architecture, and one that other CSPs have suffered from. In April 2018, attackers hijacked BGP routing to redirect Domain Name System (DNS) addresses that were part of Amazon's Route 53 (its cloud DNS offering) so as to steal cryptocurrency.[177]

Case 4: January 2019 Microsoft Azure SQL Deletion

On January 29, 2019, according to media reports, the incident began with an apparently global issue for customers attempting to authenticate and access their Microsoft Azure and Office 365 accounts, a problem that seemed to strike Australia and New Zealand particularly hard.[178] According to journalists' accounts of Microsoft Azure's status report, "An external DNS service provider [reports said CenturyLink] experienced a global outage after rolling out a software update which exposed a data

corruption issue in their primary servers and affected their secondary servers, impacting network traffic."[179] This in turn affected Azure Active Directory, which lets users authenticate into Azure DNS and other Azure services. From this, SQL databases that used a particular configuration for Transparent Data Encryption (TDE), namely, customer Key Vaults, were affected. According to media accounts of statements provided to Azure customers, "An automated process, designed to trigger when custom keys are removed from KeyVault, inadvertently caused these TDE databases to be dropped."[180] Basically, certain SQL databases were deleted because of a process set in motion by the authentication failure.

In response, Microsoft Azure reportedly said, "We are in the process of restoring a copy of these SQL DBs from a recovery point in time of less than 5 minutes before the database was dropped. These restored databases . . . are located on the same server as the original database."[181] According to journalists' accounts, the company also asked customers, "for each database, [to] identify if lost transactions, during this 5 minute timeframe, could impact business processes or applications outside the database."[182] Asking customers to verify if lost data could have a significant effect on their businesses is a real nightmare for a CSP. According to media accounts of Azure's response, it switched to a different DNS provider. Its engineers also reportedly recovered all the relevant SQL databases that had been deleted.[183]

Case 5: March 2019 Facebook Server Tweak

On March 13, 2019, Facebook and several of its core apps—Instagram, Messenger, and WhatsApp—started experiencing issues around mid-morning for users globally.[184] Facebook services were ultimately down for swaths of users around the globe, with disruptions concentrated in the U.S. east coast and the United Kingdom, for fourteen hours, one of the longest outages in the company's history. The company did not issue a full statement, but it said on Twitter that the interruption was due to a "server configuration change."[185] It also clarified the outage was not because of a distributed denial-of-service (DDoS) attack. Facebook relies on its own private infrastructure instead of a CSP, and one interesting point is that WhatsApp services, which previously relied on IBM Cloud, were also disrupted, suggesting that Facebook had completed the migration of the app to its own infrastructure.[186]

A server configuration change points to what is likely a similar scenario as the 2017 AWS US-EAST-1 outage: an internal error that caused a cascading wave of issues leading to an outage as systems were debugged and restarted. The starting location of the issues, the U.S. east coast, also suggests an infrastructure-related issue, as this is where many major data centers are located. And network analysis suggested this was not a network issue.[187]

While the AWS outage may be a better example of this type of outage because more technical details about it are available, the Facebook one also provides a case study in impact. Since Facebook now integrates a number of services (the core app, WhatsApp, Instagram, and others) into one platform, the outage disrupted user experiences across the board—a phenomenon similar to what an infrastructure failure in a major CSP would do, or even just a failure for a major customer similar to Facebook.

Case 6: The Response to the Meltdown and Spectre Vulnerabilities

In January 2018, two major vulnerabilities in computing chips called Meltdown and Spectre became publicly known. This discovery had major implications for CSPs because their data centers were among the most vulnerable to such flaws. Several independent teams of researchers—including one within Google's vulnerability research team, Project Zero—had discovered the vulnerabilities around the same time or shortly after each other, in mid-2017.[188] Coordinated efforts to root out and patch the various ways these vulnerabilities have manifested are ongoing.

The first vulnerability, Meltdown, affects central processing units (CPUs)—the hardware chips that run computers—to allow an attacker to read privileged information that should only be visible to the CPU's operating system.[189] It does this by exploiting the way modern chips process instructions by executing future operations ahead of time. The vulnerability is that information about these steps is stored in the memory cache of the chip, accessible to an attacker. Attackers can then read the secrets of other programs and users on the same chip, allowing them to steal information like passwords and protected information. According to an official resource page, Meltdown "potentially" affects Intel chips dating all the way back to 1995.[190]

The second vulnerability, Spectre, affects nearly all chips manufactured in the last twenty years, not just those of Intel, but also chips made by ARM (a firm previously known as Advanced RISC Machine) and Advanced Micro Devices (AMD).[191] Like Meltdown, Spectre also exploits the speculative execution process to discover secrets within the same program, and it is more complex to execute.[192] However, it involves a fundamental flaw in security architecture, which means that attackers could read sensitive information on CSPs' software as a service (SaaS) offerings and jump from application to application. Another variant of Spectre allows attackers to read information across nearly every isolation barrier, even permitting them to read information across virtual machines and from hypervisors, another major vulnerability to CSP architecture.[193]

As a consequence, both of these vulnerabilities have had major implications for CSPs, as well as IT infrastructure writ large. Both involve hardware vulnerabilities, and while there have been software updates that could help address Meltdown, fixing the Spectre vulnerability has involved replacing

and rearchitecting computer hardware in many cases. These vulnerabilities have created serious concerns for CSPs' security, as these shortcomings could have allowed customers sharing the same physical hardware to steal each other's secrets.[194] But there are not yet any documented instances of exploitation of either Meltdown and Spectre in the wild, and CSPs have taken steps to respond and mitigate any potential threats.

How researchers, chip makers, and the major CSPs coordinated to address these vulnerabilities is a key example of some of the barriers to successful security cooperation to tackle major incidents in the cloud space. Project Zero researchers warned chip maker Intel as well as others about the vulnerabilities on June 1, 2017—about six months before they became public in January 2018.[195] Press reporting said Intel worked with other chip makers as well as a consortium of tech firms including Amazon, Google (which housed Project Zero), and Microsoft, but it did not provide a warning to smaller cloud services firms, who said they were blindsided by the January release. In addition, other press reports indicated that Intel notified not just major U.S. CSPs but also Chinese tech companies, including Alibaba and Lenovo.[196] This is notable because it had not notified key U.S. government agencies, including the Department of Homeland Security, by the time the vulnerabilities became public. Another report said Intel failed to notify even the U.S.-Computer Emergency Readiness Team (CERT), which issued a release advising the complete removal of affected processors and which it later had to revise to advise merely fixing the equipment.[197] The NSA denied that it exploited the vulnerabilities.[198]

A hearing convened by the U.S. Senate Committee on Commerce, Science, and Transportation touched on these issues in July 2018 when it reviewed the response to the disclosure of Meltdown and Spectre.[199] In his majority statement, Senator John Thune said congressional investigators confirmed that "some Chinese manufacturers, including Huawei, were informed of the vulnerability prior to public disclosure."[200] He also noted that, "only one company—IBM—reported that it contacted the U.S. government prior to the January 3, 2018, public disclosure." Witnesses did not provide further information, but one witness (Art Manion) from U.S.-CERT outlined a three-step vulnerability notification process that should ideally be implemented: the first to be notified would be the chip manufacturers, followed secondly by operating systems vendors, and then internet infrastructure providers and CSPs would be told third, along with cybersecurity defenders like U.S.-CERT.

Another major issue was that the fixes to the vulnerabilities may have had negative effects on performance. Some of the initial patches caused operating systems to experience instability, causing Microsoft to put a halt on its rollout of patches to some AMD chips.[201] Considering that the Spectre vulnerability was a flaw in the fundamental design and operating process, some patches for older chips came with what initial observers said was a notable reduction in performance.[202] However, later

reports disputed that the patches brought any significant change in the performance of affected chips.[203] Responding to Meltdown and Spectre has been a long-term endeavor. In November 2018, researchers released seven new variants of the exploits that they had discovered.[204] Intel has stood up a team of technical experts to manage this continuous response,[205] which still seems to be in operation.

Case 7: July 2019 Capital One Data Breach

On July 19, 2019, Capital One discovered that a hacker had breached its data on its customers and individuals who had applied for Capital One credit cards.[206] This involved the information of approximately 100 million people in the United States, although only a small fraction (less than 1 percent) of people had their Social Security numbers stolen. Capital One said it discovered the breach after an external researcher contacted it through its responsible disclosure program.[207] The stolen data was apparently stored on AWS public cloud servers used by Capital One.

Soon thereafter, federal prosecutors in Seattle announced the arrest of Paige Thompson, a former engineer at a Seattle technology company (presumably AWS), who was charged with hacking into Capital One.[208] According to the filed criminal complaint, Thompson posted on GitHub about the hack of Capital One as well as data she stole, which the complaint details as "through a misconfigured web application firewall" (prosecutors' words).[209] The researcher who notified Capital One saw this post and then notified the company, while Capital One went on to notify law enforcement. Per the criminal complaint, a law enforcement investigation in collaboration with Capital One found that the breach dated to March 2019 and also discovered that there may have been additional targets "of attempted or actual network intrusions."[210]

Numerous analysts have raised the point that, as a former AWS employee, Thompson likely used her knowledge of AWS services and security to gain access through Capital One's web application.[211] However, Thompson's intrusion was not especially technically complex and would not have required specialized inside knowledge to execute.

Capital One has made a highly public shift to the cloud, as the first financial institution to announce its intention to move all services to the cloud by 2020 and as one of the first large companies, period, to wholly commit to a cloud-based model. Its own chief executive officer said it was one of the "most cloud-forward" companies in the world earlier in 2019.[212] In the aftermath of the breach, many investors, insurers, and business experts questioned the wisdom of this decision and highlighted security issues with AWS.[213] And lawmakers also got involved, as Republicans on the House Oversight Committee sent letters asking for more information from both Capital One and Amazon, raising the possibility that the incident could cast doubt on the wisdom of awarding AWS the Penta-

gon's major cloud contract.[214] Democratic Senators Elizabeth Warren and Ron Wyden also sent letters to Capital One and AWS, respectively, asking for more information about the companies' responsibility for the breach.[215]

An AWS spokesperson attempted to push back against the narrative that its move to the cloud made it vulnerable, paraphrasing Capital One's statement to say, "this type of vulnerability is not specific to the cloud." AWS also said it "was not compromised in any way."[216]

AWS has maintained that the security of the system that permitted the breach (a Capital One web application) was wholly Capital One's responsibility to secure, as was the identity and access management for the data stored in its cloud. Capital One has not publicly disputed this as of July 2020. Security experts have said it is likely that the firewall misconfiguration was in fact a type of vulnerability called a server side request forgery (SSRF). An SSRF allows an attacker to communicate with a server and obtain metadata or credentials to then access internal databases. Journalist Brian Krebs discussed how this vulnerability was used to first manipulate the AWS metadata service, then to gain access to the AWS instance where Capital One's data was stored.[217]

This vulnerability has been well known, and although it is not cloud-specific, it has been cited as a particular threat for public cloud customers, and some analysts had previously called for AWS to implement a security header or make other modifications to defend against it.[218] In November 2019, AWS said it had updated its metadata access service to provide "defense in depth against unauthorized metadata access."[219]

Case 8: Operation Cloud Hopper

In December 2018, the U.S. Department of Justice unsealed an indictment charging two Chinese hackers with conducting a years-long campaign of cyber theft of intellectual property and confidential trade secrets from dozens of companies and governments around the world. The two hackers, Zhu Hua and Zhang Shilong, were charged with working on behalf of a provincial branch of China's Ministry of State Security.[220] Their hacking organization, known within the security community as Advanced Persistent Threat 10 (APT10), was able to compromise so many corporations because they targeted a set of IT firms that provided managed cloud services to the companies. This approach allowed the hackers to steal secrets from the clients through the compromised CSPs. This campaign ran from approximately 2014 to 2018 and was termed Operation Cloud Hopper.[221]

Further press reporting, as well as an April 2017 joint study by PwC and BAE Systems that was released prior to the indictment, provided more information about the extent of the intrusions as well as the technical details of the compromises. According to a June 2019 Reuters report, the

Chinese hackers compromised eight of the world's largest technology service providers, including Computer Sciences Corporation, Dimension Data, Fujitsu, Hewlett Packard Enterprise, IBM, NTT Data, Tata Consultancy Services, and DXC Technology—Hewlett Packard Enterprise's services arm that became a separate corporate entity in 2017.[222] These providers offer services to a vast array of clients around the globe, including Allianz SE; Deutsche Bank; GlaxoSmithKline; Philipps, the healthcare giant; and Rio Tinto, the mining firm.[223] The 2018 indictment referred to at least fifteen major corporations compromised in this campaign. It is conceivable that APT10 may have had the opportunity to infiltrate the systems of other corporations (including some of the aforementioned companies), though few concrete details are publicly known with certainty. Several clients were reported to have been compromised, including Ericsson, the Swedish telecom giant; Huntington Ingalls, the U.S. Navy's main shipbuilder; and Sabre, a U.S. travel reservation service.[224]

The compromised IT firms provided something called managed cloud service and are known as managed service providers (MSPs). These firms operate both the infrastructure for cloud computing and the overlaying applications and tools running on the infrastructure and then sell clients complete cloud-based IT service.[225] This allows large corporations especially to not have to build their own applications to run on cloud services. In Operation Cloud Hopper, APT10 gained access to the MSPs' systems through a common attack pattern, first by sending spear phishing emails that deliver malware and establishing itself in their systems and then moving to gain further access and infiltrate clients' networks. The PwC/BAE report noted that APT10 focused especially on shared infrastructure between MSPs and their clients; compromising these systems facilitated movement of targeted data from clients through the MSPs and then back to the hackers.[226] As later reporting revealed, because this infrastructure also often served multiple customers simultaneously, these tactics facilitated the wide-ranging nature of the intrusion campaign.[227]

This reporting also exposed the difficulties of responding to a dedicated, persistent campaign by a state-linked hacking group. According to the 2019 Reuters report, Hewlett Packard Enterprise's management was reticent about allowing its security team full access to combat the hackers and reportedly sought to limit the knowledge provided to compromised clients like Sabre.[228] One reason for this could have been attempting to avoid losses of confidence in the security of MSPs across the industry. However, even the federal authorities investigating the breaches noted this reticence to share information, suggesting that this problem is not confined to public disclosure of compromises.

Appendix B. Abbreviations, Figures, and Tables

Abbreviations

AMD	Advanced Micro Devices
ARM	a chip-making firm formerly known as Advanced RISC Machine
API(s)	application programming interface(s)
AWS	Amazon Web Services
AZ(s)	availability zone(s)
BGP	border gateway protocol
BPaaS	business process as a service
CERT	(U.S.) Computer Emergency Readiness Team
CIA	(U.S.) Central Intelligence Agency
CIA framework	the confidentiality, integrity, and availability framework
CISO	chief information security officer
CLOUD Act	Clarifying Lawful Overseas Use of Data Act
CPU(s)	central processing unit(s)
CSP(s)	cloud service provider(s)
DDOS	distributed denial of service (cyber attacks)
DNS	domain name system
EU	European Union
FedRAMP	Federal Risk and Authorization Management Plan
GAO	(U.S.) Government Accountability Office
GDPR	(EU) General Data Protection Regulation
IaaS	infrastructure as a service
IoT	Internet of Things
IT	information technology

JEDI	(U.S.) Joint Enterprise Defense Infrastructure
MSP(s)	managed service provider(s)
NASA	(U.S.) National Aeronautics and Space Administration
NIST	(U.S.) National Institute of Standards and Technology
OPEC	Organization of the Petroleum Exporting Countries
PaaS	platform as a service
SaaS	software as a service
SSRF	server side request forgery
TDE	Transparent Data Encryption
UAE	United Arab Emirates
UK	United Kingdom
UNCLOS	United Nations Convention for the Law of the Sea
VENOM	virtualized environment neglected operations manipulation (a cyber vulnerability)
WEF	World Economic Forum

Figures and Tables

Notes

1 Heidi M. Peters, "The Department of Defense's JEDI Cloud Program," Congressional Research Service, updated August 2, 2019, https://fas.org/sgp/crs/natsec/R45847.pdf.

2 "Gartner Forecasts Worldwide Public Cloud Revenue to Grow 17% in 2020," Gartner, press release, November 13, 2019, https://www.gartner.com/en/newsroom/press-releases/2019-11-13-gartner-forecasts-worldwide-public-cloud-revenue-to-grow-17-percent-in-2020.

3 Carmen Reinicke, "3 Reasons One Wall Street Firm Says to Stick With Cloud Stocks Amid the Coronavirus-Induced Market Rout," *Business Insider*, March 30, 2020, https://markets.businessinsider.com/news/stocks/wedbush-reasons-own-cloud-stocks-coronavirus-pandemic-tech-buy-2020-3-1029045273#2-the-move-to-cloud-will-accelerate-more-quickly-amid-the-coronavirus-pandemic2.

4 "Cloud Down: Impacts on the US Economy," Lloyd's and AIR Worldwide, 2018, https://www.lloyds.com/news-and-risk-insight/risk-reports/library/technology/cloud-down.

5 This paper focuses primarily on the public cloud and its policy implications. However, to help advance a more nuanced understanding of the cloud, this primer also details other cloud deployment models.

6 Rob Joyce, "Disrupting Nation State Hackers," presentation at USENIX Enigma 2016, https://www.usenix.org/sites/default/files/conference/protected-files/engima2016_transcript_joyce_v2.pdf.

7 Arun Sundararajan, "Network Effects," NYU Stern School of Business, http://oz.stern.nyu.edu/io/network.html.

8 Rich Miller, "Who Are the Data Center's Industry's Hyperscale Players?" Data Center Frontier, September 10, 2019, https://datacenterfrontier.com/data-centers-industry-hyperscale-players.

9 Lloyd's and AIR Worldwide, "Cloud Down: Impacts on the US Economy."

10 "The Economic Impact of Cybercrime: Not Slowing Down," McAfee, 2017, https://www.mcafee.com/enterprise/en-us/solutions/lp/economics-cybercrime.html.

11 "Ninth Annual Cost of Cybercrime Study," Accenture, March 6, 2019, https://www.accenture.com/us-en/insights/security/cost-cybercrime-study.

12 Dan Lohrmann, "Cloud First Policy: What Does It Really Mean?" Government Technology, blog post, December 19, 2010, https://www.govtech.com/blogs/lohrmann-on-cybersecurity/Cloud-First-Policy--121910.html.

13 Jonathan Zittrain, "The Internet's Fort Knox Problem," Harvard University *Future of the Internet*, blog post, June 3, 2010, http://blogs.harvard.edu/futureoftheinternet/2010/06/03/fort-knox-problem.

14 Zittrain argued in 2010 against such a centralized approach out of concern that it exacerbates other public policy risks such as illegitimate government access and control to privately held data. However, he did not deny the security benefits of the centralized approach, which (given the increased costs of the deteriorating cybersecurity landscape over the past decade) merit being revisited, arguably in favor of this more centralized approach.

15 The CISO of the large bank stated that the security in the cloud is "ten times" better than the security he could achieve with his team at the bank. Author interview by phone with Alex Stamos on October 23, 2018.

16 "Capital One on AWS," Amazon Web Services (AWS), accessed May 7, 2019, https://aws.amazon.com/solutions/case-studies/innovators/capital-one.

17 Additional useful primers on the cloud include: Patricia Moloney Figiola, "Cloud Computing: Background, Status of Adoption by Federal Agencies, and Congressional Action," U.S. Congressional Research Service, updated March 25, 2020, https://crsreports.congress.gov/product/pdf/R/R46119;

"IaaS, PaaS, and SaaS: IBM Cloud Service Models," IBM Cloud, https://www.ibm.com/cloud/learn/iaas-paas-saas; and Trey Herr, "Introduction to Cloud Computing," Cybersecurity Tech Accord YouTube channel, November 5, 2018, https://www.youtube.com/watch?v=pG-b_jbF4pc.

18 IBM Cloud, "IaaS, PaaS, and SaaS: IBM Cloud Service Models."

19 Peter Mell and Timothy Grance, "The NIST Definition of Cloud Computing," U.S. Department of Commerce National Institute of Standards and Technology (NIST), special publication 800-145, September 2011, https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf.

20 Revisions to this definition are not trivial. For example, major CSPs have an incentive to ensure that any updated definition favors the incumbents and becomes a potential barrier of entry to newcomers.

21 "Amazon EC2 FAQs," AWS, https://aws.amazon.com/ec2/faqs/#What_is_an_EC2_Compute_Unit_and_why_did_you_introduce_it.

22 "What Is the OSI Model?" CloudFlare, https://www.cloudflare.com/learning/ddos/glossary/open-systems-interconnection-model-osi.

23 James Sanders and Conner Forrest, "Hybrid Cloud: What It Is, Why It Matters," ZDNet, July 1, 2014, https://www.zdnet.com/article/hybrid-cloud-what-it-is-why-it-matters.

24 The NIST definition includes one more category: *community cloud.* This refers to cloud infrastructure shared between a number of customers with a shared interest. For instance, multiple businesses in the same industry could share the same cloud infrastructure configured to meet their unique regulatory and security requirements. In practice, this approach appears to have gained less hold than other deployment models for cloud computing, but reliable data on community adoption rates is difficult to find. See Bill Kleyman, "Explaining the Community Cloud," Data Center Knowledge, October 13, 2014, https://www.datacenterknowledge.com/archives/2014/10/13/explaining-community-cloud.

25 John Patrick Pullen, "Where Did Cloud Computing Come From, Anyway?" *Time*, March 19, 2015, http://time.com/collection-post/3750915/cloud-computing-origin-story.

26 Antonio Regalado, "Who Coined 'Cloud Computing'?" *MIT Technology Review,* October 31, 2011, https://www.technologyreview.com/s/425970/who-coined-cloud-computing; and Symantec Security Response, "A Brief History of Cloud Computing," Medium, January 18, 2018, https://medium.com/threat-intel/cloud-computing-e5e746b282f5.

27 Benjamin Black, "EC2 Origins," blog post, January 25, 2009, http://blog.b3k.us/2009/01/25/ec2-origins.html.

28 By 2019, AWS had provided the majority of Amazon's operating income for the prior four years. See "Amazon's Cloud Business Reports 35% Growth in the Third Quarter, Trailing Estimates," CNBC, October 24, 2019, https://www.cnbc.com/2019/10/24/aws-earnings-q3-2019.html.

29 Jack Clark, "How Amazon Exposed Its Guts: The History of AWS's EC2," ZDNet, June 7, 2012, https://www.zdnet.com/article/how-amazon-exposed-its-guts-the-history-of-awss-ec2.

30 Michael Arrington, "Google Jumps Head First Into Web Services With Its Google App Engine," *TechCrunch*, April 7, 2008, https://techcrunch.com/2008/04/07/google-jumps-head-first-into-web-services-with-google-app-engine.

31 Les Earnest, "Who Invented Timesharing?" Stanford University, March 26, 2016, https://web.stanford.edu/~learnest/nets/timesharing.htm.

32 "Why ARPAnet Wasn't the Beginning: A Brief History of the Internet—Part 1," Paessler, blog post, September 11, 2017, https://blog.paessler.com/history-of-the-internet-part-1.

33 U.S. Department of the Interior, "Cloud Smart Strategy," https://www.doi.gov/cloud/strategy#:~:text=The%20Cloud%20First%20Policy%20was,before%20making%20any%20new%20investments.

34 Douglas McIntyre, "Cloud Computing Reaches $100 Billion, But Who Makes Money?" 24/7 Wall St., October 14, 2015, https://247wallst.com/technology-3/2015/10/14/cloud-computing-reachs-100-billion-but-who-makes-money.

35 Yury Izrailevsky, "Completing the Netflix Cloud Migration," Netflix, blog post, February 11, 2016, https://media.netflix.com/en/company-blog/completing-the-netflix-cloud-migration; "Amazon EC2 Crosses the Atlantic," AWS, December 10, 2008, https://aws.amazon.com/about-aws/whats-new/2008/12/10/amazon-ec2-crosses-the-atlantic; and "AWS Launches the Northern California Region," AWS, December 3, 2009, https://aws.amazon.com/about-aws/whats-new/2009/12/03/aws-launches-the-northern-california-region.

36 K. Sreenand, Rijo Rajan, and P. Indumathi, "IBM Cloud Computing," *International Research Journal of Advanced Engineering and Science* 1, no. 4 (2016): 89–90, http://www.irjaes.com/pdf/IRJAES-V1N4Y16/IRJAES-V1N4P181Y16.pdf.

37 Sholto Macpherson, "Microsoft Announces Azure Launch Date," IT News, February 8, 2010, https://www.itnews.com.au/news/microsoft-announces-azure-launch-date-166664.

38 James Sanders, "Micorsoft Azure: A Cheat Sheet," *TechCrunch*, June 18, 2018, https://www.techrepublic.com/article/microsoft-azure-the-smart-persons-guide.

39 U.S. Department of Commerce NIST, "Final Version of NIST Cloud Computing Definition Published," October 25, 2011, https://www.nist.gov/news-events/news/2011/10/final-version-nist-cloud-computing-definition-published.

40 Vivek Kundra, "Federal Cloud Computing Strategy," White House, February 8, 2011, https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/federal-cloud-computing-strategy.pdf.

41 "Amazon Web Services Now Available to Customers From Data Centers in the UK," Amazon, press release, December 13, 2016, https://press.aboutamazon.com/news-releases/news-release-details/amazon-web-services-now-available-customers-data-centers-uk-0; "Amazon Web Services Launches New Region in France," Amazon, press release, December 18, 2017, https://press.aboutamazon.com/news-releases/news-release-details/amazon-web-services-launches-new-region-france; Sebastian Moss, "Microsoft Opens Three UK Data Centers, Azure Cloud Now Local," Data Center Dynamics, September 7, 2016, https://www.datacenterdynamics.com/news/microsoft-opens-three-uk-data-centers-azure-cloud-now-local; and Tom Keane, "Microsoft Azure Preview With Azure Availability Zones Now Open in France," Microsoft Azure, blog post, December 12, 2017, https://azure.microsoft.com/en-us/blog/microsoft-azure-preview-with-azure-availability-zones-now-open-in-france.

42 "Amazon Web Services Launches Brazil Datacenters for Its Cloud Computing Platform," Amazon, press release, December 14, 2011, https://press.aboutamazon.com/news-releases/news-release-details/amazon-web-services-launches-brazil-datacenters-its-cloud; and "Azure Brazil South Region Now in Public Preview," Microsoft Azure, April 17, 2014, https://azure.microsoft.com/en-us/updates/azure-brazil-south-region-now-in-public-preview.

43 "Geography and Regions," Google Cloud, https://cloud.google.com/docs/geography-and-regions; "Regions, Availability Zones, and Local Zones," AWS, https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/Concepts.RegionsAndAvailabilityZones.html; and "Azure Regions," Microsoft Azure, https://azure.microsoft.com/en-us/global-infrastructure/regions.

44 "Alibaba Cloud's Global Infrastructure," Alibaba Cloud, https://www.alibabacloud.com/global-locations.

45 "IBM Commits $1.2 Billion to Expand Global Cloud Footprint," IBM, press release, January 17, 2014, https://www-03.ibm.com/press/us/en/pressrelease/42956.wss.

46 In 2011, the market research firm Gartner estimated that the global IaaS market was $3.7 billion in size. See Rachel Wheeler, "IaaS Market to Record 'Strong Growth,'" Experian, blog post, April 7, 2011, https://www.edq.com/blog/iaas-market-to-record-strong-growth; Louis Columbus, "Roundup of Cloud Computing Forecasts and Market Estimates," *A Passion for Research*, blog post, January 17, 2012, https://softwarestrategiesblog.com/2012/01/17/roundup-of-cloud-computing-forecasts-and-market-estimates-2012; and "Gartner Forecasts Worldwide Public Cloud Revenue to Grow 17.3 Percent in 2019," Gartner, press release, September 12, 2018, https://www.gartner.com/en/newsroom/press-releases/2018-09-12-gartner-forecasts-worldwide-public-cloud-revenue-to-grow-17-percent-in-2019.

47 Paddy Srinivasan, "The Real Market Size of Public Cloud Services," *Wired*, March 2013, https://www.wired.com/insights/2013/03/the-real-market-size-of-public-cloud-services.

48 "Introduction: A Bit of OpenStack History," OpenStack, last updated May 25, 2018, https://docs.openstack.org/project-team-guide/introduction.html.

49 Jack Clark, "Report: Amazon Dominates Global Cloud Spend," *Register*, March 13, 2013, https://www.theregister.co.uk/2013/03/13/amazon_has_a_third_of_global_iaas_spend.

50 Ibid.

51 Ibid.

52 Julie Bort, "Amazon Is Crushing IBM, Microsoft, and Google in Cloud Computing, Says Report," *Business Insider*, November 26, 2013, https://www.businessinsider.com/amazon-cloud-beats-ibm-microsoft-google-2013-11.

53 Kim Weins, "Cloud Computing Trends: 2015 State of the Cloud Survey," Flexera, *IT Industry Trends*, blog post, February 18, 2015, https://www.rightscale.com/blog/cloud-industry-insights/cloud-computing-trends-2015-state-cloud-survey.

54 Steven Vaughan-Nichols, "What Is Docker and Why Is It So Darn Popular?" ZDNet, March 21, 2018, https://www.zdnet.com/article/what-is-docker-and-why-is-it-so-darn-popular.

55 David Byrne, Carol Corrado, and Daniel E. Sichel, "The Rise of Cloud Computing: Minding Your P's, Q's, and K's," National Bureau of Economic Research, Working Paper 25188, October 2018, https://www.nber.org/papers/w25188.pdf.

56 "In the Cloud Price War, Cloud Storage Has Become the New Battleground," 451 Research, press release, April 20, 2017, https://451research.com/images/Marketing/press_releases/CPI_PR_04_20_2017_vf.pdf?&utm_campaign=2017_press&utm_source=twitter&utm_medium=social&utm_content=press_release&utm_term=q2_2017_cpi_pr.

57 Brandon Butler, "Rackspace Bows Out of Commodity IaaS Market in Favor of 'Managed Cloud,'" Network World, August 5, 2014, https://www.networkworld.com/article/2461361/iaas/rackspace-bows-out-of-iaas-market.html.

58 Brandon Butler, "HP Just Dropped Out of the Public Cloud—Now What?" Network World, October 22, 2015, https://www.networkworld.com/article/2996536/cloud-computing/hp-just-dropped-out-of-the-public-cloud-now-what.html; and Quentin Hardy, "HP Comes to Terms With the Cloud," *New York Times*, *Bits* (blog post), April 7, 2015, https://bits.blogs.nytimes.com/2015/04/07/hp-comes-to-terms-with-the-cloud/.

59 Steven Vaugh-Nichols, "Verizon to Shut Down Its Public Cloud," ZDNet, February 17, 2016, https://www.zdnet.com/article/verizon-closes-down-public-cloud; Shoshanna Delventhal, "Cisco to Terminate $1B Public Cloud Unit," Investopedia, December 14, 2016, https://www.investopedia.com/news/cisco-terminate-1b-public-cloud-unit; and Dan Meyer, "Verizon and AT&T Exit From Cloud Business Applauded by Analysts," SDX Central, May 5, 2017, https://www.sdxcentral.com/articles/news/verizon-and-att-exit-from-cloud-business-applauded-by-analysts/2017/05.

60 "Gartner Says a Massive Shift to Hybrid Infrastructure Is Underway," Gartner, press release, April 5, 2017, https://www.gartner.com/en/newsroom/press-releases/2017-04-05-gartner-says-a-massive-shift-to-hybrid-infrastructure-services-is-underway; Marty Chilberg, "Gartner: The IaaS Cloud Will Be a Duopoly by 2019," *Marty Chilberg's Blog*, blog post, August 19, 2017, https://seek ingalpha.com/instablog/400846-marty-chilberg/5029588-gartner-iaas-cloud-will-duopoly-2019.

61 Seth Fiegerman, "Microsoft's Cloud Bet Pushes Annual Sales Over $100 billion," CNN, July 19, 2018, https://money.cnn.com/2018/07/19/technology/microsoft-earnings/index.html.

62 Cisco, "Cisco Global Cloud Index: Forecast and Methodology: 2016–2021," 2018, https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf.

63 While Cisco doesn't name all twenty-four companies with hyperscale data operations centers, a partial list includes ADP, Alibaba, Amazon, Apple, Baidu, eBay, Facebook, Google, Microsoft, Oracle, Rackspace, Salesforce, Tencent, and Yahoo. See Yevgeniy Sverdlik, "Research: There Are Now Close to 400 Hyper-Scale Data Centers in the World," Data Center Knowledge, December 22, 2017, https://www.datacenterknowledge.com/cloud/research-there-are-now-close-400-hyper-scale-data-centers-world.

64 "Hyperscale Cloud Market Overview," Synergy Research Group, April 2017, https://srgresearch.s3.amazonaws.com/public-reports/hyperscale-market-overview-primer-april-2017.pdf.

65 "The Leading Cloud Providers Increase Their Market Share Again in the Third Quarter," Synergy Research Group, October 25, 2018, https://www.srgresearch.com/articles/leading-cloud-providers-increase-their-market-share-again-third-quarter.

66 For more on hyperscale, see Trey Herr, "Hyperscaling: The Structure of the Internet in the Second Age of Cloud," NATO Cooperative Cyber Defense Center of Excellence (CCDCE) 2018 International Conference on Cyber Conflict, NATO CCDCE YouTube channel, https://www.youtube.com/watch?v=pXQ4sl-3J-g&feature=youtu.be&t=1244.

67 Ron Miller, "Alibaba Cloud Growing Like Gangbusters, But Still Far Behind AWS and Other Market Leaders," *TechCrunch*, February 6, 2018, https://techcrunch.com/2018/02/06/alibaba-cloud-growing-like-gangbusters-but-still-far-behind-aws-and-other-market-leaders.

68 Hari Kannan and Christopher Thomas, "Public Cloud in China: Big Challenges, Big Upside," McKinsey, July 2018, https://www.mckinsey.com/industries/high-tech/our-insights/public-cloud-in-china-big-challenges-big-upside.

69 "China Public Cloud (IaaS) to Reach US$6.21Bn in 2018; Amazon Fastest Growth," China Internet Watch, October 10, 2018, https://www.chinainternetwatch.com/26900/public-cloud-iaas-2018.

70 Yevgeniy Sverdlik, "Microsoft Launches Mainland-China Azure Cloud Data Center," Data Center Dynamics, March 27, 2014, https://www.datacenterdynamics.com/news/microsoft-launches-mainland-china-azure-cloud-data-center; Doug Haugher, "Windows Azure, Operated by 21Vianet, Now Generally Available in China," Microsoft Azure, blog post, March 26, 2014, https://azure.microsoft.com/en-us/blog/windows-azure-services-now-generally-available-in-china/; and "Overview of Cloud Computing in China," Export.gov, March 16, 2016, https://www.export.gov/article?id=Overview-of-Cloud-Computing-in-China.

71 Mo Chen, Andrew McGinty, Mark Parsons, Liang Xu, and Roy Zou, "Evolving Landscape for International Cloud Providers in China: Why US Technology Giants Are Pairing Up With Local Partners," JDSupra, July 25, 2018, https://www.jdsupra.com/legalnews/evolving-landscape-for-international-16682.

72 Jon Russell, "AWS Isn't Exiting China, But Amazon Did Sell Physical Assets to Comply With Chinese Law," *TechCrunch*, November 13, 2017, https://techcrunch.com/2017/11/13/aws-exits-china.

73 Gartner, "Gartner Forecasts Worldwide Public Cloud Revenue to Grow 17.3 Percent in 2019."
74 See figure 2 on page 6 in the following report: World Economic Forum, "The Global Risk Report 2019," 2019, http://www3.weforum.org/docs/WEF_Global_Risks_Report_2019.pdf.
75 World Economic Forum, "The Global Risk Report 2019."
76 EY, "EY Global Information Security Survey," 2019, 5, https://assets.ey.com/content/dam/ey-sites/ey-com/en_ca/topics/advisory/ey-global-information-security-survey-2018-19.pdf.
77 Justina Crabtree, "Lloyd's of London CEO Says 92 Percent of European Businesses Have Experienced Cyber Breaches," CNBC, December 19, 2016, https://www.cnbc.com/2016/12/19/lloyds-of-london-ceo-says-92-percent-of-european-businesses-have-experienced-cyber-breaches.html.
78 Accenture, "Ninth Annual Cost of Cybercrime Study."
79 Ibid.
80 EY, "EY Global Information Security Survey," 2019, 16.
81 Author interview by phone with Alex Stamos on October 23, 2018.
82 Based on remarks made by the CISO of a major financial institution at a private roundtable, January 2019.
83 Liam Tung, "AWS Users Fret Over Downtime Ahead of Amazon's Massive EC2 Reboot," ZDNet, September 25, 2014, https://www.zdnet.com/article/aws-users-fret-over-downtime-ahead-of-amazons-massive-ec2-reboot.
84 Jason Zander, "Final Root Cause Analysis and Improvement Areas: Nov 18 Azure Storage Service Interruption," Microsoft Azure, blog post, December 17, 2014, https://azure.microsoft.com/en-us/blog/final-root-cause-analysis-and-improvement-areas-nov-18-azure-storage-service-interruption.
85 Ben Sullivan, "Lighting Strikes Cause of Google Cloud Outage," Silicon.co.uk, August 19, 2015, https://www.silicon.co.uk/cloud/google-cloud-outage-lightning-strikes-175135; and Google Cloud Platform, "Google Compute Engine Incident #15056," Google Cloud Status Dashboard, August 18, 2015, https://status.cloud.google.com/incident/compute/15056.
86 Charles Babcock, "Amazon Disruption Produces Cloud Outage Spiral," Information Week, September 22, 2015, https://www.informationweek.com/cloud/infrastructure-as-a-service/amazon-disruption-produces-cloud-outage-spiral/d/d-id/1322279.
87 Gavin Clarke, "Microsoft Struggles Against Self-Inflicted Office 365 IMAP Outage," *Register*, January 25, 2016, https://www.theregister.com/2016/01/25/office_365_imap_outage/.
88 Dom Nicastro, "Salesforce Customers Lose CRM Data in 20 Hour Outage," CMS Wire, May 13, 2016, https://www.cmswire.com/customer-experience/salesforce-customers-lose-crm-data-in-20-hour-outage; and Salesforce, "RCM for NA14 Disruptions of Service – May 2016," May 16, 2016, https://help.salesforce.com/articleView?language=en_US&type=1&mode=1&id=000315819.
89 "Summary of the Amazon S3 Service Disruption in the Northern Virginia (US-EAST-1) Region," AWS, https://aws.amazon.com/message/41926; Jordan Novet, "AWS Is Investigating S3 Issues, Affecting Quora, Slack, Trello (Updated)," VentureBeat, February 28, 2017, https://venturebeat.com/2017/02/28/aws-is-investigating-s3-issues-affecting-quora-slack-trello; and Yevgeniy Sverdlik, "AWS Outage That Broke the Internet Caused by Mistyped Command," Data Center Knowledge, March 2, 2017, https://www.datacenterknowledge.com/archives/2017/03/02/aws-outage-that-broke-the-internet-caused-by-mistyped-command.
90 Simon Sharwood, "Azure Storage Browns Out for Eight Hours, Nobody Notices," *Register*, March 17, 2017, https://www.theregister.com/2017/03/17/azure_storage_brownout; and Jordan Novet, "Microsoft Confirms Azure Storage Issues Around the World (Updated)," VentureBeat, March 15, 2017, https://venturebeat.com/2017/03/15/microsoft-confirms-azure-storage-issues-around-the-world.

91 Ted Greenwald and Jack Nicas, "Intel Wrestled With Chip Flaws for Months," *Wall Street Journal,* January 5, 2018, https://www.wsj.com/articles/intel-wrestled-with-chip-flaws-for-months-1515110151?mod=article_inline.

92 Jann Horn, "Reading Privileged Memory With a Side-Channel," Google, *Project Zero* blog post, January 3, 2018, https://googleprojectzero.blogspot.com/2018/01/reading-privileged-memory-with-side.html.

93 Thomas Claburn, "Azure North Europe Downed by the Curse of the Irish – Sunshine," *Register*, June 22, 2018, https://www.theregister.com/2018/06/22/azure_north_europe_downed_by_pleasant_weather; and Sead Fadilpašić, "Microsoft Azure Suffers Major Outage," ITProPortal, June 20, 2018, https://www.itproportal.com/news/microsoft-azure-suffers-major-outage.

94 "Postmortem: VSTS 4 September 2018," Microsoft Azure, *DevOps* blog post, September 10, 2018, https://devblogs.microsoft.com/devopsservice/?p=17485; and Richard Speed, "Microsoft Azure: It's Getting Hot in Here, So Shut Down All Your Cores," *Register*, September 4, 2018, https://www.theregister.co.uk/2018/09/04/azure_its_getting_hot_in_here.

95 Dan Goodin, "Google Goes Down After Major BGP Mishap Routes Traffic Through China," Ars Technica, November 13, 2018, https://arstechnica.com/information-technology/2018/11/major-bgp-mishap-takes-down-google-as-traffic-improperly-travels-to-china.

96 "Azure Active Directory Outage and RCA for Azure Cloud Hiccups," Born's Tech and Windows World, February 2, 2019, https://borncity.com/win/2019/02/02/azure-active-directory-outage-rca-for-azure-cloud-hicups; and Danny Bradbury, "Microsoft Azure Data Deleted Because of DNS Outage," Naked Security, February 1, 2019, https://nakedsecurity.sophos.com/2019/02/01/dns-outage-turns-tables-on-azure-database-users.

97 Alex Hern, "How Did an Amazon Glitch Leave People Literally in the Dark?" *Guardian,* March 1, 2017, https://www.theguardian.com/technology/2017/mar/01/amazon-web-services-outage-smart-homes.

98 This approach has obvious limitations as the past does not necessarily predict the future.

99 Juan Carlos Crisanto, Conor Donaldson, Denise Garcia Ocampo, and Jermy Prenio, "Regulating and Supervising the Clouds: Emerging Prudential Approaches for Insurance Companies," Financial Stability Institute (FSI) Insights no. 13, December 2018, https://www.bis.org/fsi/publ/insights13.pdf.

100 The authors thank Trey Herr for his feedback and for drawing their attention specifically to hypervisor vulnerabilities.

101 The authors thank Trey Herr for drawing attention to the particular role that hypervisors play with respect to the overall architecture and security of the cloud.

102 Joe Warminsky, "Apple, Oracle, VMware Products Successfully Hacked at Pwn2Own," Cyberscoop, March 21, 2019, https://www.cyberscoop.com/pwn2own-2019-day-one-apple-oracle-vmware.

103 Daniel Gray, "Hyperjacking—Future Computer Server Threat," SysChat, September 2, 2009, http://www.syschat.com/hyperjacking-future-computer-server-threat-4917.html.

104 "Virtualized Environment Neglected Operations Manipulation (VENOM)," Crowdstrike, updated May 21, 2015, https://venom.crowdstrike.com.

105 John Patrick Barrowclough and Rameez Asif, "Securing Cloud Hypervisors: A Survey of Threats, Vulnerabilities, and Countermeasures," *Security and Communication Networks* 18 (June 11, 2018), https://www.hindawi.com/journals/scn/2018/1681908/.

106 In 2018, NIST published security guidance for "server-based hypervisor platforms," which basically means CSPs, although this could include other data center management setups. The document discusses some of the major security ramifications of hypervisors. It outlines three classes of threats: breach of process isolation, breach of network isolation, and denial of service. See Ramaswamy Chandramouli,

"Security Recommendations for Server-Based Hypervisor Platforms," U.S. Department of Commerce NIST, Special Publication 800-125A revision 1, June 2018, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-125Ar1.pdf.

107 "Cyber Incident Serverity Schema," White House, July 26, 2016, https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Cyber%2BIncident%2BSeverity%2BSchema.pdf.

108 In 2018, the U.S. Census Bureau failed to secure critical data in AWS storage buckets. See Tajha Chappellet-Lanier, "Cloud Security Weaknesses Put 2020 Census Prep at 'Potentially Catastrophic Risk,'" FedScoop, June 24, 2019, https://www.fedscoop.com/census-bureau-cloud-security-audit.

109 Paul Schwartz, "Legal Access to the Global Cloud," *Columbia Law Review* 118, no. 6 (2018): 1681–1762, https://columbialawreview.org/content/legal-access-to-the-global-cloud; and Jennifer Daskal, "The Un-Territoriality of Data," *Yale Law Review* 125 (2015): 326–398, https://www.yalelawjournal.org/pdf/a.326.Daskal.398_qrhgeoar.pdf.

110 Andrew Keane Woods and Peter Swire, "The CLOUD Act: A Welcome Legislative Fix for Cross-Border Data Problems," *Lawfare*, blog post, February 6, 2018, https://www.lawfareblog.com/cloud-act-welcome-legislative-fix-cross-border-data-problems.

111 Shaul Brazil and Jonthan Flynn, "Overseas Production Orders—A New Tool for UK Law Enforcement," BCL Solicitors, February 2019, https://www.bcl.com/overseas-production-orders-a-new-tool-for-uk-law-enforcement.

112 Kati Suominen, "No Choice? GDPR's Impact on the US, UK, and the EU," Center for Strategic and International Studies (CSIS), blog post, January 31, 2018, https://www.csis.org/blogs/future-digital-trade-policy-and-role-us-and-uk/no-choice-gdprs-impact-us-uk-and-eu.

113 "Freedom on the Net 2019: The Crisis of Social Media," Freedom House, 2019, https://freedomhouse.org/report/freedom-net.

114 Freedom House, "Bahrain," in *Freedom in the World 2020: A Leaderless Struggle for Democracy*, 2020, https://freedomhouse.org/country/bahrain/freedom-world/2020; and Freedom House, "United Arab Emirates," in *Freedom in the World 2020: A Leaderless Struggle for Democracy*, 2020, https://freedomhouse.org/country/united-arab-emirates/freedom-world/2020.

115 Werner Vogels, "As-Salaam-Alaikum: The Cloud Arrives in the Middle East!" *All Things Distributed*, blog post, September 25, 2017, https://www.allthingsdistributed.com/2017/09/aws-region-middle-east.html.

116 Jason Zander, "Microsoft Expands Cloud Services in Europe and Into Middle East to Meet Growing Customer Demand," Microsoft, blog post, March 14, 2018, https://blogs.microsoft.com/blog/2018/03/14/microsoft-expands-cloud-services-in-europe-and-into-middle-east-to-meet-growing-customer-demand.

117 Sebastian Moss, "Google Cloud Continues to Grow, Is Coming to Saudi Arabia," DataCenter Dynamics, April 24, 2018, https://www.datacenterdynamics.com/news/google-cloud-continues-to-grow-is-coming-to-saudi-arabia.

118 Sebastian Moss, "Google Still Pursuing Saudi Arabian Data Centers, After Khashoggi's Murder," Data Center Dynamics, April 18, 2019, https://www.datacenterdynamics.com/en/news/google-still-pursuing-saudi-arabian-data-centers-after-khashoggis-murder.

119 Aditya Kalra, "Exclusive: U.S. Senators Urge India to Soften Data Localization Stance," Reuters, October 13, 2018, https://www.reuters.com/article/us-india-data-localisation-exclusive/exclusive-u-s-senators-urge-india-to-soften-data-localization-stance-idUSKCN1MN0CN.

120 Yuxi Wei, "Chinese Data Localization Law: Comprehensive But Ambiguous," University of Washington Henry M. Jackson School of International Relations, February 7, 2018, https://jsis.washington.edu/news/chinese-data-localization-law-comprehensive-ambiguous.

121 Wei, "Chinese Data Localization Law: Comprehensive but Ambiguous."
122 Eset Dedezade, "Microsoft to Deliver Cloud Services From New Datacentres in Germany in 2019 to Meet Evolving Customer Needs," Microsoft News Center Europe, 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.
123 Aidan Finn, "Changes to Azure Germany Operations," Petri, October 5, 2018, https://www.petri.com/changes-to-azure-germany-operations.
124 Josh Kallmer, "China: Challenge to U.S. Commerce," testimony, U.S. Senate Committee on Commerce, Science, and Transportation, Subcommittee on Security, March 7, 2019, https://www.itic.org/dotAsset/7267b522-28c6-4bb7-9719-c9f1f20fd9a3.pdf.
125 Yifan Yu, "Amazon Prepares to Battle With Alibaba in Asia's Cloud," *Nikkei Asian Review,* May 4, 2019, https://asia.nikkei.com/Business/Companies/Amazon-prepares-to-battle-with-Alibaba-in-Asia-s-cloud; and Mercedes Ruehl, "US and Chinese Cloud Companies Vie for Dominance in South-East Asia," *Financial Times,* May 19, 2020, https://www.ft.com/content/1e2b9cd9-f82e-4d3b-a2d8-f20c08bdc3aa.
126 U.S. Office of the Director of National Intelligence, "Threats to Undersea Cable Communications," September 28, 2017, https://www.dni.gov/files/PE/Documents/1---2017-AEP-Threats-to-Undersea-Cable-Communications.pdf; https://www.reuters.com/article/us-huawei-tech-usa-cable/chinas-huawei-to-sell-undersea-cable-business-buyers-exchange-filing-shows-idUSKCN1T40BS.
127 Agence France-Presse, "Solomon Islands Drops Chinese Tech Giant Huawei for Billion-Dollar Undersea Cable, Signs Australia," *South China Morning Post*, June 13, 2018, https://www.scmp.com/news/asia/diplomacy/article/2150616/solomon-islands-drops-chinese-tech-giant-huawei-billion-dollar.
128 Tom Westbrook, "PNG Upholds Deal With Huawei to Lay Internet Cable, Derides Counter-Offer," Reuters, November 26, 2018, https://www.reuters.com/article/us-papua-huawei-tech/png-upholds-deal-with-huawei-to-lay-internet-cable-derides-counter-offer-idUSKCN1NV0DR.
129 Jeremy Page, "Huawei Selling Stake in Undersea-Cable Firm as U.S. Pressure Mounts," *Wall Street Journal,* June 3, 2019, https://www.wsj.com/articles/huawei-selling-stake-in-undersea-cable-firm-as-u-s-pressure-mounts-11559562892; and Jeremy Page, Kate O'Keeffe, and Rob Taylor, "America's Undersea Battle with China for Control of the Global Internet Grid," *Wall Street Journal,* March 12, 2019, https://www.wsj.com/articles/u-s-takes-on-chinas-huawei-in-undersea-battle-over-the-global-internet-grid-11552407466.
130 Tony Romm, "Google Fined Nearly $1.7 Billion for Ad Practices That E.U. Says Violated Antitrust Laws," *Washington Post,* March 20, 2019, https://www.washingtonpost.com/technology/2019/03/20/google-fined-nearly-billion-ad-practices-that-violated-european-antitrust-laws/?utm_term=.57f8971cd3f1.
131 Elizabeth Warren, "Here's How We Can Break Up Big Tech," Medium, March 8, 2019, https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c.
132 Lina M. Khan, "Amazon's Antitrust Paradox," *Yale Law Journal* 126 (2017): 754, https://digitalcommons.law.yale.edu/cgi/viewcontent.cgi?article=5785&context=ylj.
133 Tae Kim, "Amazon Should Split Into Two Companies to Avoid Antitrust Scrutiny From Trump Administration: Citi," CNBC, September 17, 2018, https://www.cnbc.com/2018/09/17/amazon-should-split-its-retail-and-cloud-computing-businesses-citi.html.
134 Jon Bateman, "The Antitrust Threat to National Security," *Wall Street Journal,* October 22, 2019, https://www.wsj.com/articles/the-antitrust-threat-to-national-security-11571784197.
135 Crisanto, Donaldson, Ocampo, and Prenio, "Regulating and Supervising the Clouds: Emerging Prudential Approaches for Insurance Companies."

136 Trey Herr's argues that "The U.S. strategy to secure cloud computing is incomplete and, unless there is a shift in regulatory thinking, the push for cloud computing as a solution to security shortfalls in small and medium-sized organizations will only produce more risk." See Trey Herr, "Better to Be Realistic About the Security Opportunities of Cloud Computing," *Lawfare*, blog post, March 17, 2020, https://www.lawfareblog.com/better-be-realistic-about-security-opportunities-cloud-computing.

137 Paresh Dave, "Google's New Cloud Boss Has Big Task to Catch Rivals, Reuters Data Show," Reuters, February 21, 2019, https://www.reuters.com/article/us-alphabet-google-cloud-focus/googles-new-cloud-boss-has-big-task-to-catch-rivals-reuters-data-show-idUSKCN1QA0HB.

138 Lixian Loong Hantover, "The Cloud and The Deep Sea: How Cloud Storage Raises the Stakes for Undersea Cable Security and Liability," *Ocean and Coastal Law Journal* 19, no. 1 (2014): 1–28, https://digitalcommons.mainelaw.maine.edu/cgi/viewcontent.cgi?article=1028&context=oclj.

139 The working group of the Federal Communications Commission in the United States wrote a report in 2014 on protecting cables through spatial separation. See "Final Report—Protection of Submarine Cables Through Spatial Separation," Federal Communications Commission, Communications, Security, Reliability and Interoperability Council, Working Group 8 Submarine Cable Routing and Landing, December 2014, https://transition.fcc.gov/pshs/advisory/csric4/CSRIC_IV_WG8_Report1_3Dec2014.pdf. In 2017, the U.S. Office of the Director of National Intelligence published a report on threats to undersea cables providing an in-depth review of the submarine cable industry. See "Threats to Undersea Cable Communications," U.S. Office of the Director of National Intelligence, September 28, 2017, https://www.dni.gov/files/PE/Documents/1---2017-AEP-Threats-to-Undersea-Cable-Communications.pdf. A British parliamentarian also wrote a major report on the vulnerabilities of cables in 2017. See Rishi Sunak, "Undersea Cables: Indispensible, Insecure," Policy Exchange, November 2017, https://policyexchange.org.uk/wp-content/uploads/2017/11/Undersea-Cables.pdf.

140 John K. Crain, "Assessing Resilience in the Global Undersea Cable Infrastructure," Naval Postgraduate School, thesis, June 2012, https://apps.dtic.mil/dtic/tr/fulltext/u2/a562772.pdf.

141 Hantover, "The Cloud and the Deep Sea."

142 "Government Cloud First Policy," UK Government, February 3, 2017, https://www.gov.uk/guidance/government-cloud-first-policy.

143 Aaron Boyd, "White House Outlines Move From 'Cloud First' to 'Cloud Smart,'" NextGov, September 12, 2018, https://www.nextgov.com/it-modernization/2018/09/white-house-outlines-move-cloud-first-cloud-smart/151498/; and U.S. Chief Information Officer (CIO), "Federal Cloud Computing Strategy," accessed August 16, 2020, https://cloud.cio.gov/.

144 U.S. Government Accountability Office, "Agencies Have Increased Usage and Realized Benefits, but Cost and Savings Data Need to Be Better Tracked," GAO-19-58, 2019, https://www.gao.gov/products/GAO-19-58.

145 U.S. CIO, "From Cloud First to Cloud Smart," https://cloud.cio.gov/strategy/#security; and "FedRAMP," FedRAMP, https://www.fedramp.gov/.

146 "Buying and Selling on the Digital Marketplace," UK Government, June 27, 2019, https://www.gov.uk/guidance/the-g-cloud-framework-on-the-digital-marketplace.

147 David Braue, "Australian Government Signs Up with AWS," Information Age, July 2, 2019, https://ia.acs.org.au/article/2019/australian-government-signs-up-with-aws.html; Aaron Tan, "Singapore Government Makes Bold Cloud Move," *Computer Weekly*, October 2, 2018, https://www.computerweekly.com/news/252449829/Singapore-government-makes-bold-cloud-move; and Chris Thornett, "German Government Goes Open Source With Cloud Firm Nextcloud," Techradar.pro, April 17, 2018, https://www.techradar.com/news/german-government-goes-open-source-with-open-source-cloud-firm-nextcloud.

148 Figiola, "Cloud Computing: Background, Status of Adoption by Federal Agencies, and Congressional Action."
149 Frank Konkel, "The Details About the CIA's Deal With Amazon," *Atlantic,* July 17, 2014, https://www.theatlantic.com/technology/archive/2014/07/the-details-about-the-cias-deal-with-amazon/374632.
150 Charles McLellan, "Multicloud: Everything You Need to Know About the Biggest Trend in Cloud Computing," ZDNet, July 1, 2019, https://www.zdnet.com/article/multicloud-everything-you-need-to-know-about-the-biggest-trend-in-cloud-computing.
151 "Flexera 2020 State of the Cloud Report," Flexera, 2020, https://info.flexera.com/SLO-CM-REPORT-State-of-the-Cloud-2020.
152 McLellan, "Multicloud: Everything You Need to Know about the Biggest Trend in Cloud Computing."
153 "Cloud Security Complexity," Cloud Security Alliance, May 2019, https://cloudsecurityalliance.org/artifacts/cloud-security-complexity.
154 Aaron Gregg, "Pentagon Doubles Down on 'Single-Cloud' Strategy for $10 Billion Contract," *Washington Post,* August 5, 2018, https://www.washingtonpost.com/business/capitalbusiness/pentagon-doubles-down-on-single-cloud-strategy-for-10-billion-contract/2018/08/05/352cfee8-972b-11e8-810c-5fa705927d54_story.html.
155 Billy Mitchell, "DOD Defends Its Decision to Move to Commercial Cloud With a Single Award," FedScoop, March 8, 2018, https://www.fedscoop.com/dod-pentagon-jedi-cloud-contract-single-award.
156 Barry Rosenberg, "Oracle Vs. Pentagon: Why Judge Approved a Single Vendor for JEDI Cloud," Breaking Defense, July 29, 2019, https://breakingdefense.com/2019/07/oracle-vs-pentagon-why-judge-approved-a-single-vendor-for-jedi-cloud.
157 Daniel Gouré, "The Critics Are Wrong: A Single Award for the JEDI Contract Is the Right Approach," Lexington Institute, blog post, October 16, 2018, https://www.lexingtoninstitute.org/the-critics-are-wrong-a-single-award-for-the-jedi-contract-is-the-right-approach.
158 Adam Mazmanian, "CIA Plans Multibillion Cloud Buy for Intelligence Community," FCW, April 1, 2019, https://fcw.com/articles/2019/04/01/cia-cloud-c2e-multivendor.aspx.
159 "Summary of the Amazon S3 Service Disruption in the Northern Virginia (US-EAST-1) Region," AWS, https://aws.amazon.com/message/41926/.
160 Kenneth Hui, "AWS 101: Regions and Availability Zones," Rackspace Technology, blog post, February 16, 2017, https://www.rackspace.com/blog/aws-101-regions-availability-zones.
161 AWS, "Summary of the Amazon S3 Service Disruption in the Northern Virginia (US-EAST-1) Region."
162 Jordan Novet, "AWS Is Investigating S3 Issues, Affecting Quora, Slack, Trello (Updated)," *VentureBeat*, February 28, 2017, https://venturebeat.com/2017/02/28/aws-is-investigating-s3-issues-affecting-quora-slack-trello.
163 Alex Hern, "How Did an Amazon Glitch Leave People Literally in the Dark?"
164 Yevgeniy Sverdlik, "AWS Outage That Broke the Internet Caused by Mistyped Command," Data Center Knowledge, March 2, 2017, https://www.datacenterknowledge.com/archives/2017/03/02/aws-outage-that-broke-the-internet-caused-by-mistyped-command.
165 "Amazon S3," Similar Tech, https://www.similartech.com/technologies/amazon-s3.
166 "Amazon S3 Storage Classes," AWS, https://aws.amazon.com/s3/storage-classes.
167 "Postmortem: VSTS 4 September 2018," Microsoft Azure, *DevOps* blog post, September 10, 2018, https://devblogs.microsoft.com/devopsservice/?p=17485
168 Ibid.
169 Richard Speed, "Microsoft Azure: It's Getting Hot in Here, So Shut Down All Your Cores," *Register*, September 4, 2018, https://www.theregister.co.uk/2018/09/04/azure_its_getting_hot_in_here.

170 "Postmortem: VSTS 4 September 2018."
171 Tom Krazit, "Some Microsoft Azure Issues Stretch Into a Second Day, as Active Directory and Office 365 Stabilize," GeekWire, September 5, 2018, https://www.geekwire.com/2018/microsoft-azure-issues-stretch-second-day-active-directory-office-365-stabilize.
172 Thomas Claburn, "Azure North Europe Downed by the Curse of the Irish—Sunshine," *Register*, June 22, 2018, https://www.theregister.co.uk/2018/06/22/azure_north_europe_downed_by_pleasant_weather/; and Yevgeniy Sverdlik, "Microsoft Says Azure Outage Caused by Accidental Fire-Suppression Gas Release," DataCenter Knowledge, October 4, 2017, https://www.datacenterknowledge.com/uptime/microsoft-says-azure-outage-caused-accidental-fire-suppression-gas-release.
173 Aftab Siddiqui, "Route Leak Causes Major Google Outage," Internet Society, November 15, 2018, https://www.internetsociety.org/blog/2018/11/route-leak-caused-a-major-google-outage.
174 Dan Goodin, "Strange Snafu Misroutes Domestic US Internet Traffic Through China Telecom," Ars Technica, November 6, 2018, https://arstechnica.com/information-technology/2018/11/strange-snafu-misroutes-domestic-us-internet-traffic-through-china-telecom.
175 Dan Goodin, "Google Goes Down After Major BGP Mishap Routes Traffic Through China," Ars Technica, November 13, 2018, https://arstechnica.com/information-technology/2018/11/major-bgp-mishap-takes-down-google-as-traffic-improperly-travels-to-china.
176 Google, "Google Cloud Networking Incident #18018," Google Cloud Platform, https://status.cloud.google.com/incident/cloud-networking/18018.
177 Dan Goodin, "Suspicious Event Hijacks Amazon Traffic for 2 Hours, Steals Cryptocurrency," Ars Technica, April 24, 2018, https://arstechnica.com/information-technology/2018/04/suspicious-event-hijacks-amazon-traffic-for-2-hours-steals-cryptocurrency.
178 Chris Williams, "Trying to Log into Office 365 Right Now? It's a Coin Flip, Says Microsoft: Service Goes TITSUP as Azure Portal Wobbles," *Register*, January 29, 2019, https://www.theregister.co.uk/2019/01/29/office_365_outage.
179 "Azure Active Directory Outage and RCA for Azure Cloud Hiccups," Born's Tech and Windows World, February 2, 2019, https://borncity.com/win/2019/02/02/azure-active-directory-outage-rca-for-azure-cloud-hicups; and Danny Bradbury, "Microsoft Azure Data Deleted Because of DNS Outage," Naked Security, February 1, 2019, https://nakedsecurity.sophos.com/2019/02/01/dns-outage-turns-tables-on-azure-database-users.
180 Richard Speed, "Forget Snowmageddon, It's Dropageddon in Azure SQL World: Microsoft Accidentally Deletes Customer DBs," *Register*, January 30, 2019, https://www.theregister.co.uk/2019/01/30/azure_sql_delete.
181 Ibid.
182 Ibid.
183 "Born's Tech and Windows World, "Azure Active Directory Outage and RCA for Azure Cloud Hiccups."
184 Thomas Ricker, "Facebook Returns After Its Worst Outage Ever," *Verge*, March 14, 2019, https://www.theverge.com/2019/3/14/18265185/facebook-instagram-whatsapp-outage-2019-return-back; and Hamza Shaban, "Facebook, Instagram, and WhatsApp Suffered a Global Outage. What Happened?," *Washington Post*, March 14, 2019, https://www.washingtonpost.com/technology/2019/03/14/facebook-instagram-whatsapp-suffered-global-outage-what-happened.
185 Ibid.
186 Peter Judge and Will Calvert, "Facebook, Instagram, WhatsApp Suffer Global Outage," DataCenter Dynamics, March 14, 2019, https://www.datacenterdynamics.com/en/news/facebook-instagram-whatsapp-suffer-global-outage.

187 "Network Measurements Provide Insight Into Global Facebook Outages," NetBlocks, March 14, 2019, https://netblocks.org/reports/global-facebook-outage-analysis-Prybp2A7.

188 Jann Horn, "Reading Privileged Memory With a Side-Channel," Google, *Project Zero* blog post, January 3, 2018, https://googleprojectzero.blogspot.com/2018/01/reading-privileged-memory-with-side.html.

189 Brad Robinson, "A Simplified Explanation of the "Meltdown" CPU Vulnerability," Hackernoon, January 11, 2018, https://hackernoon.com/a-simplified-explanation-of-the-meltdown-cpu-vulnerability-ad316cd0f0de.

190 "Meltdown and Spectre," MeltdownAttack.com resource page, https://meltdownattack.com/#faq-systems-meltdown; and Zach Whittaker, "Critical Flaws Revealed to Affect Most Intel Chips Since 1995," ZDNet, January 3, 2018, https://www.zdnet.com/article/security-flaws-affect-every-intel-chip-since-1995-arm-processors-vulnerable.

191 Laura Hautala, "Spectre and Meltdown: Details You Need on Those Big Chip Flaws," CNet, January 8, 2018, https://www.cnet.com/news/spectre-meltdown-intel-arm-amd-processor-cpu-chip-flaw-vulnerability-faq.

192 Paul Miller, "Explaining Meltdown With Parallel Worlds, Libraries, and a Bank Heist," *Verge*, January 6, 2018, https://www.theverge.com/2018/1/6/16854668/meltdown-spectre-hack-explained-bank-heist-analogy.

193 Artem Dinaburg, "An Accessible Overview of Meltdown and Spectre, Part 2," *Trail of Bits* blog post, March 22, 2018, https://blog.trailofbits.com/2018/03/22/an-accessible-overview-of-meltdown-and-spectre-part-2.

194 Russell Brandom, "The CPU Catastrophe Will Hit Hardest in the Cloud," *Verge*, January 4, 2018, https://www.theverge.com/2018/1/4/16850120/meltdown-spectre-vulnerability-cloud-aws-google-cpu.

195 Ted Greenwald and Jack Nicas, "Intel Wrestled With Chip Flaws for Months," *Wall Street Journal,* January 5, 2018, https://www.wsj.com/articles/intel-wrestled-with-chip-flaws-for-months-1515110151?mod=article_inline.

196 Robert McMillan and Liza Lin, "Intel Warned Chinese Companies of Chip Flaws Before U.S. Government," *Wall Street Journal,* January 28, 2018, https://www.wsj.com/articles/intel-warned-chinese-companies-of-chip-flaws-before-u-s-government-1517157430.

197 Stephen Nellis, "Intel Did Not Tell U.S. Cyber Officials About Chip Flaws Until Made Public," Reuters, February 22, 2018, https://www.reuters.com/article/us-cyber-intel/intel-did-not-tell-u-s-cyber-officials-about-chip-flaws-until-made-public-idUSKCN1G62PS; and Tom Warren, "Intel Didn't Warn US Government About CPU Security Flaws Until They Were Public," *Verge*, February 23, 2018, https://www.theverge.com/2018/2/23/17043768/intel-meltdown-spectre-no-us-goverment-warning.

198 Patrick Tucker, "How Long Did the US Government Know About Spectre and Meltdown?" Defense One, February 6, 2018, https://www.defenseone.com/technology/2018/02/how-long-did-us-government-know-about-spectre-and-meltdown/145776/.

199 U.S. Senate, "Complex Cybersecurity Vulnerabilities: Lessons Learned From Spectre and Meltdown," U.S. Senate Committee on Commerce, Science, and Transportation hearing, July 11, 2018, https://www.commerce.senate.gov/2018/7/complex-cybersecurity-vulnerabilities-lessons-learned-from-spectre-and-meltdown; and Sean Lyngaas, "Senators Question Vulnerability Disclosure Process After Spectre and Meltdown Stumbles," Cyberscoop, July 11, 2018, https://www.cyberscoop.com/senators-question-vulnerability-disclosure-process-spectre-meltdown-stumbles.

200 John Thune, "Majority Statement," in "Complex Cybersecurity Vulnerabilities: Lessons Learned From Spectre and Meltdown," U.S. Senate Committee on Commerce, Science, and Transportation hearing, July 11, 2018, https://www.commerce.senate.gov/2018/7/complex-cybersecurity-vulnerabilities-lessons-learned-from-spectre-and-meltdown.

201 Tom Warren, "Microsoft Halts AMD Meltdown and Spectre Patches After Reports of Unbootable PCs," *Verge*, January 9, 2018, https://www.theverge.com/2018/1/9/16867068/microsoft-meltdown-spectre-security-updates-amd-pcs-issues.
202 Tom Warren, "Epic Games Blames Meltdown CPU Performance Issues for Fortnite Downtime," *Verge*, January 6, 2018, https://www.theverge.com/2018/1/6/16857878/meltdown-cpu-performance-issues-epic-games-fortnite.
203 John Leyden, "Meltdown and Spectre, One Year On: Feared CPU Slowdown Never Really Materialized," Daily Swig, January 31, 2019, https://portswigger.net/daily-swig/meltdown-and-spectre-one-year-on-feared-cpu-slowdown-never-really-materialized.
204 Catalin Cimpanu, "Researchers Discover Seven New Meltdown and Spectre Attacks," ZDNet, November 14, 2018, https://www.zdnet.com/article/researchers-discover-seven-new-meltdown-and-spectre-attacks.
205 Lily Hay Newman, "The Elite Intel Team Still Fighting Meltdown and Spectre," *Wired*, January 3, 2019, https://www.wired.com/story/intel-meltdown-spectre-storm.
206 "Information on Capital One Cyber Incident," Capital One, updated September 23, 2019, https://www.capitalone.com/facts2019.
207 Riyaz Walikar, "An SSRF, Privileged AWS Keys and the Capital One Breach," Appsecco, August 4, 2019, https://blog.appsecco.com/an-ssrf-privileged-aws-keys-and-the-capital-one-breach-4c3c2cded3af.
208 "Seattle Tech Worker Arrested for Data Theft Involving Large Financial Services Company," U.S. Department of Justice, press release, July 29, 2019, https://www.justice.gov/usao-wdwa/pr/seattle-tech-worker-arrested-data-theft-involving-large-financial-services-company.
209 U.S. Department of Justice, "Seattle Tech Worker Arrested for Data Theft Involving Large Financial Services Company."
210 *U.S. v. Paige Thompson*, United States District Court for the Western District of Washington at Seattle, Criminal Complaint, July 29, 2019, https://www.justice.gov/usao-wdwa/press-release/file/1188626/download.
211 James Rundle and Catherine Stupp, "Capital One Breach Highlights Dangers of Insider Threats," *Wall Street Journal,* July 31, 2019, https://www.wsj.com/articles/capital-one-breach-highlights-dangers-of-insider-threats-11564565402.
212 Jenny Surane and Lananh Nguyen, "Capital One Touted the Data Cloud's Safety. Then a Hacker Breached It," *Los Angeles Times,* July 31, 2019, https://www.latimes.com/business/story/2019-07-30/capital-one-cloud-safety-hacker-breach.
213 Robert McMillan, "Capital One Breach Casts Shadow Over Cloud Security," *Wall Street Journal,* July 30, 2019, https://www.wsj.com/articles/capital-one-breach-casts-shadow-over-cloud-security-11564516541; and David Henry, "Capital One Customer Data Breach Rattles Investors," Reuters, July 30, 2019, https://www.reuters.com/article/us-capital-one-fin-cyber-amazon-com/capital-one-customer-data-breach-rattles-investors-idUSKCN1UP1LD.
214 Jim Jordan, Michael Cloud, and Mark Meadows, letter to Richard D. Fairbank, U.S. House of Representatives, Committee on Oversight and Reform, August 1, 2019, https://republicans-oversight.house.gov/wp-content/uploads/2019/08/2019-08-01-JDJ-MC-MM-to-Fairbank-Cap-One-re-Capital-One-Breach.pdf; and Jim Jordan, Michael Cloud, and Mark Meadows, letter to Jeff Bezos, U.S. House of Representatives, Committee on Oversight and Reform, August 1, 2019, https://republicans-oversight.house.gov/wp-content/uploads/2019/08/2019-08-01-JDJ-MC-MM-to-Bezos-AWS-re-Capital-One-Breach.pdf.

215 Ron Wyden, letter to Jeff Bezos, U.S. Senate, August 5, 2019, https://online.wsj.com/public/resources/documents/AmazonLetter080519.pdf?mod=article_inline; and Elizabeth Warren, letter to Richard D. Fairbank, U.S. Senate, August 8, 2019,https://www.warren.senate.gov/imo/media/doc/2019.08.08%20Letter%20to%20Capital%20One%20re%20recent%20massive%20data%20breach.pdf.

216 Jason Murdock, "Amazon Refuses Blame for Capital One Data Breach, Says Its Cloud Services Were 'Not Compromised in Any Way,'" *Newsweek*, July 30, 2019, https://www.newsweek.com/amazon-capital-one-hack-data-leak-breach-paige-thompson-cybercrime-1451665.

217 Brian Krebs, "What We Can Learn From the Capital One Hack," *Krebs on Security*, blog post, August 19, 2019, https://krebsonsecurity.com/2019/08/what-we-can-learn-from-the-capital-one-hack.

218 Rob Wright and Chris Kanaracus, "Capital One Hack Highlights SSRF Concerns for AWS," Tech Target, August 5, 2019, https://searchsecurity.techtarget.com/news/252467901/Capital-One-hack-highlights-SSRF-concerns-for-AWS.

219 "Announcing Updates to Amazon EC2 Instance Metadata Service," AWS, November 19, 2019, https://aws.amazon.com/about-aws/whats-new/2019/11/announcing-updates-amazon-ec2-instance-metadata-service.

220 "Two Chinese Hackers Associated With the Ministry of State Security Charged With Global Computer Intrusion Campaigns Targeting Intellectual Property and Confidential Business Information," U.S. Department of Justice, press release, December 20, 2018, https://www.justice.gov/opa/pr/two-chinese-hackers-associated-ministry-state-security-charged-global-computer-intrusion.

221 Jack Stubbs, Joseph Menn, and Christopher Bing, "Inside the West's Failed Fight Against China's 'Cloud Hopper' Hackers," Reuters, June 26, 2019, https://www.reuters.com/investigates/special-report/china-cyber-cloudhopper.

222 Ibid.

223 Rob Barry and Dustin Volz, "Ghosts in the Clouds: Inside China's Major Corporate Hack," *Wall Street Journal*, December 30, 2019, https://www.wsj.com/articles/ghosts-in-the-clouds-inside-chinas-major-corporate-hack-11577729061.

224 Jack Stubbs, Joseph Menn, and Christopher Bing, "Inside the West's Failed Fight Against China's 'Cloud Hopper' Hackers."

225 Prashanth Chandrasekar, "What Is Managed Cloud?" *Rackspace,* blog post, May 4, 2019, https://www.rackspace.com/blog/what-is-managed-cloud#:~:text=What%20is%20a%20managed%20cloud%20services%20provider%3F,infrastructure%20and%20application%20level%20support; and "Managed Service Provider Partners" AWS, accessed August 16, 2020, https://aws.amazon.com/partners/msp.

226 "Operation Cloud Hopper," PwC and BAE Systems, April 2017, https://www.pwc.co.uk/cyber-security/pdf/cloud-hopper-report-final-v4.pdf.

227 Stubbs, Menn, and Bing, "Inside the West's Failed Fight against China's 'Cloud Hopper' Hackers."

228 Stubbs, Menn, and Bing, "Inside the West's Failed Fight against China's 'Cloud Hopper' Hackers."

1779 Massachusetts Avenue NW | Washington, DC 20036 | P: + 1 202 483 7600

CarnegieEndowment.org