CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

JANUARY 2024

Cloud Reassurance: A Framework To Enhance Resilience And Trust

John Pendleton, Ariel (Eli) Levite, and Bob Kolasky

Cloud Reassurance: A Framework to Enhance Resilience and Trust

John Pendleton, Ariel (Eli) Levite, and Bob Kolasky

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

Summary

Moving information and services to the cloud offers agility and security and allows users to outsource their information technology (IT) needs. Instead of maintaining their own computer servers and IT staff, consumers are increasingly relying on cloud service providers (CSPs) to store and process their information as well as the software to support their work. Over time, however, cloud services have become concentrated in a handful of providers, with around two-thirds of cloud services being managed by three "hyperscale" providers.

This paper explores the challenges associated with this concentration of cloud services, the risks that can result, and actions to manage those risks. The assessment focuses on *resilience*—that is, the ability to anticipate and prepare for hazards, reduce their impact, and recover from them.

We recommend that private-sector organizations adopt a *Cloud Resilience Framework*. The framework developed as part of this project lays out foundational policy commitments and suggests actions that would enhance both the resilience of the cloud system and trust in that system. The Cloud Resilience Framework provides a comprehensive and proactive way for cloud providers to build on existing efforts while incorporating more stakeholders—including insurers and governments—with a common goal of ensuring resilience. The four-part framework is summarized below:

Framework area	What is it?
Foundational commitments	Public commitments to advance cloud-related security and resilience and minimize digital harms.
Resilience of the cloud system	Actions that cloud providers can take to demonstrate and increase resilience of their cloud services.
Resilience of customers	Working with customers, insurers, and other stakeholders to develop a standards-based Resilience Maturity Model.
Exercises and stress tests	Scenario-based exercise programs to validate contingency plans and test capabilities as well as identify best practices and lessons learned.

Implementing the framework will require sustained senior-level attention at the CSPs; moreover, some parts of the framework require partnership with external stakeholders. For example, developing and implementing a Resilience Maturity Model based on agreed-upon standards could help cloud customers better understand their own level of resilience while providing insurers and policymakers the information they need to effectively help stakeholders manage risk. The final part of the framework—expanding the scope and rigor of exercises and including outside participants in them—can ultimately serve as a stress test of the systems and build confidence among outside stakeholders going forward.

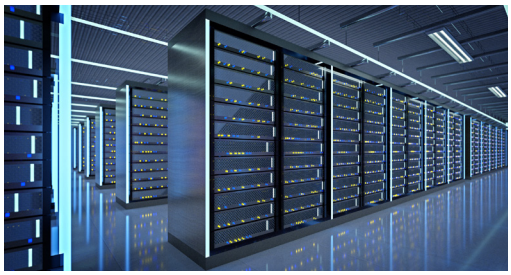
The paper also offers guidance for government policymakers wishing to enable and support the proposed Cloud Resilience Framework. Regulatory actions can have welcome benefits; they can also cause inefficiencies if not well designed and implemented. Given the challenges around harmonizing regulations, the path forward should emphasize the cross-sectoral criticality of cloud services; insist on improved transparency of risk information; and support the development of a functioning re/insurance market; among others. Continuing to collaborate with industry will be crucial. Government can help catalyze cloud resilience and has a responsibility to protect communities and citizens from digital harm. That role can be partially fulfilled by regulation, but it must also involve streamlining authorities, taking steps to harmonize requirements, and utilizing mandates to promote better risk understanding and address market failures. Cloud services are a shared benefit to societies and their availability and resilience needs to be recognized as a shared goal.

Looking to the future, the rapid emergence of artificial intelligence (AI) is intensifying the spotlight on the potential risks of technology. The large language models being developed require massive computer resources, which are provided by cloud services. As society and commerce become even more reliant on an AI-enmeshed cloud, the resilience of that cloud will be crucial. This paper offers ways to ensure that resilience going forward.

Introduction

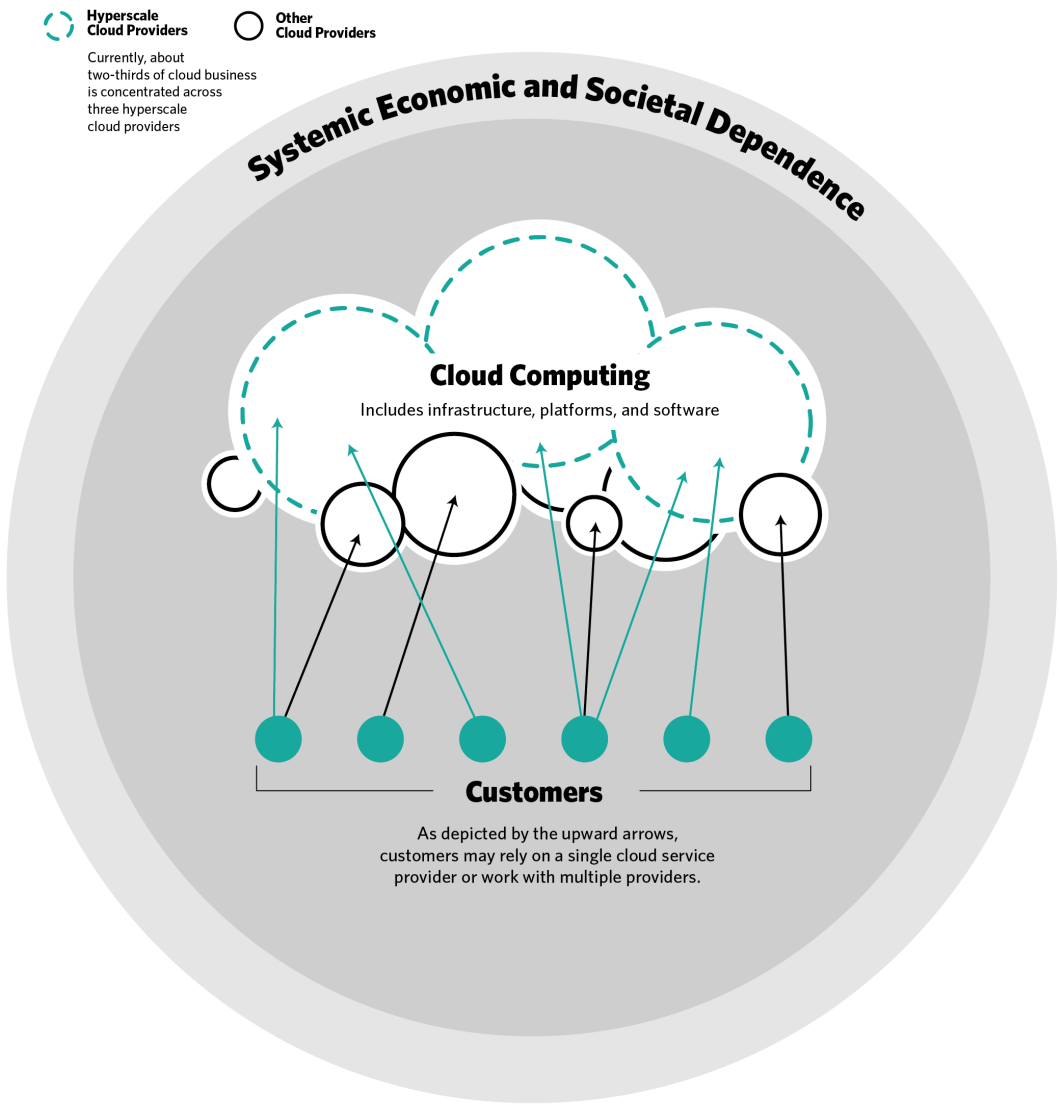
Society's dependence on cloud computing seemed to grow slowly at first, then rapidly, as the COVID-19 pandemic accelerated an ongoing transformation. With an internet connection, anyone can connect to colleagues worldwide to work from home and then stream entertainment in leisure time. Banking and shopping are increasingly efficient with mobile devices, which seamlessly switch to provide pinpoint accurate GPS navigation with real-time traffic and weather forecasts. Each of these modern conveniences is enabled by a massive technology infrastructure of energy and telecommunications and finance (see Figure 1)—increasingly powered by cloud services to make them work.

Figure 1: Server Facilities



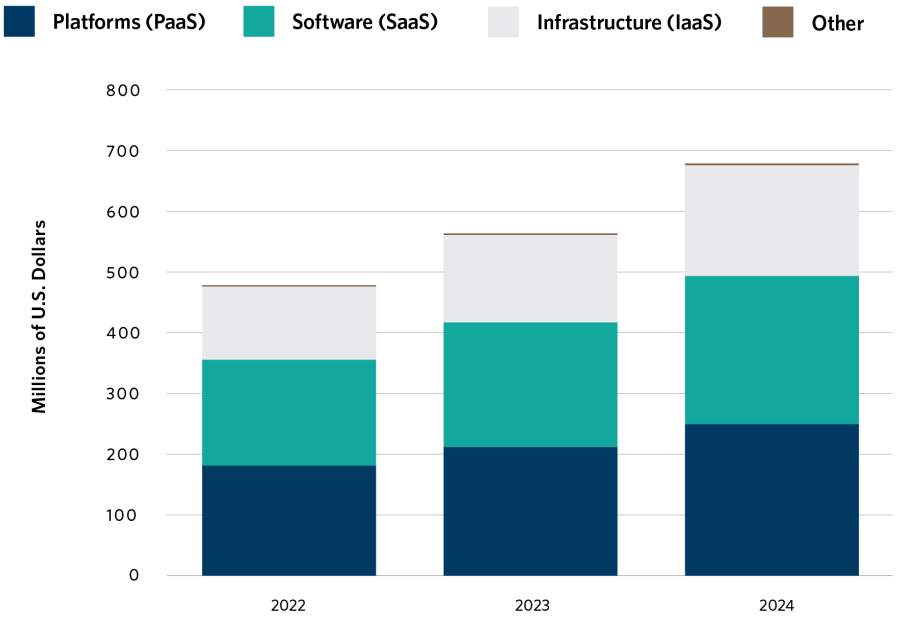
A June 2023 interim report on this project described how moving information and services to the cloud offers greater agility, lower costs, and better basic security for businesses and governments.[1] No longer must businesses and governments maintain servers—along with the information technology (IT) staff necessary for that maintenance—exclusively on their

Figure 2: Growing Economic and Societal Dependence on Hyperscale Cloud Services Providers



premises. Instead, customers can purchase services from cloud providers who do much of the underlying IT work behind the scenes. Over time, cloud service providers (CSPs) have grown to offer a range of services (storage, processing, and applications); these range from simply providing computing infrastructure, to hosting a customer's software and data and offering multiple applications (some of their own, others by third parties) and, of late, access to generative artificial intelligence (AI) services. CSPs promise security, reliability, and safety as key selling points. While the transition away from primarily on-premises IT is generally viewed as broadly improving cybersecurity, the concentration of services in a few "hyper-scale" providers has also resulted in growing economic and societal dependence, as shown in Figure 2 below.

Figure 3: Worldwide Spending on Public Cloud Services Is Rapidly Increasing



Source: "Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach $679 Billion in 2024," https://www.gartner.com/en/newsroom/press-releases/11-13-2023-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-reach-679-billion-in-20240.
Note: "Platforms (Paas)" also includes business cloud process service.

In 2023, the three hyperscale U.S. cloud service providers that participated in our study accounted for about two-thirds of the worldwide market for Infrastructure as a Service (IaaS) and Platform as a Service (PaaS).[2] By 2025, enterprise IT spending on public cloud computing, within addressable market segments, will overtake spending on traditional IT. Almost two-thirds of spending on application software will be directed toward cloud technologies in 2025.[3] The market for cloud services (including IaaS, PaaS, and Software as a Service [SaaS]) has been growing rapidly and is expected to reach nearly $680 billion in 2024 as shown in Figure 3. The deployment of generative artificial intelligence services is expected to further increase reliance on cloud services, given the scale of the infrastructure required.[4]

Given high-profile breaches and ransomware incidents such as SolarWinds and Log4Shell, most discussions about technology use, including cloud usage, have revolved around cybersecurity, often including concerns about unauthorized access of confidential data. We have taken a broader perspective to identify the potential for *residual* risks in the cloud beyond cybersecurity. Such residual risks might stem from natural disasters or operational issues including "misconfiguration" (in other words, both external and internal risks). Such events are low frequency but high impact, and they are difficult to predict or mitigate but could potentially render cloud facilities inoperable at the regional or system-wide level. The aim of this inquiry is to make the overall system more *resilient*, and our work aims to *reassure* the range of stakeholders that might be affected that extensive cloud dependence has not created significant unmanaged risk that presents an imminent challenge for national and economic well-being.

What do we mean by overall cloud system resilience?

The ability of communities, businesses, and nations vulnerable to cloud-related incidents to anticipate and prepare for, reduce the impact of, cope with, and recover from the effects of shocks and stresses without compromising their long-term prospects.

In the context of the cloud, resilience includes the capacity to withstand incidents and the capacity to recover from them, on the part of both cloud providers and the customers who rely on them. For example, a resilient cloud system would be able to contain an incident to avoid cascading failures and would enable speedy recovery and remediation with minimal impact on customers. In addition, incidents would inspire adaptations and evolution to mitigate future risks.

To consider this in depth, Carnegie convened a group of participants including the hyperscale cloud providers and reinsurers, along with independent experts with a broad range of previous government experience both in the United States and Europe. This group brought together leading market players in cloud services and insurance to discuss the potential risks that accompany the myriad benefits of growing cloud dependence while putting the conversation into a broader context. In the June interim report, we considered the potential for systemic risks created by cloud concentration and the scenarios that might be faced; the report also provided an overview of the complex policy challenges facing governments as they consider potential regulatory and nonregulatory approaches to mitigating resultant risks.[5] This final report moves beyond those considerations to address the following key questions:

- What are the challenges, if any, associated with growing cloud dependence?
- How do the hyperscale cloud providers ensure resilience at present?
- What are the residual risks, if any?
- What is the policy environment?
- What actions might be taken in the private sector?
- What actions should policymakers consider?

Challenges Associated with Growing Cloud Dependence

As discussed in the introduction, the benefits of the cloud are vast and growing, but society's increased adoption of cloud computing has also resulted in increasing scrutiny of the potential risks involved. To date, cloud service outages caused by human errors, natural disasters, or malicious activity have generally not resulted in widespread, long-term interruptions or loss of availability. However, cloud services have been targeted by criminal groups and foreign intelligence actors who have (among other tactics) exploited identity management and encryption key management in cloud systems to access multiple customers; such incidents have raised concerns about data integrity and reliability. Moreover, the market concentration among the hyperscale cloud providers has raised apprehension about the potential for a single point-of-failure incident where a breakdown of hardware or common software vulnerability leads to a prolonged disruption with insufficient contingency options or backup plans.

Gaps in Expertise and Knowledge Make Cloud Risk Assessment Difficult

Cloud providers have varying access into their customers' cloud environments. Cloud providers and customers have commercially sensitive intellectual property (IP), and providers are bound by their agreements with customers to protect customer information, safeguard their privacy, and prevent unauthorized access. As a result, the cloud providers' detailed knowledge of customer workloads is generally at the customer's discretion.

Some customers, especially larger enterprise customers, partner with CSPs to build and configure their environments for their most critical workloads, thus giving the CSP in-depth knowledge of their architecture and internal controls. By contrast, other customers essentially rent compute and storage capabilities on a "self-service" basis. The CSP's knowledge of such environments is far more limited, such as analysis of billing data and network traffic patterns. In these cases, cyber hygiene measures—such as managing supply chain risks or ensuring application resilience—are managed by the customer. However, as customers have relied more on the cloud for IT services, some enterprises (especially small and medium-sized ones) have pared down their in-house IT staff to save money. As a result, many customers may not have the expertise in-house to fully understand their risks and evaluate, deploy, and maintain the appropriate services for their risk profile.

Against this backdrop, stakeholders such as insurers and governments desire information about cloud operations so they can assess the risks to which customers, shareholders, and citizens are exposed. Cloud provider participants noted that they are subject to international standards and multiple certification regimes, as well as participating in information sharing and analysis centers, among other things. However, these measures provide only a partial

view of how a CSP and its customers would respond if faced with a major, unforeseen event. This makes it especially difficult for third parties to manage risks when it comes to cloud operations, as they do not have extensive visibility into CSP and customer vulnerabilities, but could nevertheless be affected by a disruption.[6] For example, insurers have expressed concern that they could suffer large losses in the event of a major disruption, and some governments are concerned about the real-world impact on their constituents in the wake of a natural disaster or other force majeure if cloud services were significantly affected.

In a February 2023 report about its ongoing work to assess the implications of cloud dependence in the financial sector, the U.S. Treasury Department found that many financial institutions—which are tightly regulated and perhaps best equipped to appreciate and manage risk—are still struggling to obtain and digest information related to the risks associated with cloud services. These risks include software dependencies within the concentrated cloud market where a vulnerability might affect many customers; supply chain risks such as those encountered in SolarWinds; lack of information about the results of testing of resilience and security capabilities; and lack of information regarding operational incidents, including real-time updates and after-action reports.

Evolving Risk Profile Challenges Insurability

In many sectors, insurance serves as a depository of knowledge on incidents and risk analysis, incentivizes and shares good practices, mitigates losses, and frees up capital, among other things. These functions all contribute to resilience.[7] But cybersecurity risks have thus far proven to be a complex analytical challenge that is not fully understood and, as a result, such cyber risks are likely significantly underinsured.[8] In late 2022, a report by a reinsurer estimated that about 90 percent of cybersecurity risks remain uninsured even as worldwide cybersecurity losses (most not attributable to cloud services) approached a trillion dollars.[9]

Notably, increased reliance on cloud services—ranging from infrastructure to common cloud-based software—appears to have fundamentally transformed the insurance risk profile. Overall, insurers believe that cloud adoption has diminished the likely number and size of relatively small losses due to the presumably improved level of cybersecurity and resilience posture it enables. However, due to the aggregation/accumulation effects, the largest possible losses could be much larger than they were prior to widespread adoption of the cloud. This is more fully discussed in Appendix 1 via a technical analysis that describes reinsurers' point of view on cloud-related risks.

As a practical matter, the reinsurers who participated in this study—who provide insurance for insurers—told us that cloud risks are exceptionally difficult to assess and underwrite.[10] Lacking precise information about cloud operations and potential risks, insurers must develop their own estimates or rely on third-party estimates, which vary widely but typically paint a sobering picture of the likely financial losses associated with cloud failure. The challenges associated with estimating cyber risks are discussed in a November 2023 report by

The Geneva Association, a global association of insurance and reinsurance companies. The report notes that some estimates indicate that insured losses from a rare industry-wide cyber incident (such as a cloud outage) could approach that of a natural disaster, but that estimates vary widely and are extremely sensitive to the assumptions used. Also, the report describes several actions that cloud providers take that may serve to mitigate risk, such as separating operations into multiple, regional "availability" zones.[11]

The reinsurer participants in our study were particularly concerned with their ability to estimate the accumulation of losses from "tail risks" that may result from cloud concentration. Tail risks refer to events that are low-probability (perhaps as rare as once in 200 years) but high-consequence: weighing these risks is integral to insurance assessments and risk tolerance and can be part of insurers' regulatory requirements. Some have called for governments to step in as an "insurer-of-last-resort" to alleviate concern about such cyber-catastrophe tail risks. In the wake of the September 11 terrorist attacks, for example, the United States created the Terrorism Risk Insurance Program to assure insurance markets that large-scale losses would be covered by the U.S. government. Currently, the U.S. Treasury Department is assessing whether a federal backstop is similarly warranted to free up insurance capital by allaying the concerns about cyber tail risks.[12] The November 2023 Geneva Association report concluded that "suitably designed, state-sponsored backstops could encourage re/insurers to extend coverage, promote good cyber hygiene and ensure that governments only face extreme losses above an agreed threshold." Even if a federal backstop is put in place, it would not overcome some of the consequences of the information gaps described in the previous section, which present an arguably more fundamental challenge to both insurers and governments. Unless and until these information gaps are filled, the cyber insurance market may continue to be subject to exclusions and limitations that make it less attractive to customers and less effective at managing cyber risks in the commercial sector.

Ongoing Efforts to Ensure Cloud Resilience

Cloud resilience is simultaneously the responsibility of the CSPs and the customers. For our purposes, we recognize this distinction as resilience of the cloud (CSPs) and resilience in the cloud (customers).

Recognizing that this reality involves a measure of shared responsibility, CSPs are tasked with ensuring the confidentiality, integrity, and availability of the underlying infrastructure, networks, and servers, as well as the building blocks of cloud-based applications. Customers, for their part, are responsible for implementing security controls to defend the applications they build on top of cloud infrastructure, as well as for architecting resilient applications. For example, cloud providers can offer an infrastructure consisting of several interconnected data centers, with service level agreements (SLAs) that specify guaranteed

metrics for performance, availability, and security. But it is ultimately up to the customer to determine how and where data and applications should be hosted and processed and how many data centers should hold backup copies, among other considerations. This will vary based on the offerings provided by the CSP and the customer's performance needs, risk appetite, and budget.

The CSPs in our study identified a range of steps that they take to ensure security and reliability of the cloud, which are detailed below.

Resilience of the Cloud

CSPs' main actions to ensure that their services are robust and reliable include training CSP employees to understand the importance of security and privacy; monitoring for threats; developing incident response plans; maintaining physical security; vetting third-party suppliers; encrypting data; and conducting audits and system testing. Table 1 lists efforts that are common across the CSPs who participated in this study.

Table 1: Efforts to Maintain Secure and Reliable Cloud Services

Area	Description
Security culture	Employees undergo ongoing security and privacy training and agree to a code of conduct to ensure the confidentiality, integrity, and availability of user systems and data.
Vulnerability management	Internal vulnerability management processes actively scan CSPs' global infrastructure for threats using a) software security reviews, b) automated and manual penetration testing, c) external audits, d) bug bounty programs, and e) threat intelligence.
Security monitoring	CSPs analyze network activity and collect security telemetry at many points across the global network to identify potential threats using proprietary and third-party tools.
Incident management	CSPs maintain detailed incident response plans, including procedures for a) incident detection, b) escalation, c) investigation, d) containment and recovery, and e) continuous improvement Incident management teams notify impacted customers and provide customer support, per pre-defined SLAs.
Physical data center security	Data centers rely on a layered physical security model including a) perimeter fencing and vehicle barriers, b) electronic access cards, c) access logs, d) 24/7 activity monitoring, and e) server and hardware hardening against tampering. In addition, staff undergo rigorous background checks and training.
Data center operations	All critical components within data centers have primary and alternate power sources and backup generators and cooling systems.
Hardware supply chain security	CSPs vet hardware suppliers (such as servers and network equipment) and track the location and status of all hardware in their supply chain to ensure reliability and security of components.
Software development practices	CSPs limit the potential for vulnerabilities by using secure development techniques such as a) source control protections, b) two-party code reviews, c) pre-vetted libraries and frameworks, and d) automated code scanning tools.
Encryption	CSPs encrypt customer data at rest (within the data center) and in transit (between data centers).
Network security	Firewalls and front-end servers prevent unauthorized access to CSPs' global networks and mitigate the risk of DDoS attacks.
Identity and access management	CSPs track and manage access to systems and data by all users, devices, and applications on their networks. The principle of least privilege (often described as "zero trust") means users receive the minimum amount of access necessary to perform their functions.
High-availability networking	Networks are designed to be highly redundant and to dynamically reroute data flows, minimizing dependencies on any single server, data center, or network connection. Data centers are geographically distributed to minimize the impacts of disruptions caused by natural disasters, accidents, or other events.
Audit and compliance	Internal audit teams review products with security laws and regulations and develop processes to meet new requirements.
Systems testing	CSPs undertake structured testing of their systems and infrastructure to ensure continued availability in a range of outage scenarios, which are categorized into tiers based on organizational breadth and anticipated levels of simulated operational disruption.

Source: Microsoft Azure, Google Cloud Services, and Amazon Web Services

As previously noted, CSPs also typically seek to conform to international and insurance standards, such as ISO 27001, ISO 27017, and ISO 27110, as well as participate in other programs such as the Cloud Security Alliance's Security, Trust, and Assurance Registry and the Center for Internet Security Benchmarks.

Resilience in the Cloud

CSPs explained that they do not have in-depth knowledge of how every customer uses their cloud services due to default security measures that prevent CSPs from having access to customer data. Within that context, however, they identified several areas where they provide resources for or work with some customers to assist in designing secure and resilient cloud deployments. The SLA contracts between CSPs and customers identify the terms and agreed-to performance parameters. Beyond SLAs, CSPs increasingly offer services (complimentary or paid) in the form of trainings, white papers, dashboards, and even technical specifications. These offerings are intended to help customers build knowledge by contributing to their understanding of the cloud and thereby reduce the likelihood of issues or data incidents because of misconfiguration. Some variation exists in customers' (especially small- to medium-sized customers') capacity to fully understand the information made available to them to enhance their resilience, as well as in the resources they commit to implementing best practices or seeking additional support. Nonetheless, while CSPs must abide by privacy restrictions, they are in the best position to understand their customers and those customers' level of dependency on and resilience in the cloud, especially for critical functions.

Table 2 lists efforts that are common across the CSPs who participated in this study.

Table 2: Efforts to Ensure and Enhance Customer Security and Resilience

Area	Description
Service level agreements	Contractual terms between CSP and customer delineating CSP responsibility for meeting key metrics for performance, availability, customer support, incident management, and others.
Infrastructure security blueprints	Curated guidance and automated tools to optimize cloud-native controls and reduce the risk of infrastructure misconfigurations.
Application security blueprints	Curated guidance and automated tools to help customers design, build, operate, and configure applications for security in the cloud and meet specific compliance requirements.
Security management tools	Dashboards that enable customers to monitor the health of their cloud environments, through a) providing visibility into resources and policies, b) identifying misconfigurations and application vulnerabilities, and c) detecting compliance violations based on industry standards.
Architecture frameworks	Frameworks and design patterns that help customers build reliable cloud architectures designed for high availability and resilience in the event of regional outages.
Security posture reviews	Security and resilience posture reviews in the form of audits, penetration testing, and/or tabletop exercises to identify gaps in customers' security controls or incident response playbooks.
Training and certificate programs	Free training and certificate programs help customer teams build cloud security skills and expertise.
Threat intelligence	Complimentary or paid threat intelligence services to provide insight into common attack tools and techniques, enabling customers to implement countermeasures.

Source: Microsoft Azure, Google Cloud Services, and Amazon Web Services

Residual Risk Assessment

Serious effort to enhance the resilience of cloud services must start with an assessment of the risks involved in cloud dependency that does not lose sight of the benefits that accrue from reliance on these services. It then ought to highlight deficiencies and point to desirable practices that could provide an acceptable level of cloud resilience without either stifling innovation or diminishing cloud benefits. Yet notwithstanding the appeal of such a logical progression, implementing resilience assessments is challenging.

As discussed, cloud services' resilience is a function of both resilience of the cloud and resilience in the cloud. The latter is heavily dependent on customers' decisions and practices in configuring their cloud connectivity and dependence including the functions they assign to the cloud. Further complicating the picture is the rapid evolution of cloud services and

providers' concerns about sharing proprietary information about cloud operations. And while CSPs report data on uptime and other metrics, no consensus standard exists to measure or assess cloud resilience, thereby limiting efforts to test it.

Assessments of any technology must also accept a sobering reality: a measure of risk accompanies any human activity, especially technology-dependent ones. Efforts to drive out all risk, or even come close to it, however well intentioned, may not be possible—indeed, they are potentially wasteful and may stifle innovation. Trying to regulate away technological risk can have other adverse consequences, like inducing complacency or encouraging a compliance approach to a complicated and evolving problem. For all these reasons, this study has adopted a different approach toward assessing and presenting the risks associated with cloud dependency.

At the center of our approach lies the determination to focus on the *residual cloud risks* that meet three conditions:

1. Systemic consequences: These are risks that could produce adverse systemic consequences (materially effecting numerous entities critically dependent on cloud services) or less widespread adverse impacts that nonetheless result in socially unacceptable effects;

2. Triggered from external or internal factors: These risks could be triggered by any source, from natural disasters to malicious external actors to insider threats, or by technical failures and human errors; and,

3. Potential for digital harm: These risks may manifest themselves in some form of digital harm, ranging from loss of availability, to compromise of confidentiality and/or integrity, to loss of trust in the entire cloud service.

Importantly, our approach takes into consideration that the hyperscalers partaking in this project are already deploying some significant mitigation measures (as outlined in the previous section). These measures diminish the probability that some grave threats to cloud services will materialize while also diminishing the impact if they do. Put differently, in considering residual risk, we try to assess, however crudely, which risk scenarios might still reach the level of systemic concern, even after these mitigation measures are deployed.

Here, a key factor is the distinct possibility that risks could arise not solely because the resilience *of the cloud* falls short but also because of inadequate resilience measures by customers *in the cloud*. In addition, in contrast to large, sophisticated, and well-resourced cloud service providers, there are innumerable small and medium-sized cloud customers who may lack the sophistication and resources to understand the potential risks they incur if they do not actively participate in maintaining resilience.

Given the difficulties in massing quantifiable data about the level of cloud resilience, coupled with the challenge of providing a generic assessment of the level of residual risk, we are focusing here primarily on the residual risk assessment for the major cloud providers. Furthermore, although we do consider the impact of cascading effects of cloud disruptions on consumers and societies at large, our study group did not include customers of cloud services, so we were unable to fully consider consumers' views on the level of cloud dependence and resilience.[13]

Toward that end, we have worked with the project's expert participants to brainstorm, debate, and finally settle on the following summary of our residual risk assessment, recognizing that to do so credibly we have to settle for an ordinal scale of measurement. Additionally, taking into consideration the generic nature of our exercise, as well the centrality of the distinction between risks in and of the cloud, we elected to orient our residual risk assessment toward generating a strategic situational awareness rather than creating a blueprint for attributing responsibility and, consequently, liability for addressing cloud risks. The primary purpose is to highlight outstanding concerns and encourage actions for all those stakeholders whose fortunes are tied to the cloud to undertake in the context of "a shared fate model."

The Residual Cloud Risk Assessment

Our assessment considered a variety of factors and elements.

Availability disruption: This is the most common cloud risk; in fact, such disruptions have happened already. Consequently, it is also by far the best understood cloud risk scenario and the one CSPs are most experienced with and prepared to tackle. Their sustained efforts and considerable investments mean that prolonged cloud service disruptions appear to be unlikely. Should these disruptions occur in the future, their geographical scope could be wide, but the disruption of cloud services is unlikely to last more than a day, in most cases. It is, however, possible to envisage scenarios that take longer to sort out, such as those associated with authentication/credential management and/or software supply chain attack. The ransomware threat is growing and evolving and warrants special attention because of its potential to severely disrupt availability. A small cloud provider in Norway, for instance, suffered a devastating ransomware attack in August 2023 that resulted in the loss of all customer data.[14] Such a scenario seems highly unlikely with hyperscalers, whose sophistication, resources, and business strength should suffice to withstand and recover from such an event.

An outage caused by a *logical* trigger (like a software glitch or malware) is likely to have only a temporary impact on cloud service availability. Yet significant *physical* events that affect core cloud infrastructure (such as the chillers that manage the heat generated by servers and other infrastructure) or the utilities cloud services depend upon (like energy, telecommunications, or water) could potentially affect cloud service availability for much longer unless significant redundancy is available and rapidly deployable. Damage to underwater data

cables might be the most likely scenario in this category. That said, in circumstances other than war and natural disaster, these disruptions of CSP services are likely to be localized, or at most regional. Nonetheless, both physical and logical disruptions of availability have the potential to have a significant impact on consumers' operations. The duration and gravity of such impacts would likely vary greatly among consumers, depending on the services those consumers have transferred to the cloud as well as their level of resilience to cope with such scenarios.

Confidentiality: Less attention has been given to scenarios associated with breach of data stored in the cloud. But such breaches happened before and may well occur again. Extensive breach of data confidentiality in cloud services, were it to happen, would probably be short-lived once detected and treated. But it may take a while to detect the vulnerabilities that the intruders exploit to produce such compromise and to stop them entirely. The effects of breaches will likely be uneven depending on the level of encryption of data stored on the cloud. Once data are compromised, however, the effect will be irreversible and the impact on its original owners (in terms of privacy as well as intellectual property) could be significant.

Integrity: Perhaps the least understood risk scenario is one involving integrity compromise. Discovery of heretofore unknown vulnerabilities in the cloud could result in improper, accidental, or malicious modification of data stored on the cloud, potentially triggering concerns about the integrity of cloud-based applications. If an integrity compromise is exploited by a persistent and sophisticated attacker, it could take many months to sort out its impact and restore full confidence in the integrity of the data and applications. Longer dwell time (the period an integrity compromise has been present in the system) could significantly extend the mitigation timeline. During this time, the functionality of the cloud service as well as the services based on it could be significantly slowed, especially if no backups are available. Importantly, some vulnerabilities (especially in hardware and open-source software) may be common to several CSPs and require involvement of third parties to resolve, potentially multiplying the effects of integrity compromises.

Fragility of trust: Societies, enterprises, and individuals have become heavily dependent on cloud services for many of their tasks, benefitting widely from cloud services' widespread availability, functionality, and affordability. This reliance is lubricated by users' trust in the ability of cloud-based services to withstand logical and physical shocks of all kinds. The demonstrated capacity of the largest cloud providers to minimize disruptions to their services has been indispensable to enhancing this trust, making cloud services appealing and ubiquitous. Yet a serious breach of trust in cloud services (networks, computing, and storage) after an especially acute display of their vulnerability could be highly consequential and might be difficult and time-consuming to correct**.** Trust is an elusive and fragile commodity; therefore, we reason that the importance of mitigating residual risks should motivate providers to increase their efforts to maintain and enhance trust of all key stakeholders in this ecosystem.

Cascading effects: Although it is far simpler to treat these risks as isolated possibilities, we ought to consider the likelihood that some of the scenarios discussed above might have trickle-down effects that trigger additional scenarios. For example, a confidentiality breach might

impact availability, integrity, and ultimately trust. Similarly, as noted for the individual harms, the impact of disruptions on the CSPs themselves ("of the cloud") would be modest and short-lived, but their impact on at least some of their customers ("in the cloud") could be far more profound, depending on the nature of the customer's business, the level of their dependence on cloud services, and the level of their individual resilience to such shocks.

It is impossible to know which of these risks might rise to the level of "societally unacceptable" effects. However, we can note that disruptions to some sectors such as health care or financial services have the potential to intensify to a societally unacceptable level and beyond, through a combination of cloud disruptions and downstream effects suffered by those dependent on the cloud.

Possible Sources of Residual Cloud Risk

In the previous sections we have explored the specific manifestations of residual risk that this project has reviewed and assigned them a crude measure of probability and impact. We presently are in no position to assign a more precise probability or assessment of impact.

Here we consider possible sources of residual cloud risk. For heuristic purposes we have lumped these underlying factors into six interrelated categories. Many will require future in-depth consideration.

- **Concentration:** As discussed previously, the cloud represents a significant potential for cross-sectoral risk because so many consumers across sectors depend on it and because many functions (for example, credentials management) are with single-service providers.

- **Consequences of aggregation:** The cloud service market is concentrated in the hands of a handful of hyperscalers**.** This phenomenon is largely the outcome of the huge resource requirements needed to sustain sophisticated and reliable cloud services, which make it extremely difficult for providers to enter the market. While there are numerous benefits to having sophisticated and resourceful providers, concentrating services among just a few hyperscalers that serve the lion's share of the quickly evolving market does present some risks, especially if there could be common modes of failure among them.

- **Information gaps:** There are significant information gaps between various stakeholders that hamper accurate understanding of exposure to cloud related risks and of the level of resilience to the various types of potential cloud disruptions. These gaps extend to stakeholders' respective roles and responsibilities in mitigating cloud dependence risks and the available mitigation measures to enhance resilience. Especially when considering the important role of third parties in the cloud service supply chain and the reliance on utilities for service provisions.

- **Coordination in a competitive market:** Some mitigation proposals require participants to coordinate their actions. For example, CSPs would need to share lessons learned and mitigation strategies, or even work together in emergencies. Some collaboration takes place, but there are political, commercial, operational, and other barriers to broad-based coordination and cooperation. As a result, the plans for how the cloud system would absorb and recover from a major disruption (whether the result of an unexpected surge in demand or of degradation of the capability to supply services) are not fully understood. Several policy and technological innovations are currently being contemplated or rolled out that, if properly implemented, could help to address some (but certainly not all) of the critical coordination challenges.[15] Notably, several companies are working to address barriers to cooperation with technical solutions such as multi-cloud architectures, which would allow data to be portable and interoperable across public cloud providers.

- **Complacency:** The fact that we have thus far avoided a massive and extended disruption of the cloud is certainly inspiring. This success is no doubt a tribute to the major cloud providers' efforts to anticipate and plan for challenges and apply lessons from less severe incidents. However, this track record, if taken for granted, could induce complacency about resilience, especially among consumers.

Finally, the resiliency of cloud services is heavily dependent on providers' ability to shift resources and traffic to manage unanticipated peak demand and/or sudden shortfall of supply. Rigid restrictions on CSPs requiring them to divert or reroute traffic or data localization requirements might serve other national purposes but could undercut the providers' operational agility and ability to innovate solutions and end up impeding resilience. As discussed in the next section, regulatory action must be carefully calibrated to avoid unintended consequences.

The Policy Environment

The perceived current state of cloud security and risk concentration have made policymakers increasingly active in this realm. Governments across the globe are taking steps to encourage government agencies, critical infrastructure providers, and industries alike to enhance their cybersecurity measures—including in their use of the cloud—and are considering regulatory requirements for CSPs and their customers as well. The U.S. National Cybersecurity Strategy, for example, states that "The Administration will identify gaps in authorities to drive better cybersecurity practices in the cloud computing industry and for other essential third-party services, and work with industry, Congress and regulators to close them."[16] However, harmonizing such efforts to better drive security and resilience—while ensuring that benefits outweigh cost—is proving to be challenging.

Current State of Government Oversight

While much of the policy debate has focused on cloud security and mandating practices, there remain gaps, overlaps, and ambiguities in the responsibility of government agencies for driving cloud resilience. Recognizing and navigating those authorities is critical for an effective oversight regime. Given the scope and scale of CSP operations, there are multiple national regulatory authorities that potentially can set requirements on cloud security and resilience. Regulators in Europe, the United States, and the Asia-Pacific region are all taking steps in this direction. In Europe, the EU's Digital Operational Resilience Act sets operational resilience requirements for cloud and other information and communications technology providers in the financial sector. Asia does not have a centralized regulatory structure, so countries are pursuing country-specific data localization and competitive assurance rules.

Meanwhile, although cloud services increasingly underpin all sectors in the United States, there is no single digital authority that has clear authority to set requirements for the cloud ecosystem. Cloud companies are generally identified as critical infrastructure in the IT sector, where the Cybersecurity and Infrastructure Security Agency (CISA) serves as the agency tasked with risk management. Yet CISA is clear that its authorities in this space involve voluntary cooperation and not regulation.[17] Other government agencies play a role in creating guidance and policy, including:

- Through the Federal Risk and Authorization Management Program (FedRAMP), administered by the General Services Administration, the government uses its purchasing power to set meaningful security requirements on CSPs and other elements of the IT ecosystem as part of contract requirements. These controls and related oversight regime have spillover impacts in the practices that CSPs use for nonfederal customers. They are also significant drivers of compliance activity given the market share CSPs have with the Federal government.

- FedRAMP has proven an enduring driver of security practices and there is a movement to extend the model into many states via "StateRAMP" programs as well as consideration about the implications this would have for critical infrastructure sectors. At this point, however, the alignment between federal and state government requirements is not always consistent.

- The U.S. Department of Commerce's National Institute of Standards and Technology (NIST) develops core guidance on cloud services, including the NIST Cloud Computing Reference Architecture,[18] which is intended to provide a framework for use of cloud computing. It was developed in 2011 and has not been updated to reflect many of the changes in the current provision of cloud services. NIST also develops and propagates additional critical guidance including NIST 800-53 on Security and Privacy Controls for Information Systems and the NIST Cybersecurity Framework.

- The Biden administration is also implementing cloud security and software assurance requirements through Executive Order 14028 and follow-on activities assigned to government agencies, such as NIST and CISA, related to zero trust and information sharing, among other areas.

- Other critical infrastructure sectors, including communications, energy, financial, health, and public health, have regulators who are interested in ensuring continuity of service delivery for critical functions that are dependent on cloud computing, but whose authorities are not directly related to the resilience of the cloud.

- Market and consumer protection authorities are also diffuse. The Securities and Exchange Commission (SEC) operates on behalf of shareholders; [19] the Federal Trade Commission operates on behalf of consumers.[20] Both actively look at options for pushing cyber security requirements, including those protecting competition amongst service providers, on behalf of their constituents.

- The insurance sector is largely regulated at the state level and there is no clear federal preemption or backstop to support insurance markets in the face of concentration risk associated with the cloud.

This noncomprehensive list of the policy activity related to cloud resilience and the relevant government programs demonstrates the significant existing and potential regulatory burden put on the cloud ecosystem. The list does not include many aspects of data privacy, certification programs, and information collection and sharing requirements that are also relevant to the cloud industry; suffice it to say that the major industry players that enable the provision of cloud services already work closely with government agencies and face a potential regulatory whirlwind going forward; influencing that activity via education about existing resilience-building efforts is an important element of building an effective policy framework.

Ongoing Challenges

Given this situation, one key question is how to better harmonize government authorities and policies to drive cloud security and resilience and more effectively and efficiently reassure critical consumers about the availability and security of cloud services. Toward that end, in summer 2023 the Office of the National Cyber Director released a request for information in the Federal Register that included a query about "the costs to third-party service providers of conflicting, mutually exclusive, or inconsistent cybersecurity regulatory requirements that are passed on them through contracts with regulated customers."[21] The consensus from industry in response to the RFI is that regulatory harmonization is critical and long overdue. At the same time, however, the paths for achieving it aren't immediately obvious and as additional requirements are formulated there is scant evidence that they are being developed via processes that encourage and incentivize harmonization.

Harmonization in this case starts with the question of which standards CSPs will be held to, what regimes are in place to certify adherence to standards, the degree to which such regimes rely on attestations and audits, and the portability of demonstrated compliance. The more regulations with different requirements are created, the greater the compliance cost and burden on the regulated entities. This plays out for cloud services both in terms of the requirements placed on the service providers themselves as well as the steps they need to take to help their customers demonstrate compliance to regulation.

Any government intervention through regulatory means must ensure that the societal benefits—which can largely be conceptualized as residual risk reduction—outweigh the costs. These costs include direct and indirect monetary costs, as well as costs associated with limiting an open, competitive market that allows for innovation and the ability to compete on a global scale with CSPs emanating from China and other regions of the world. Lack of harmonization increases the costs and reduces the benefits.

One case where the costs imposed on the cloud ecosystem by regulation may exceed the benefits is the unintended consequence of rules inspired by the European Union's (EU's) General Data Protection Regulation (GDPR)—in particular, rules that require cloud providers to take all reasonable technical, legal, and organizational measures to avoid international transfers or governmental access to personal data held in the EU. These so-called "data localization" laws come with a price for resilience, especially CSPs' capacity to flexibly manage data loads to other locales to avoid disruptions. Data localization laws also contribute to the increase in the number and location of data centers, which add costs and complexity to CSPs operations, and some experts argue that these laws cause more harm than good.[22]

The debate over the impact of the European data localization requirement on the security and resilience of the cloud ecosystem is a useful cautionary case for U.S. regulators as they continue to implement the National Cybersecurity Strategy. A focus on gaps and overlaps in authority and on harmonization of requirements is useful, but whether additional requirements are needed remains an open question. We will return to this question in a later section.

Actions Needed to Ensure Resilience and Enhance Trust: A Cloud Resilience Framework

The massive investments and comprehensive security practices taken by cloud providers have averted catastrophic failures to date and led to generally higher expectations for resilience, but the cloud nevertheless faces potential residual risks as previously discussed. Therefore, to reassure all those now dependent on cloud services, actions are needed to ensure that the cloud ecosystem continues to remain resilient—reliable to withstand almost all acute challenges, and able to rebound with only minimal operational impact.

To address these residual risks, we are putting forward a set of interrelated and complementary recommendations. These recommendations offer a series of actions and best practices that will enhance resilience *of the cloud* and *in the cloud* over time. These actions include resilience testing and demonstrating, publicly, to those providing, consuming, and insuring cloud services, as well as to governmental agencies and regulators, that the resilience of and in the cloud is subjected to sustained testing and that shortfalls, if and when identified, are effectively addressed.

Toward these ends, this study proposes a Cloud Resilience Framework that lays out a foundation of policy commitments and follow-up actions to ensure resilience and enhance trust. Some of the proposed actions are already underway. Some steps can be taken in the near term while others will take longer to implement. The framework is intended to be as comprehensive as possible while being flexible to fit diverse business models and cultural differences, as well as to incorporate rapid technological evolution. The framework areas and a summary of actions needed are shown in Table 3.

Table 3: Overview of Cloud Resilience Framework

Framework area	What is it?	What needs to be done?
Foundational commitments	Commitments to advance cloud-related security and resilience.	Adopt Statement of Commitments developed during this project and commence implementing the obligations listed there.
Resilience "of" the cloud	Actions cloud providers can take to increase resilience of their cloud services.	In addition to making ongoing efforts more transparent, seek to (or continue to) build resilience culture into cloud development and operations and expand planning for contingencies and testing of mitigation plans.
Resilience "in" the cloud	Working with customers, insurers, and other stakeholders to increase the resilience of customers who rely on cloud services.	In addition to making ongoing efforts more transparent, work with a broad range of stakeholders to develop and promote a user-friendly Resilience Maturity Model and develop options to enhance stakeholders' resilience to cloud setbacks.
Exercises	Scenario-based exercises to validate contingency plans and test capabilities as well as identify best practices and lessons learned.	Expand internal programs to encompass more rigorous validation of design objectives and resilience in crisis scenarios and to include a broader range of stakeholders.

Each of the areas are discussed more fully below, along with specific near- and medium-term recommendations for each.

Foundational Commitments

The relationship between CSPs and customers is often described as a "shared fate" or "shared responsibility" model. The framing of this model maintains that the CSP provides the infrastructure and the broad security envelope while the customer is responsible for cyber hygiene locally. Depending on their sophistication level, customers may or may not understand the full implications of a large-scale, cascading event. Moreover, the expectations of other stakeholders—most notably governments but also critical service providers for the cloud such as energy, telecommunications, and water utilities—are not included in these agreements. Carnegie explored the policy issues in a previous report, noting the potential confusion about roles and that CSPs have broad responsibilities in the relationship especially as on-premises IT expertise is growing thinner, in particular in small- and medium-sized companies.[23] Recognizing this and proactively addressing it along the lines suggested in the Statement of Commitments would help avert some of the more severe consequences of misplaced expectations and misunderstandings; it could also head off the imposition of suboptimal regulations if governments step in to fill the breach.

In the near term, we recommend that our participants adopt the proposed Statement of Commitments to clearly calibrate expectations and provide another source of confidence to reassure stakeholders. This statement also provides a platform to subsequently invite those in the broader technology ecosystem to adopt these commitments. The Statement of Commitments is summarized in Table 4 below and included in full in Appendix 2.

Table 4: Summary of Proposed Statement of Commitments for Cloud Technology Providers

Topic	Elements
Situational awareness	• Cloud services add significant value including ease of use; enhanced cybersecurity; and sophistication of services and applications. • Global economy and society are increasingly cloud reliant. • CSPs invest significant resources to ensure reliable, secure, confidential, and dependable services. • Hostile actions, natural disasters, human errors, and technical failures create risk of cascading/enduring effects that could result in significant harms.
Roles, responsibilities, and stakeholders	• CSPs and customers share responsibility for the security, robustness, and resilience of cloud-based services. • CSPs are in a privileged position not only to address resilience *of* the cloud but also to inform and assist in enhancing resilience *in* the cloud. • Customers must maintain good cyber hygiene and configure their workloads in a secure and resilient manner. • Cloud service continuity is reliant on energy, telecommunications, and water utilities to sustain service. • Collaboration with insurers will be needed to help mitigate potential losses from cloud disruptions. • Governments need to spell out the threshold conditions for support (an insurer of last resort) should cloud services be severely disrupted.
Commitments	• Endeavor to prevent and mitigate the risk of cloud-based harms of compromise of availability, confidentiality, integrity, and trust. • Lead stakeholders to recognize, share, and address risks as they emerge. • Communicate about significant risks to consumers and other stakeholders in a timely way. • Ensure that policies, procedures, and practices are aligned with and support these principles.

This statement is intended to serve as a first step in developing a shared comprehensive commitment that invites all pertinent stakeholders in the cloud ecosystems to maintain and further enhance the robustness and resilience of cloud services.

In summary, Carnegie sought and incorporated feedback on the Statement of Commitments from the participants of the Carnegie Cloud Reassurance Project, and we believe it serves as a blueprint. We encourage other pertinent stakeholders to consider endorsing this statement and invite them to make corresponding and complementary pledges to enhance the resilience of cloud provisions in a manner consistent with their role and expertise.

Cloud Providers: Resilience of the Cloud

The second part of the Cloud Resilience Framework aims to ensure the resilience of cloud services and enhance trust in them. As previously discussed, CSPs already have several ongoing efforts to enhance the security and resilience of their services. However, most of these efforts appear focused more on security than on overall resilience; moreover, they are neither well known nor understood by other stakeholders.

Making these ongoing efforts more visible will be critical but will require some shifts in thinking. During our study, for example, we set out to model the scope and potential financial impact of potential outages using insurance modeling techniques. However, developing this impact model proved difficult because the cloud providers struggled to provide nonproprietary information about such things as the "blast radius" of a disruption—that is, which and how many users could be affected. Working through such obstacles will be challenging but an opportunity exists to increase understanding through increased sharing of information.

Looking forward, we recommend that CSPs work more with outside stakeholders to develop their respective risk profiles, including working with insurers to help develop better estimates of potential losses. Better information exchange is needed to ensure that the systems are resilient and to specify what actions CSPs would take in the face of contingencies. Such contingencies may have a serious bearing on CSP customers, other societal entities ("nonmarket participants") and nations writ large. These stakeholders need information about how CSPs will gear up should they experience an unexpected surge in demand or a shortfall in their capacity to supply full services, including their priorities for service restoration as well as "fail-over" or reserve capacity arrangements. Some of the actions, such as setting relative priorities for workloads in critical sectors like healthcare and energy during emergencies, may require guidance (and perhaps even consent) from governments. CSPs should identify those areas and seek such guidance where needed. They must also be open to adopting innovations that address these issues as well as to increased partnership with governmental entities.

Actions to improve CSPs' resilience are outlined in Table 5.

Table 5: Actions Needed to Improve Cloud Providers' Resilience

Recommendation	Near-term (< 2 years)	Medium-term (2+ years)
Make ongoing efforts more transparent and enhance trust by developing, periodically testing, and sharing contingency plans	• Build awareness among stakeholders of ongoing resilience efforts. • Work with insurers to identify information needed to evaluate and improve loss estimates. • Work with governments to identify prioritization plans should CSPs face unanticipated demands or other contingencies that diminish their capacity to supply services. • Expand and test load balancing/shedding strategies during contingencies. • Identify needed technological solutions that address reserve resource capacity, fail-over arrangements, and portability and interoperability of data across cloud services.	• Continue to work with insurers with a goal to expand coverage. • Periodically test and review the mitigation measures; adjust as necessary.

Customers: Resilience in the Cloud

The third part of the Cloud Resilience Framework pivots to focus on what customers can do to ensure their resilience, and how CSPs can assist in this. From the customer's perspective, moving IT to the cloud offers myriad benefits. No longer do they face large, recurring capital investments in hardware, software, and IT personnel; instead, they can largely rely on cloud services to provide that capacity and manage it. This allows them to focus less on IT, and more on their core operations. Nevertheless, customers should still use proper cyber hygiene, including encryption of their data, and be cognizant of supply chain risks, among other things, to ensure that their systems are safe and reliable. But many have not planned for situations in which their critical cloud services might be adversely affected. In other words, enhanced cloud resilience does not guarantee "resilience in the cloud."

Cloud customers often need assistance, however, to fully quantify their vulnerabilities, the resilience level of their deployments, and the capabilities available to help enhance their resilience. The CSPs are in the best position to provide such guidance and assistance. They can work closely with customers to help them identify and better understand the level of their resilience; they can also provide customers with options for configuring their systems and contract in ways that make their critical functions more resilient in the face of cloud disruptions. Participants see value in developing a common set of benchmarks—a "maturity model" for customer resilience.

Developing such a model will require considerable work: what constitutes a resilient system has so far received much less public attention than cybersecurity. However, the effort could pay huge dividends, both increasing resilience levels and enhancing trust. Developing such a model is in the public interest, as well as the interest of a range of stakeholders, and could help assuage concerns associated with the information asymmetries previously discussed. To wit, not only would customers better understand their own resilience, and options available to upgrade it, they should also receive more favorable insurance terms if they can demonstrate their higher levels of resilience. Therefore, once a maturity model is developed, it could help close a virtuous circle. A more transparent system would not only improve resilience but also reassure stakeholders, including governments, not least by generating a broad situational awareness of key sectors and societal resilience overall. In addition, the model would help the insurance market function better in allaying the financial risks.

Working to explicitly balance the privacy needs of customers—especially small to medium-sized customers who possess limited cloud expertise—with their resilience needs will be critical going forward. CSPs should address this challenge directly with customers to find ways to ensure adequate privacy and resilience. Customers may need to waive or otherwise revise some privacy restrictions to ensure they can access a full range of assistance from CSPs to achieve their desired levels of resilience in the cloud, especially for their critical functions.

These efforts could be both timely and influential, especially if they are synchronized with other ongoing efforts to develop technology standards and improve resilience. For example, in May 2023, the United States published the "United States Government National Standards Strategy for Critical and Emerging Technology," which pledged to invest in the development of standards that address risk, security, and resilience.[24] The actions outlined in this report could greatly inform the development of the maturity model. In addition, NIST is also making significant updates to its cybersecurity framework (last published in 2014). Most prominently, the scope of the framework has expanded from critical infrastructure such as hospitals and power plants to include cybersecurity for all organizations regardless of type or size. The NIST framework, which is and will remain voluntary, has been expanded to include a governance function to reflect the importance of cyber to management of risks.[25]

Building such a "Resilience Maturity Model" will require a stepwise approach. While it could start with unilateral efforts by individual CSPs along with their customers and insurers, over time the model should be harmonized across these efforts and would need to include the full range of stakeholders in addition to CSPs. While many paths could be taken, here are waypoints to guide the efforts:

- **Inventory** what CSPs already do to help customers build resilience, including any formal or informal resources or programs.
- **Consult** customers and other stakeholders to determine how they view resilience and what they consider important.
- **Identify** commonalities and best practices across CSPs.

- **Collaborate** with customers and insurers to identify the key elements of resilience that should be fundamental to the model.
- **Pilot** different approaches and models.
- **Build** a generic Resilience Maturity Model with levels based on agreed-upon standards.
- **Deploy** the model for use in assessing resilience across technology sectors.
- **Develop** ways to share composite information and lessons learned levels with insurers and other stakeholders, including anonymizing the data to the extent needed.

Our specific recommendation and implementation actions are shown in Table 6.

Table 6: Actions Needed to Improve Resilience in the Cloud

Recommendation	Near term (< 2 years)	Medium Term (2+ years)
Develop a Resilience Maturity Model (RMM) to allow customers to better assess their resilience level as well as provide visibility to other stakeholders, including insurers.	• Catalogue, organize, and advertise existing efforts. • Ensure that efforts balance privacy concerns with resilience needs. • Work together (CSPs/customers/insurers) to develop a framework for an RMM.	• Develop RMM standards to describe resilience levels. • Harmonize across cloud providers • Use resilience maturity level to explain resilience level to third parties, including developing ways to anonymize information if necessary.

Exercises and Stress Testing

Complex systems are exceptionally difficult to comprehensively and reliably assess in an ongoing way. Our CSP participants undergo numerous audits and other customer-specific validations to ensure that they have proper processes and procedures to withstand security challenges and support the services they offer. Although such audits and internal tests are necessary and helpful, they often focus on security or pieces of the larger system rather than resilience more broadly. In general, these processes do not test how the broader system will perform under stress, nor do they assess how a cloud disruption could affect swaths of customers, sectors, and societies writ large. Although results are shared with customers, concerns about data privacy and protection of proprietary business information understandably make CSPs reticent to share the detailed results of these tests with outside stakeholders. A middle ground must be found here for trust to be enhanced.

As part of its ongoing financial-sector cloud project, the U.S. Treasury Department conducted a tabletop exercise in April 2022 that brought together CSPs, financial institutions, and various government agencies in a simulated outage of an IaaS provider. The

exercise revealed that more work needs to be done to examine the potential impact of such an outage on the financial sector. The Treasury Department intends to conduct additional tabletop exercises and has formed an interagency steering group.[26]

Stress testing put in place after the financial crash of 2008 provides an analog for assessing complex systems that are integral to commerce and society. This testing was intended to assess whether "too-big-to-fail" financial institutions could maintain sufficient reserve capital amid shocks. Stress tests include a standard set of scenarios that financial institutions might experience, such as a natural disaster, an unexpected conflict, or a sudden surge in unemployment combined with a plunge in stocks and real estate.[27] Such tests are performed periodically, usually semiannually, and have strict reporting requirements. Like large financial institutions, the hyperscale cloud providers may have crossed the "too-big-to-fail" threshold, and the ecosystem could benefit from a neutral, well understood exercise, or set of exercises, that include strong protections for CSPs'—and their customers'—intellectual property.

From the outset of our study, our participants agreed that the idea of such an exercise had merit, both to help them test their own resilience and identify shortfalls if any, as well as to reassure the various stakeholders. We explored this throughout our study; in fact, we had planned to conduct a tabletop exercise to walk through a scenario-based disruptive situation during the project. However, myriad issues precluded us from doing so; this remains a high priority for future work.

Nonetheless, the merit of the underlying idea endured, and participants discussed scenarios that could form the basis for such exercises. Four plausible scenarios were identified that had the potential for residual risk and would thus be appropriate to test through an exercise. These included:

- **Code vulnerability** to malevolent intent (such as a supply chain attack), unintended error, or unanticipated failure.
- **Compromise of credentials** that gives unauthorized parties privileged access.
- **Availability loss** from malevolent or unintended misconfiguration.
- **Connectivity problems** triggered by loss of telecommunications, power outage, or architecture issues.

To address operational sensitivities and avoid potential duplication, we are recommending a "crawl-walk-run" approach in developing an exercise and stress-test program. This approach would ultimately build to scenarios designed to test critical parts of the system (especially in circumstances that deprive CSPs of the ability to provide their full services to all their customers). Exercises would include independent observers and participants from other stakeholder organizations, including governments. These exercises should be designed to assess the plans, procedures, and capabilities that make it possible to prevent, protect, mitigate, respond to, or recover from a threat or incident. Recommended actions are shown in Table 7 below.

Table 7: Recommended Actions to Expand CSP Exercise Programs

Recommendation	Near term (< 2 years)	Medium Term (2+ years)
Expand exercise programs to test contingency plans against increasingly stressful scenarios, evaluate capabilities, and identify best practices and lessons learned.	• Inventory individual CSP exercise programs and identify opportunities for expansion. • Identify and include outside observers who can add both expertise and independence • Conduct progressive series of exercises with increasing complexity over time. • Test mitigation measures to manage stress, such as load shedding arrangements and load balancing.	• Continue to expand into broader exercises with more stakeholders and more stressful scenarios to test resilience more fully. • Implement system-wide fixes to major weaknesses revealed by the stress tests. • Identify needed capacity for large-scale scenarios and explore options for attaining reserve capacity. • Develop methods to share outcomes and lessons learned.

Policy Paths Forward

Although our primary focus has been on what the private sector might do to ensure resilience and enhance trust, continued policy and regulatory actions are expected. These can have welcome benefits; they can also cause inefficiencies if not well designed and implemented. Our focus on policy actions is related to the things that the government can do to enable the proposed Cloud Resilience Framework.

Given the challenges of harmonizing regulation and the negative impact that a lack of harmony has on efficient resilience measures by CSPs, the path forward will need to be carefully plotted and coordinated. Our work suggests that several underlying factors should guide policy actions to ensure that new requirements consider resilience and risk reduction as well as security:

- **Reinforcing the idea that cloud computing is critical across many sectors, not just IT, and enabling cross-sector coordination.** When it comes to critical infrastructure, law and policy as well as international norms prioritize security *and* resilience through government-industry collaboration, national preparedness, risk mitigation, and national defense. Cloud computing, and the underlying function of data management, are critical infrastructure and are increasingly recognized as essential to business and government continuity. At the same time, it needs to be acknowledged that other critical infrastructure—particularly telecommunications, energy, and water—is interdependent with cloud computing. Expectations that cloud services will be assured for critical infrastructure should be balanced with the expectation that CSPs have priority access to that infrastructure to support

continuity efforts. Testing the resilience of interdependencies should be a national preparedness priority enforced by governments in a streamlined way. There also should be clear expectations that cloud services must not be targets of nation-state attacks to cause mass availability outages and that if they are so attacked, there will be commensurate responses from national governments. This will have the benefit of protecting critical infrastructure and critical functions.

- **Improving transparency of risk information to support shared risk management.** Our efforts to fully understand residual risks associated with cloud computing were limited by the availability and sensitivity of information. One of the roles that governments play effectively is mandating the provision of information to support more informed decisionmaking by risk managers. Much of the federal focus thus far has been focused on incident reporting—including through the Cyber Incident Reporting for Critical Infrastructure Act (CIRCIA) and SEC-mandated requirements—but risk information is more than a descriptive analysis of incidents that occur. It also needs to enable risk assessment. While much progress has been made in information sharing and operational collaboration, currently the ability to do data-driven assessments of residual risk across the cloud ecosystem is difficult because of data collection and availability issues. One area of focus should be enhancing information sharing about threat activity, significant near-miss incidents, known and observed vulnerabilities, and effective controls, all of which will support holistic risk assessments. Creating the ability to aggregate anonymized information through a mechanism such as a Bureau of Cyber Statistics would be a positive step; so too would be continued public-private efforts to enhance information-sharing and enhanced collaboration among insurance providers and CSPs. Any such efforts, however, must be accompanied by strong liability, intellectual property, and other safe harbor provisions for the public sector participants.

  Not only do CSPs need to learn from each other's experience, their customers need to have access to risk information so they can adjust their security and resilience postures as needed to protect sensitive information and ensure continuity of operations. It is not the responsibility of CSPs to manage their customers' risk, but creating mechanisms for information-sharing that allow customers to take on that responsibility is an imperative—particularly for organizations that provide critical functions that communities and the economy depend on. The recent announcement that the U.S. Department of Homeland Security's Cyber Safety Review Board will conduct a review of the malicious targeting of cloud computing environments presents an opportunity for recommendations on how to enhance transparency of risk information.[28]

- **Developing tools that make it easier to value both security and resilience by default.** Deploying secure and resilient cloud services is a shared responsibility between CSPs and customers. Both benefit from the push being made for security by default. CISA defines "secure by default" as products "that are secure to use out of the box, with little to no configuration changes and are available at no additional cost."[29] To that, we add the imperative "resilience by default," which means that cloud services are available to customers with a clear option for how they can be deployed and utilized to ensure resilience of service provisions. For a movement like secure/resilient by default to succeed, industry and government should partner to create a definition of what these terms mean for cloud services at the technical level. Customers of cloud services should have the expectation that their systems will be configured to meet their standards of risk acceptance, and mechanisms like Resilience Maturity Models should be in place to allow customers to have a conversation about the right level of security and resilience as part of deployment. It is especially critical that work to develop Resilience Maturity Models be a multistakeholder effort, involve all relevant parties, and not stifle competition.

- **Leaning toward innovation and market-based solutions rather than restrictive regulatory solutions.** Governments can help facilitate better transparency and resilience in and of the cloud through support for technological and commercial solutions to address the scenarios and concerns surrounding cloud services rather than through regulatory mandates. One current example is the U.S. Department of Defense's embrace of—and requirement for—data portability and interoperability across clouds (also known as multi-cloud technology) in the $10 billion Joint Warfighting Cloud Capability contract.

- **Ensuring a functioning (re)insurance market.** Our review has found that it remains unclear whether insurance is positioned to address the concentration of residual risk that remains in the current cloud ecosystem. This leaves governments as insurers of last resort on a *de facto* basis as opposed to by design. This needs to change and work needs to be done to address the findings of the Treasury Department's 2022 request for information regarding the establishment of a federal cyber insurance backstop.[30] Creating a backstop would proactively define maximum risk exposure for the insurance market and provide greater certainty as to when the federal government would step in to manage tail risk should a catastrophic incident occur. That has the resilience benefit of contributing to continuity of the economy and enhancing the current insurance and reinsurance markets.

- **Maintaining a collaborative model.** The policy discussion around cloud resilience needs to include both policymakers and industry participants, and participants need to prioritize establishing and maintaining a relationship based on trust. Governments, key providers, and users of cloud services need to effectively share information and collaborate on risk mitigation to be prepared for response and recovery in the event of a major cloud disruption—especially one caused by a

foreign adversary aiming to exploit gaps and sow confusion while causing significant harm. Government has a role in catalyzing cloud resilience and has a responsibility to protect communities and citizens from digital harm. That role can be partially fulfilled by regulation, but it must also involve streamlining authorities, taking steps to harmonize requirements, and utilizing mandates to promote better risk understanding and address market failures. Cloud services are a shared benefit to societies and their availability and resilience needs to be recognized as a shared goal. The Statement of Commitments proposed in Appendix 2 is a step in that direction, but living up to those commitments will be crucial.

Conclusions and Future Considerations

This study brought together some of the largest cloud providers and reinsurers in the world, along with cloud-savvy enterprises and multiple independent experts with rich industry, technology, and government experience, all of whom were invited by Carnegie to assist with the project. The aim was to assess the risks associated with the growth and consequent concentration of cloud services against the backdrop of rising policymaker anxiety—anxiety that will only intensify given the explosion of cloud-dependent AI.

The bottom line is that while many beneficial effects stem from widespread use of cloud services and applications, the risks involved cannot be eliminated. Attempts to eliminate these risks would be inefficient at best and would likely both reduce innovation and increase costs. Nonetheless, the residual risks should be made transparent and addressed. This study offers a roadmap for doing this.

Because their innovations and market successes have made them such an integral part of our lives, cloud providers going forward will need to think beyond their customers to broader society. Such a change can occur only with a sustained commitment from top management to bring together the technical, operational, business, and policy elements of their huge organizations. But the potential rewards are significant. For one, developing such norms can be a credible alternative to regulatory interventions. Moreover, it would represent a good faith effort that would be visible to the governments entrusted with protecting their citizens. The time to begin this shift is now. We should not wait for a catastrophe to occur and exact its cost, because the policymaker response would likely be outsized and serve to diminish the benefits of cloud services.

All stakeholders in the ecosystem have critical roles going forward. CSPs must take more responsibility for enhancing the resilience of the cloud. At the same time, customers must make a parallel shift in their attitude and recognize that they play a key role in their own resilience. This will require C-suite appreciation for their own responsibilities as well as a

clear-eyed and continuous evaluation of the risks they face and what they can do to mitigate them. In addition, insurers want to help customers manage their cyber risks, but the market is somewhat stalled amid recent losses and the fear of tail risks. Insurers need better information to offer better coverage and will need to work with governments to design an effective backstop for catastrophic losses.

Significantly, many of the recommendations in this report can be implemented simply by increasing openness or sharing information about what is already occurring at the CSPs. Others will require additional efforts, or further study, such as expanding exercise programs and establishing standards to underpin a Resilience Maturity Model. However, some issues will require careful monitoring and likely additional study. For instance, the growing centrality of cloud services amid the labyrinth of critical infrastructure is insufficiently understood. Cloud providers can withstand a limited power outage, for example, but an extended or widespread disruption could ripple through cloud infrastructures and potentially cascade into other sectors. Moreover, the growing fusion between cloud and telecommunications services also has considerable implications and must be carefully considered.

Finally, the rapid emergence of AI is intensifying the spotlight on the potential risks of technology. AI and cloud services are already interwoven and difficult to disentangle or even distinguish as large-language models often rely on massive cloud-based datasets. But one thing is clear: as society and commerce become increasingly reliant on an AI-enmeshed cloud, the resilience of that cloud will be crucial.

Appendix 1: Implications of Cloud Dependence for Insurance

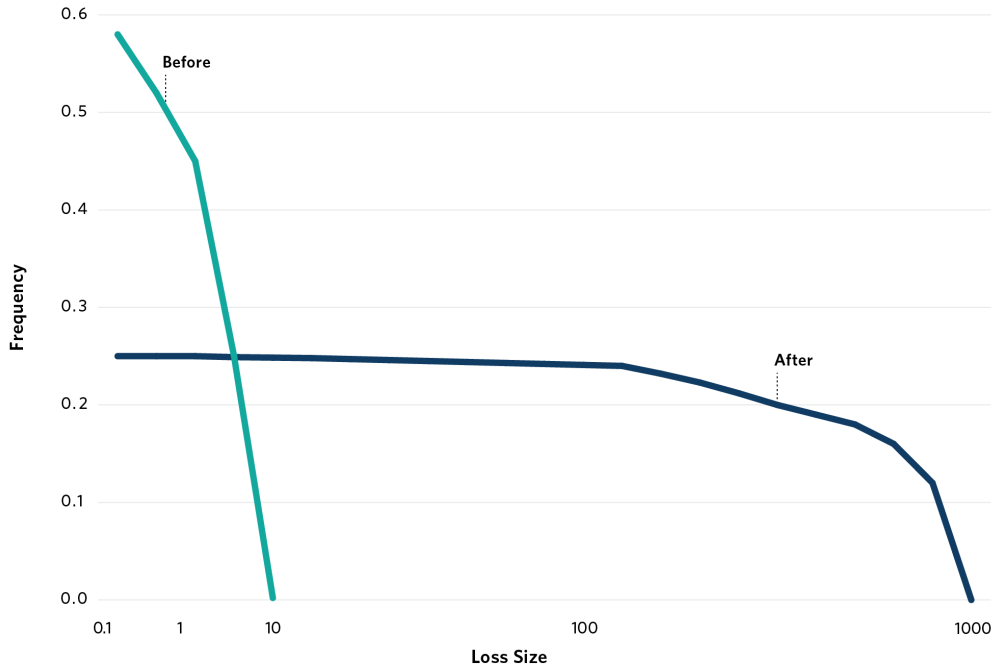
Increasing reliance on the cloud presents an apparent paradox: namely, that risks are likely less overall, but the "tail risk" associated with a contagion event that impacts a broad swath of cloud-based customers is increased. One can visualize these assumptions using graphical representations. The Loss Frequency Curve shown in Figure 4, for example, indicates how often losses occur before and after the introduction of cloud services and how costly those losses are. The basic assumption imagines a hypothetical society that is totally without cloud services ("before") and compares it to the situation where 100 percent of potential cloud-based activity takes place on the cloud ("after"), all other things being the same. We also assume in the first graph that our society is free of any accumulation risk in the "before" state, and that the only accumulation of risk in the "after" state is due to the introduction of cloud services. In other words, the largest loss before the introduction of cloud services corresponds to a total loss of just the single largest risk in our society. Although this is unrealistic, it helps to illustrate the possible impacts of a transition to widespread cloud adoption.

The Loss Frequency Curve is an exceedance curve, i.e. the vertical axis describes the frequency of events that result in the loss values on the horizontal axis. It thus takes into consideration not only the relative size of the losses but also the total number of losses in a year. For the sake of representation, we have chosen a logarithmic scale on the horizontal axis.

In broad terms, complete reliance on cloud services should mean fewer small losses (attritional and medium sized) due to the assumed higher cyber security/resilience posture achieved through cloud adoption. However, due to the aggregation/accumulation

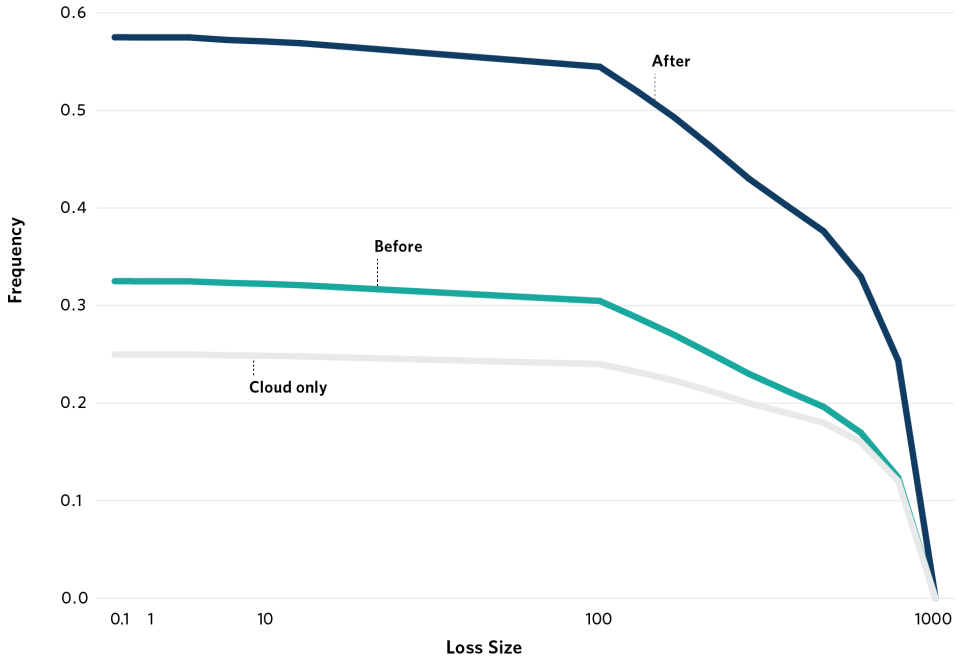
effects—because more people will be affected by any event—the largest possible losses are now much larger than they were in the non-cloud society.[31] In this representation we have not made any attempt to reflect on the "total risk" of our society. Loss Frequency Curves allow us to show (and calculate) such an annual total risk, which would be represented by the surface under each curve.

Figure 4. Comparison of Notional Loss Frequency Curves Before and After Cloud Adoption (No Accumulated Risk)



Source: SwissRe and MunichRe

Figure 5. Comparison of Notional Loss Frequency Curves Before and After Cloud Adoption (With Previous Accumulated Risk)



Source: SwissRe and MunichRe

In order to be able to represent data closer to reality both for the frequencies and for the loss sizes, and consequently also for the "total risk," one would need many more insights into the assumed improved security that cloud adoption brings with respect to small and medium-sized losses and into the systemic risk introduced by the very adoption of the cloud.

Figure 5 represents a somewhat more realistic situation in our hypothetical "society." Even prior to the introduction of the cloud, there existed some accumulation risk (for instance zero-day vulnerabilities in widely used pieces of software) and this is reflected in the blue line. The introduction of the cloud thus represents an additional accumulation of risk. Here, for the sake of simplification, we will assume that

The largest possible loss due to the introduction of the cloud is the same size as the largest possible loss due to the already existing accumulation.

The existing accumulation risk (in blue below) remains untouched by cloud adoption (in other words, computers in our hypothetical society are still performing the same local tasks and have not been replaced by the cloud).

This graph should be considered a schematic: it is not possible for us to represent the true change in the shape and position of the curves due to the real, existing adoption of cloud services. Clearly the new accumulation effect will increase the expected frequency of large

losses, but it is unclear if the accompanying reductions in attritional and mediums losses will over-compensate and reduce the total risk or undercompensate and result in a net increase in risk. This uncertainty is due in part to our limited insight into the improved security presumably brought about by cloud adoption as well as our limited ability to understand and quantify the replacement effect of existing computer use reducing the accumulated risk that existed before cloud adoption. Furthermore, the shape transformation will be dependent on the specific types of loss analyzed (for example, nonavailability of services, data breaches, and compromise of confidential information).

As already mentioned in this report, stress tests of the cloud services and transparent communication of those outcomes would be a first step toward a quantitative representation of effective transformation of the changes brought about by widespread cloud adoption.

Appendix 2: Statement of Commitments on Cloud Services

Seeking to enhance trust in the robustness and resilience of cloud services, the participants in the Carnegie Cloud Reassurance Project offer the following Statement of Commitments. This statement is intended to serve as a first step in developing a shared comprehensive commitment that invites all pertinent stakeholders in the cloud ecosystems to maintain and further enhance the robustness and resilience of cloud services.

Cloud Benefits and Potential for Harms

- Cloud services and the overall cloud ecosystem add significant value to the people and organizations that use them, in both the private and public sectors. Such value includes the increased ease of computing services, improved cybersecurity, the benefits of scale (including increased access to products from developers), and sophistication in applications.

- Governments, organizations, and individuals are becoming increasingly reliant on cloud services for their everyday business, making the global economy and society more dependent on the robustness and resiliency of cloud services.

- Cloud service providers (CSPs) recognize the importance that customers and governments place on the confidentiality, availability, and integrity of their cloud services, as well as on their ability to trust these services. They are also conscious of the need to be able to track and learn from setbacks to cloud services that occur through actual incidents as well as to anticipate those through tests and simulations.

- Thus, CSPs expend considerable effort and resources to ensure reliable, secure, confidential, resilient, and dependable services.

- Notwithstanding these efforts, there is a residual risk that cloud-related incidents could be more enduring and widespread (especially if triggered by sophisticated hostile actions or large-scale natural disasters), producing cascading and/or enduring effects on individuals, institutions, and societies writ large.

- The ubiquity of cloud services and the critical dependence that entities and societies have on their reliability mean that such events, while improbable, could produce considerable harm by triggering cascading and/or enduring effects.

- Such harms may be triggered by events in the logical or physical realms that result in compromises of confidentiality,[a] availability,[b] integrity,[c] and not least trust,[d] and may be further complicated by accountability challenges.[e] Such impacts might originate in causes that lie partially or wholly beyond the control of cloud services providers.

- Such rare severe events are neither fully predictable nor preventable. But prudent planning and contingency plans by all the participants in the cloud ecosystem can make such events less likely and less severe even though they cannot be entirely eliminated.

Roles, Responsibilities, and Stakeholders

- CSPs and their customers share responsibility for the security, robustness, and resiliency of cloud-based services. Having a clear understanding of the respective roles and responsibilities is critical to maintaining and enhancing their respective resilience to disruptions.

- The CSP's role is to provide underlying cloud infrastructure and services that are robust and resilient; to enhance the risk awareness of cloud customers; and to offer cloud customers capabilities to deploy more advanced features and resilient service options.

- Bearing in mind that CSPs by default have neither access to their customers' data nor cognizance of the critical functions customers assign to the cloud, it is the customers' role to configure and execute their workloads in a manner that meets their security and resilience requirements.

---

a Encroachment on individual privacy, commercially sensitive information.
b Adverse impact on the ability of customers to access the service they have contracted for whenever they need to do so.
c Susceptibility to improper accidental or malicious manipulation of assets and services entrusted to the cloud.
d Loss of confidence in the cloud service overall or in the reliability of some of its features or services.
e Diminished ability to detect, locate, and trace the origin of adverse impacts.

- Continuity of cloud services is heavily reliant on availability of electricity, water, and communications services. CSPs maintain a reserve capacity to withstand temporary shortfalls in these services but longer duration disruptions of these services stress cloud resilience.

- Insurers are important facilitators of risk mitigation and risk channeling services, and as such could have an important role to play both in mitigating most risks associated with cloud disruptions and in enhancing the capacity of their customers to recover from such incidents. Realizing this role requires intimate collaboration between the stakeholders to enhance the understanding of cloud services and dependence therein, as well as to bound the residual systemic risks inherent in cloud services.

- Governments typically actively support early warning from, response to, and recovery from natural disasters and assume the lead role in forestalling and responding to external attacks on their citizens, enterprises (especially those enterprises performing or supporting critical national functions), and territory. In extreme cases governments even serve as "insurers of last resort." They can be expected to play similar roles in the event of similar cyber induced scenarios.

Participants' Commitments

Noting that their legal duties are confined to compliance with both law and customer agreements, CSPs and any other cloud stakeholders joining this pledge commit to the following principle:

"**First, endeavor to prevent and mitigate the risk of harm. In doing so, consider and prioritize those risks that could result in systemic and other socially unacceptable harms, inter alia through cascading, and enduring effects on critical services.**"

CSPs are hardly the sole relevant players capable of mitigating such harms. They nevertheless recognize that they are both in a unique position to tackle risks that could degrade their services, and in a privileged position to continuously identify the potential for extensive harm stemming from the lack of security and resilience of their infrastructure. CSPs are thus in a position to take steps to understand both systemic and other socially unacceptable cloud risks, as well as to mitigate against them. They can do so by sensitizing and encouraging actors in their supply chains as well as other stakeholders to mitigate such risks. To fulfill this principle, CSPs commit to making a sustained effort toward the following:

- **Recognize risks:** Regularly assess the key risks (as described above) of significant harm in and stemming from uses of technology within the ecosystem, especially experienced by customers critically dependent on cloud services; and make sustained and concerted efforts to prevent and mitigate those risks that fall within their jurisdiction;

- **Articulate risks:** Cautiously and appropriately share with consumers of cloud services and other pertinent stakeholders the known risks both within the cloud ecosystem and stemming from uses of technologies within this ecosystem;

- **Address risks:** Undertake to redress in a timely fashion all critical vulnerabilities exposed in their cloud services while also providing commercially reasonable assistance to participants in the ecosystem to mitigate against risks within the cloud ecosystem, of cloud services, but also empower these participants to enhance their overall level of digital resiliency (which likely will include opportunities for education, technical assistance, and information sharing); and

- **Update** all consumers of cloud services and other pertinent stakeholders in a timely manner on developments that could end up causing them serious harm while prioritizing communication with those consumers that provide critical services.

Recognizing that these commitments are not to be taken lightly, CSPs and others endorsing this Statement of Commitments undertake to have in place policies, procedures, and practices and allocate commensurate resources that would allow them to appreciate, articulate, and address cloud-based risks that accompany their innumerable benefits. CSPs further commit to regularly assess and as warranted, update such policies and procedures for effectiveness. They also acknowledge the importance of periodically engaging their customers and other defined stakeholders to update them on their practices in this realm and solicit feedback on ways to upgrade them.

The participants of the Cloud Reassurance Project encourage all other pertinent stakeholders to endorse this statement and invite them to make corresponding and complementary pledges to enhance the resilience of cloud provisions in a manner consistent with their role and expertise.

About the Authors

John H. Pendleton serves as a nonresident scholar in the Technology and International Affairs Program. Prior to joining Carnegie, he served almost thirty-five years at the U.S. Government Accountability Office (GAO).

Ariel (Eli) Levite was the principal deputy director general for policy at the Israeli Atomic Energy Commission from 2002 to 2007.

Bob Kolasky is a nonresident scholar in the Technology and International Affairs Program and is senior vice president for critical infrastructure at Exiger, where he focuses on developing cutting-edge third-party risk management solutions for the critical infrastructure community.

Acknowledgements

The authors wish to acknowledge extensive contributions by project participants, as well as a number of independent experts.

Participating private-sector organizations included Amazon Web Services, Axio, Microsoft Azure, Google Web Services, Munich Re, Swiss Re, and VMware. The participant organizations provided a range of technical and policy experts who were generous with their time and deep knowledge. Particular thanks to Karen Barth and Chris Cornilie, who were instrumental in compiling the section on ongoing resilience efforts, as well as Dale Gonzales and Eric Durand, who advised on our assessment of risks and authored the technical appendix on insurability challenges, respectively.

Carnegie scholars included Robert Kolasky, who authored the policy chapters and lent his expertise across a range of issues. Peter Armstrong was instrumental in leading our discussions of technical risks and how they might be mitigated. John Pendleton envisioned the Cloud Resilience Framework and served as the overall editor of the report. Finally, Ariel (Eli) Levite authored the section on residual risks and has driven the Cloud Reassurance Project from its beginning. The project enjoyed outstanding legal support from Gare Smith and Chris Hart from Foley Hoag LLC who brought extensive corporate social responsibility expertise as well as ensured that the project did not violate antitrust provisions, among many other contributions. The authors sought consensus among project participants but relied on our scholars and other experts to ensure independence and retained editorial control for the final report.

Steering Group

Nathan Bird
Rachelle Celebrezze
Edna Conway
Matt Gaschel
Chris Hart
Scott Kannry
Robert Kolasky
Ariel (Eli) Levite
George Perkovich
Jürgen Reinhart
Jordana Siegel
Gare Smith
Bobbie Stempfley
David Tennenhouse
Phil Venables
Stephan von Watzdorf

Other Contributors

Anitha Abrahim
Peter Armstrong
Merritt Baer
Karen Barth
Nick Beecroft
James Bono
Christiana Briggs
Stephan Brünner
Chris Cornilie
Amanda Craig
Debashis Das
Eric Durand
Chris Finan
Dale Gonzales
Hila Hanif
Tara Knapp
Martin Kreuzer
Marc Nussbaumer
John Pendleton
Kevin Reifsteck
Jurgen Reinhart
Mark Ryland
Val Sandoval
David Simpson
David White
Hudi Zack

Finally, the authors wish to thank staff at the Carnegie Endowment for International Peace, which has supported the multiyear Cloud Reassurance Project from its beginning. The publications staff, including Alana Brase, Jocelyn Soly, and Amanda Branom were patient and helpful throughout. Isabella Furth edited and proofread the paper, offering many suggestions that improved the report. Aiysha Kirmani Zafar and Reshad Amini helped us manage our grants and kept our books straight. Emma Landi and Arthur Nelson were instrumental in assisting the team throughout the project, providing considerable support for our range of project participants. George Perkovich advised throughout and offered many insightful suggestions on the approach and final report. Finally, a project of this ambitious scope could not have been completed without the unwavering support of Carnegie President Mariano-Florentino (Tino) Cuéllar.

Funding for the project was provided by the participating organizations. Special thanks to Omidyar Network, who provided generous support which made it possible to complete the project.

Notes

1 Ariel (Eli) Levite and John Pendleton, "Cloud Reassurance Project: Interim Report," Carnegie Endowment for International Peace, June 12, 2023, https://carnegieendowment.org/2023/06/12/cloud-reassurance-project-interim-report-pub-89927.

2 Felix Richter, "Amazon Maintains Lead in the Cloud Market," Statista, August 8, 2023, https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/.

3 "Gartner Says More Than Half of Enterprise IT Spending in Key Market Segments Will Shift to the Cloud by 2025," press release, Gartner, February 9, 2022, https://www.gartner.com/en/newsroom/press-releases/2022-02-09-gartner-says-more-than-half-of-enterprise-it-spending.

4 "Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach $679 Billion in 2024," press release, Gartner, November 13, 2023, https://www.gartner.com/en/newsroom/press-releases/11-13-2023-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-reach-679-billion-in-20240.

5 Levite and Pendleton, "Cloud Reassurance Project: Interim Report."

6 We use the term disruption to mean an event that affects not only the availability of services but also their confidentiality and integrity.

7 Levite and Pendleton, "Cloud Reassurance Project: Interim Report."

8 Levite and Pendleton, "Cloud Reassurance Project: Interim Report."

9 Ryan Smith, "Staggering 90% of Cyber Risk Remains Uninsured," Insurance Business, November 8, 2022, https://www.insurancebusinessmag.com/us/news/cyber/staggering-90-of-cyber-risk-remains-uninsured-426713.

10 Levite and Pendleton, "Cloud Reassurance Project: Interim Report."

11 "Cyber Risk Accumulation: Fully Tackling the Insurability Challenge," The Geneva Association, November 2023, Cyber Risk Accumulation: Fully tackling the insurability challenge (genevaassociation.org)

12 "Potential Federal Insurance Response to Catastrophic Cyber Incidents," U.S. Department of the Treasury, September 29, 2022, https://www.federalregister.gov/documents/2022/09/29/2022-21133/potential-federal-insurance-response-to-catastrophic-cyber-incidents.

13 Thus far we are familiar with only one survey that examines this level of resilience, including the discrepancy between big versus small and medium financial institutions in the United States. See "The Financial Services Sector's Adoption of Cloud Services," U.S. Department of the Treasury, February 8, 2023, https://home.treasury.gov/system/files/136/Treasury-Cloud-Report.pdf.

14 Zack Whittaker, "Danish Cloud Host Says Customers 'Lost All Data' After Ransomware Attack," TechCrunch, August 23, 2023, https://techcrunch.com/2023/08/23/cloudnordic-azero-cloud-host-ransomware.

15 In 2022, the U.S. Cybersecurity and Infrastructure Security Agency (CISA) launched a public-private collaborative group called the Joint Cyber Defense Collective (JCDC) to bring cybersecurity experts together from industry and government. JCDC membership includes CSPs, security companies, and government agencies.

16 "National Cybersecurity Strategy," The White House, March 2023, https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Cybersecurity-Strategy-2023.pdf.

17 "FTC Seeks Comment on Business Practices of Cloud Computing Providers that Could Impact Competition and Data Security," press release, Federal Trade Commission, March 22, 2023, https://www.ftc.gov/news-events/news/press-releases/2023/03/ftc-seeks-comment-business-practices-cloud-computing-providers-could-impact-competition-data.

18 Fang Liu et al., "NIST Cloud Computing Reference Architecture," National Institute of Standards and Technology, September 2011, https://doi.org/10.6028/NIST.SP.500-292.

19 "SEC Adopts Rules on Cybersecurity Risk Management, Strategy, Governance, and Incident Disclosure by Public Companies," press release, U.S. Securities and Exchange Commission, July 26, 2023, https://www.sec.gov/news/press-release/2023-139.

20 "FTC Seeks Comment on Business Practices."

21 "Request for Information on Cyber Regulatory Harmonization," Executive Office of the President, Office of the National Cyber Director, July 19, 2023, https://www.whitehouse.gov/wp-content/uploads/2023/07/ONCD-Reg-Harm-RFI-Final-July-19.2023.pdf.

22 Emily Wu, "Sovereignty and Data Localization," Belfer Center for Science and International Affairs, Harvard Kennedy School, July 2021, https://www.belfercenter.org/sites/default/files/2021-07/SovereigntyLocalization.pdf.

23 Ariel (Eli) Levite and Gaurav Kalwani, "Cloud Governance Challenges: A Survey of Policy and Regulatory Issues," Carnegie Endowment for International Peace, November 9, 2020, https://carnegieendowment.org/2020/11/09/cloud-governance-challenges-survey-of-policy-and-regulatory-issues-pub-83124.

24 "United States Government National Standards Strategy for Critical and Emerging Technology," The White House, May 2023, https://www.whitehouse.gov/wp-content/uploads/2023/05/US-Gov-National-Standards-Strategy-2023.pdf.

25 "NIST Drafts Major Update to Its Widely Used Cybersecurity Framework," news release, National Institute of Standards and Technology, August 8, 2023, https://www.nist.gov/news-events/news/2023/08/nist-drafts-major-update-its-widely-used-cybersecurity-framework.

26 "U.S. Department of the Treasury Kicks Off Public-Private Executive Steering Group to Address Cloud Report Recommendations," press release, U.S. Department of the Treasury, May 25, 2023, https://home.treasury.gov/news/press-releases/jy1503.

27 Troy Segal, "What Is a Bank Stress Test? How It Works, Benefits, and Criticism," Investopedia, October 28, 2021, https://www.investopedia.com/terms/b/bank-stress-test.asp.

28 Department of Homeland Security's Cyber Safety Review Board to Conduct Review on Cloud Security," press release, U.S. Department of Homeland Security, August 11, 2023, https://www.dhs.gov/news/2023/08/11/department-homeland-securitys-cyber-safety-review-board-conduct-review-cloud.

29 "Secure by Design," Cybersecurity and Infrastructure Security Agency (CISA), https://www.cisa.gov/securebydesign.

30 "Summary of Comments on Request for Comment: Federal Insurance Response to Catastrophic Cyber Incidents," U.S. Department of the Treasury, Federal Insurance Office, March 29, 2023, https://home.treasury.gov/system/files/311/2023-03-27%20Presentation%20Summary%20of%20Cat%20Cyber%20RFI%20Responses.pdf.

31 For the sake of simplification and of the graphical representation we assume that the accumulation risk represents a 100-fold increase compared to the largest single loss. Effectively this factor should be roughly equal to the largest number of cloud users impacted by the same cloud incident and hence would in reality be much larger than 100. Furthermore, again for the sake of representation, we have used here frequencies which are not representative of real portfolios. The frequency of small losses would be much larger than 1, showing that there are multiple incidents in a year in such portfolios.

Carnegie Endowment for International Peace

The Carnegie Endowment for International Peace is a unique global network of policy research centers around the world. Our mission, dating back more than a century, is to advance peace through analysis and development of fresh policy ideas and direct engagement and collaboration with decisionmakers in government, business, and civil society. Working together, our centers bring the inestimable benefit of multiple national viewpoints to bilateral, regional, and global issues.

Technology and International Affairs

The Carnegie Technology and International Affairs Program (TIA) helps governments and industries reduce large-scale international risks of new technologies and related services. Recognizing that commercial actors control many of the most germane technologies, TIA identifies best practices and incentives that can motivate industry stakeholders to pursue growth by enhancing rather than undermining international relations.

CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

CarnegieEndowment.org