Perspective

Sovereign cloud and the 5G network: an assessment

February 2023

Bence Szeidl, Joseph Attwood and Caroline Chappell

Contents

List of figures

This perspective was commissioned by 5GDNA. Usage is subject to the terms and conditions in our copyright notice. Analysys Mason does not endorse any of the vendor's products or services.

1. Executive summary

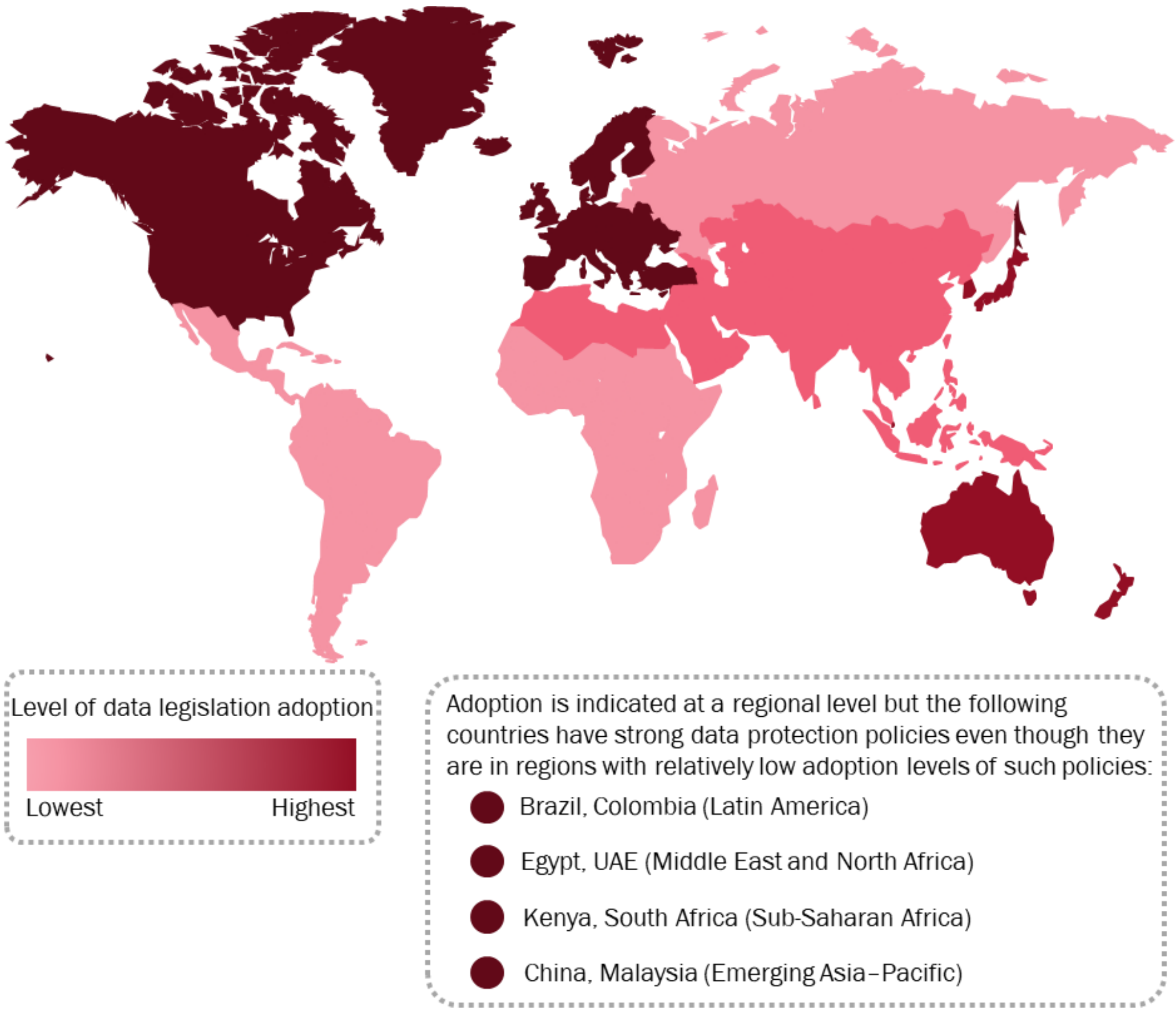
1.1 Global interest in data sovereignty and privacy is rapidly increasing

Most countries have data protection regulations in place or are drafting data protection policies. The EU's General Data Protection Regulation (GDPR) and its predecessor, the Data Protection Directive (DPD), have been instrumental in guiding discussions on data protection and privacy worldwide. Both pieces of legislation have had an impact on attitudes to data protection, prompting other countries and regions to adopt similar frameworks to safeguard the information of their residents.

The publication of the GDPR was a watershed moment in the history of data protection legislation, showcasing the possibility of enforcing comprehensive laws across a large geographical area. This is leading to further supranational efforts to protect sensitive information. However, although countries generally agree on the principles of data protection, their legislation varies in scope and detail and the introduction of supranational legislation remains a difficult and expensive process. As a result, many companies err on the side of caution and introduce restrictive data policies that are not well-aligned with those of their neighbours and trading partners.

Renewed interest in data sovereignty is being driven by the expansion of the global digital economy and the ease with which digital mechanisms generate, process and use customer data, including across borders. Data sovereignty, which Analysys Mason defines as data that is subject to the regulations of the country of origin, is fundamental for providing adequate levels of data protection to digital economy participants. Data sovereignty is associated with the ideas of data localisation and data residency, which set out what, how and if at all data can leave the jurisdiction where it was produced. Laws based on these concepts can make it difficult for enterprises to operate internationally due to different regulations they need to meet in each country.

Figure 1.1: Adoption of data protection legislation, worldwide, 2023



Adoption is indicated at a regional level but the following countries have strong data protection policies even though they are in regions with relatively low adoption levels of such policies:

- ● Brazil, Colombia (Latin America)
- ● Egypt, UAE (Middle East and North Africa)
- ● Kenya, South Africa (Sub-Saharan Africa)
- ● China, Malaysia (Emerging Asia–Pacific)

Source: Analysys Mason

1.2 Public clouds pose multiple challenges for enterprises that want to continue to comply with data sovereignty rules

Public clouds are emerging as the engine rooms of the digital economy, home to the vast amount of data needed to power it. Public clouds have gained popularity over the past decade due to their flexibility, scale and the cost benefits they provide, but they may also pose a threat to data sovereignty. Because public cloud providers (PCPs) are physically present in a limited number of countries, enterprises face challenges associated with moving workloads and their associated data across borders or between regions to access PCP cloud infrastructure, if these enterprises want to continue to comply with increasingly stringent and fragmented national and supranational privacy laws. Enterprises are also concerned about a growing trend for extraterritorial legislation that allows a government to exercise its country’s laws beyond its borders. This gives such countries the power to access the data of citizens in other jurisdictions in some circumstances. The largest PCPs that hold and manage data from around the world are all based in the USA and are, therefore, subject to the US CLOUD Act. This is a threat to data privacy and may result in conflicts of interest between the USA and other nations’ sovereignty laws.

Despite significant efforts on the part of PCPs to make their infrastructure secure, occasional vulnerabilities can still expose sensitive customer information. Such leaks can cause reputational and financial damage to the enterprises that own the data and are subject to data protection legislation, while the PCPs are somewhat shielded from the negative publicity. The same applies to public cloud outages, which are rare, but require enterprises to design their systems for relatively low levels of availability (99.9% uptime). Public cloud availability is not guaranteed by service level agreements (SLAs), so enterprises have no redress if the infrastructure fails and they are responsible for the resilience of their own systems. PCPs typically address resilience by migrating workloads and data to unaffected availability zones in case of failure. If these are out of region, this may violate data protection legislation.

Availability and resilience are not sovereignty considerations per se, but regulated industries, such as telecoms, need to take them into account when considering whether or not to use third-party infrastructure, because any factor that may cause a communications service provider (CSP) to fail to comply with regulation of any kind is a business red flag. A particular concern voiced by CSPs is the mismatch between the level of availability that CSPs and PCPs provide. CSPs are required to provide 99.999% uptime on their mission-critical networks and SLAs backed by financial penalties for failure. PCPs are not required to provide SLAs and are unable to match CSP levels of uptime.

1.3 Evaluating the viability of the public cloud for 5G core workloads

CSPs recognise the benefits that the public cloud can provide and, like most other enterprises, have begun to move IT workloads to it. Advanced CSPs are now understandably evaluating the suitability of public clouds to support network workloads. 5G networks are built to run on the cloud, and as CSPs start to deploy 5G networks, they are evaluating whether to run their 5G cores on public cloud infrastructure. However, CSP networks are highly regulated because they provide critical national infrastructure and CSPs are therefore encountering three main challenges when trying to migrate their networks to the public cloud.

- CSPs need to comply with all data protection and privacy regulations that are effective in their region of operation. Depending on the country and its data protection regime, the current operational model and geographical presence of PCPs may not be adequate enough to meet a CSP’s sovereignty requirements.

- CSPs cannot afford data leaks nor the risk of being subject to extraterritorial jurisdiction due to the sensitive customer information they handle. These risk factors may be exacerbated if network cloud infrastructure is outsourced to PCPs.

- CSPs are subject to governmental mandates regarding the availability and resilience of their networks. Many CSPs would face regulatory barriers if they tried to use a PCP that only has a single data centre in a country or region, even if that data centre offers multiple availability zones (AZs), and they struggle with the PCPs’ ongoing lack of support for carrier-grade SLAs.

Many CSPs are considering deploying the cloud-native 5G standalone (SA) core initially for enterprise use cases and not to support their highly regulated macro networks that carry consumer traffic. For some enterprise use cases, the public cloud may provide a viable platform for the 5G SA core. However, enterprises are still subject to national and/or regional data protection legislation and, in many cases, they are considering private 5G networks to support mission-critical use cases associated with operational transformation. Enterprises in key sectors, such as finance, healthcare and manufacturing, will be as concerned about the compliance, availability and security aspects of their private networks as CSPs are about their own networks.

For many reasons, CSPs need to understand the potential issues that will affect the deployment of a 5G core in the public cloud, whether that core is destined to support consumer mobile broadband at scale or valuable enterprise use cases and network slices. This paper considers the complexities involved in meeting data protection regulation in countries across the world and the cost of meeting additional availability and security requirements that the public cloud imposes if CSPs are to avoid reputational damage. It reveals the questions that CSPs should ask themselves when considering whether to build an on-premises cloud to run their 5G mobile core or to buy infrastructure as a service from a PCP to host this critical network function.

2. Governments are increasingly demanding data sovereignty guarantees to protect citizens

2.1 GDPR is waking up the world to the value of protecting data and privacy as a national asset and human right

The introduction of new, and the strengthening of existing, data protection legislation is gaining momentum worldwide. The European Union's GDPR, which came into force in 2018, has had a major influence on attitudes towards data protection and is prompting renewed interest in promoting it. Data protection is not a new area of concern and legislation that safeguards individual privacy has existed for decades. The EU's predecessor to GDPR, the DPD, was enacted in 1995 and was subsequently adopted as the basis for other nations' data protection legislation, including Singapore's and Malaysia's Personal Data Protection Acts introduced in 2012 and 2013, respectively. Japan, however, was one of the first countries in Asia to introduce data protection legislation, introducing its Act on the Protection of Personal Information in 2003, which it revised in 2015.

The introduction of GDPR has changed the global discourse around data protection for several reasons. In general, the regulation of personal data prior to GDPR was limited in scope and lacked a strong framework for enforcing compliance. GDPR is different because it is supranational in nature and has proven that it is possible to define a broad piece of legislation that can, and will, be enforced by the countries that are subject to its jurisdiction. Since 2018, countries in all parts of the world, including Australia, Brazil, Chile, China and Egypt, have been inspired to introduce new data privacy regulation or significantly reworked their existing rules. Today over 70% of all countries have some sort of data protection law in place and another 10% have legislation on the way.[1] For example, the Commission of the African Union started developing a supranational Data Policy Framework in 2021. Figure 2.1 provides examples of legislation in different countries and regions of the world.

Figure 2.1: Notable national and supranational data protection policies

Region	Country/area	Policy name	Date of approval or introduction
North America	Canada	The Personal Information Protection and Electronic Documents Act (PIPEDA)	2000
	USA (federal)	US Privacy Act of 1974	1974
	USA (state of California)	California Consumer Privacy Act (CCPA)	2018
	USA (state of Virginia)	Virginia Consumer Data Protection Act (VCDPA)	2021

[1] Data from the United Nations Conference on Trade and Development, 4Q 2021.

Region	Country/area	Policy name	Date of approval or introduction
	USA (state of Colorado)	Colorado Privacy Act (ColoPA)	2021
Latin America	Argentina	Argentina Personal Data Protection Act (PDPA)	2000
	Brazil	Lei Geral de Proteção de Dados Pessoais (LGPD)	2018
	Chile	Law 19.628 (Chilean Data Protection Law)	1999
	Mexico	Ley General de Protección de Datos Personales	2009
	Peru	Law No. 29733 on the Protection of Personal Data	2011
Europe	EU	The General Data Protection Regulation (GDPR)	2016
Western Europe	Switzerland	Federal Act on Data Protection (FADP)	1992
	Turkey	Law on Protection of Personal Data No. 6698 (DPL)	2016
	UK	United Kingdom General Data Protection Regulation (UK GDPR)	2018
Developed Asia–Pacific	Australia (federal)	Privacy Act 1988	1988
	Australia (Capital Territory)	Information Privacy Act	2014
	Australia (Northern Territory)	Information Act	2002
	Australia (New South Wales)	Privacy and Personal Information Protection Act	1998
	Australia (Queensland)	Information Privacy Act	2009
	Australia (Tasmania)	Personal Information Protection Act (PIPA)	2011
	Japan	The Act on the Protection of Personal Information (APPI)	2003
	Singapore	The Personal Data Protection Act (PDPA)	2012
	South Korea	Personal Information Protection Act (PIPA)	2011
	China	PRC Cybersecurity Law	2017
	India	Digital Personal Data Protection Act (draft)	2022
	Indonesia	Personal Data Protection Bill (PDP Bill)	2022
	Malaysia	Personal Data Protection Act (DPDA)	2010
The Middle East and North Africa	Egypt	Law No. 151 (Law on the Protection of Personal Data)	2020
	Israel	Protection of Privacy Law, 5741-1981	1981
	Qatar	Law No. (13) of 2016 Concerning Personal Data Protection	2016
	Saudi Arabia	The Personal Data Protection Law (PDPL)	2022
	UAE	Federal Decree Law No. 45, Protection of Personal Data (DPL)	2021
Sub-Saharan Africa	Ghana	Data Protection Act (Act 843)	2012
	Kenya	Data Protection Act No. 24 (DPA)	2019
	Nigeria	Nigerian Data Protection Regulation (NDPR)	2019

Source: Analysys Mason

2.2 Countries define data protection in similar ways but attitudes to legislation vary across the world

Countries largely agree on the guiding principles for data protection, which focus on limiting the collection and storage of personal data, providing transparency around the type of information that is being is collected and why, and requirements for the security of and consumer control over personal data. Data protection legislation can also define what constitutes the safe transfer of data and the penalties that will occur if companies do not follow the legislation, such as prosecution and fines. GDPR is considered to be the world benchmark for each of these principles. However, there is significant variation in the scope and granularity in the way that different countries implement these principles. Most countries that are currently aligning their privacy laws with GDPR do not follow its example fully. They have a strong focus on defining sensitive information and how data can be transferred between organisations but are weaker on enforcement. For example, New Zealand's Privacy Act and Nigeria's Data Protection Regulation (NDPR) build on the GDPR but are more lenient when it comes to providing directives on fines for non-compliance. South-East Asian countries have also recognised the importance of transferring data securely across borders and have attempted to define a common standard for protecting data across member states: the ASEAN Framework on Personal Data Protection (2016). Since the ASEAN countries do not have an overarching supranational authority with legislative power, as the EU does, they rely on individual members for enforcement, resulting in varying levels of data protection across the region. Different regions of the world have fundamentally different attitudes to data privacy. The state of California and the European Union both have comprehensive data privacy laws, but they differ over the default right of companies to process data. In the EU, consumers must specifically grant companies permission to use their information, while under the California Consumer Privacy Act (CCPA) the processing of personal data is permitted by default and consumers must explicitly opt out. The definition of what constitutes sensitive data can vary as can the rules and requirements affecting how, when and where data can be transferred. The introduction of supranational legislation is a lengthy and resource-intensive process. Since countries want to protect their citizens in the meantime, there is evidence that many countries introduce more restrictive data protection regimes with stricter controls on what information, if any, can leave their country, than supranational regulation requires. Privacy is considered a human right in developed regions, including the EU, Canada and South Korea, although developing countries, such as Chile, are also introducing this concept into their constitutions. As a result, countries are under increasing pressure to take responsibility for the safety of their citizens' data and thus are tightening regulation around data leaving the country

2.3 Participation in the digital economy is a key driver for new data protection legislation

It is no coincidence that national and supranational interest in data protection is intensifying as the global digital economy is expanding. The digital economy poses an increased threat to customer data and is a key driver for new data privacy legislation for the following two reasons.

- **The amount of personal data that can be collected about customers is increasing.** Digital products and services target the collection of personal data and the digital nature of data storage and connectivity has made it easier for stores of personal data to be attacked. Breaches of the First American Financial Corporation (2019), Marriott International (2018) and Equifax (2017) resulted in the exposure of over 1.5 billion data points combined, including highly confidential information such as credit card details, Social Security numbers and addresses. These events have shaken the confidence of consumers in the digital storage of their data by third parties, especially as loss of that data can seriously affect their personal lives.

- **The global nature of the digital economy means that personal data can easily be collected and sent across borders.** Digital products and services are easy to sell across national borders, which can result in conflicts not only over where consumer data is processed but which countries have access to it under which legislation.

There is growing consensus around legislative principles and the example of GDPR to follow, but country-specific regulatory nuances make navigating the global data protection environment difficult. This situation will get worse as more countries not only introduce legislation around data privacy but also increase their propensity to prosecute breaches of their laws.

2.4 Data sovereignty is a key pillar of data protection

Enterprises and public bodies must ensure that they can process and store data in a way that is compliant with the local data protection legislation they are subject to. This is giving rise to the concepts of data sovereignty and sovereign cloud.

Analysys Mason defines sovereign data as data that is subject to the legislative and governance requirements of the specific region or jurisdiction in which the data has been produced. The concept of data sovereignty is associated with the relatively recent idea of data residency or data localisation. Data localisation requires data about residents of a country to be collected, processed and stored in that country under that country's jurisdiction. Data subject to data residency cannot be transferred out of the country's jurisdiction, in some cases, without the consent of the data owner. In other cases, where countries have a stricter interpretation of data localisation, personal data cannot leave the jurisdiction at all.

Data localisation encourages local approaches to data management in order to reduce the risks associated with non-compliance, including fines and threats to reputation and consumer confidence. All three risks can severely damage a company's economic performance in a national market due to the amount of media attention they can attract in a data-conscious age. A sovereign cloud that operates within countries or supranational jurisdictions under their specific data residency/localisation laws can provide a solution to the requirement for local shared data processing and storage capacity. Sovereign clouds are perceived to have key role to play in keeping the data and privacy of residents safe.

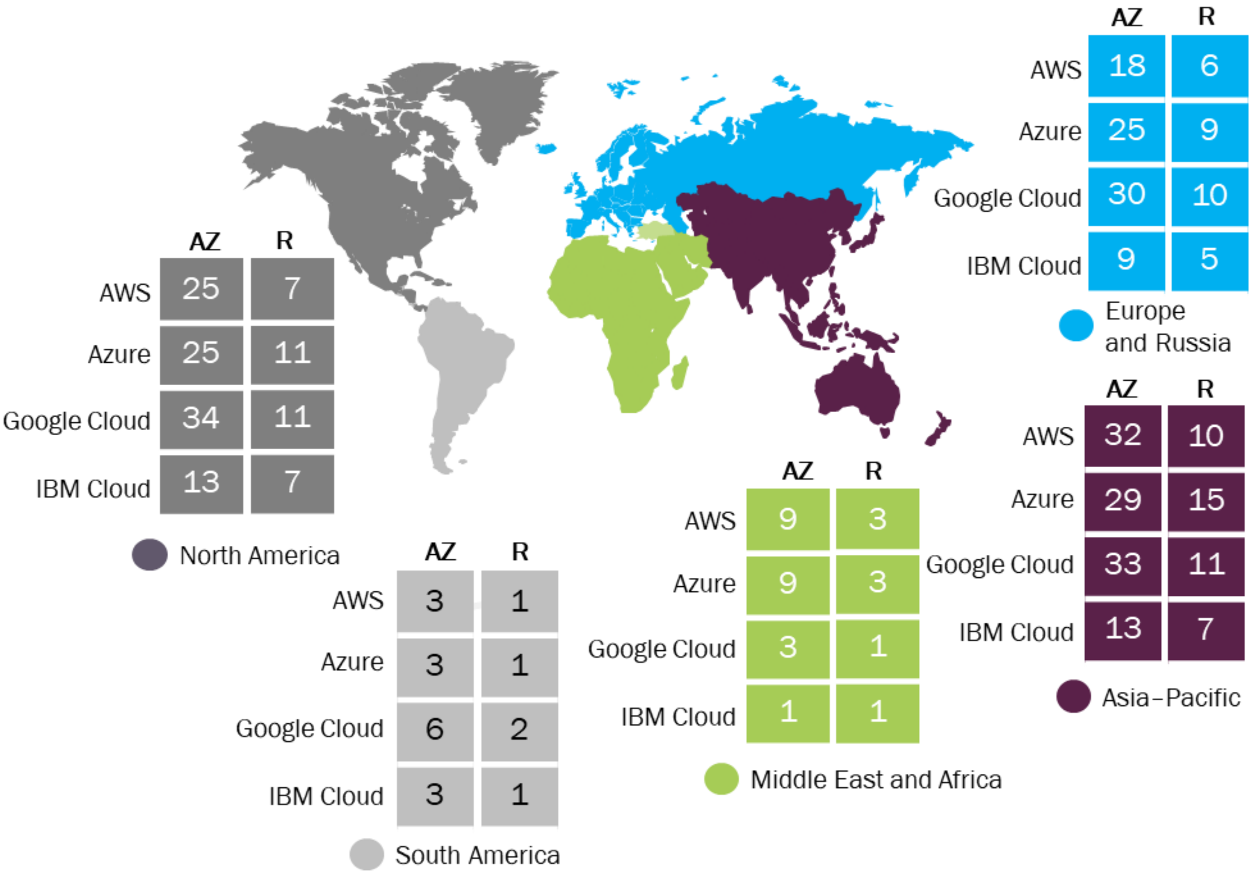
3. Public cloud poses challenges for data sovereignty

3.1 Public cloud threats to data sovereignty

The concept of data localisation has emerged contemporaneously with the growing use of the public cloud. Enterprises and public bodies increasingly collect, process and store data in public clouds which can provide more flexibility around cost and scale than an organisation's own data centre(s). However, the public cloud can pose challenges for organisations that need to comply with local data protection legislation. These threats manifest themselves in the following ways:

Public cloud footprints do not provide comprehensive coverage of data protection jurisdictions. The largest PCPs have a global reach, but this is achieved by efficiently moving data across their infrastructures to maximise its utilisation and by keeping multiple copies of the same information in different physical locations. In reality, PCPs have highly centralised and geographically limited data centre footprints distributed unevenly across the globe (Figure 3.1). As we have seen, 70% of the world has some sort of data protection legislation in place but the top three PCPs only cover less than 30% of these countries with their data centres. This means that as the data sovereignty requirements of different regions becomes increasingly fragmented and detailed, these PCPs will face challenges around migrating workloads and their associated data across borders or between regions if they want to remain compliant with national and supranational privacy laws. Leading PCPs will also find it increasingly onerous to accommodate the sovereign requirements of many countries as they expand their footprints.

Figure 3.1: Geographical distribution of availability zones (AZ) and regions (R) for AWS, Google Cloud, IBM and Microsoft Azure, 4Q 2022[2]



● North America

	AZ	R
AWS	25	7
Azure	25	11
Google Cloud	34	11
IBM Cloud	13	7

● South America

	AZ	R
AWS	3	1
Azure	3	1
Google Cloud	6	2
IBM Cloud	3	1

● Middle East and Africa

	AZ	R
AWS	9	3
Azure	9	3
Google Cloud	3	1
IBM Cloud	1	1

● Europe and Russia

	AZ	R
AWS	18	6
Azure	25	9
Google Cloud	30	10
IBM Cloud	9	5

● Asia–Pacific

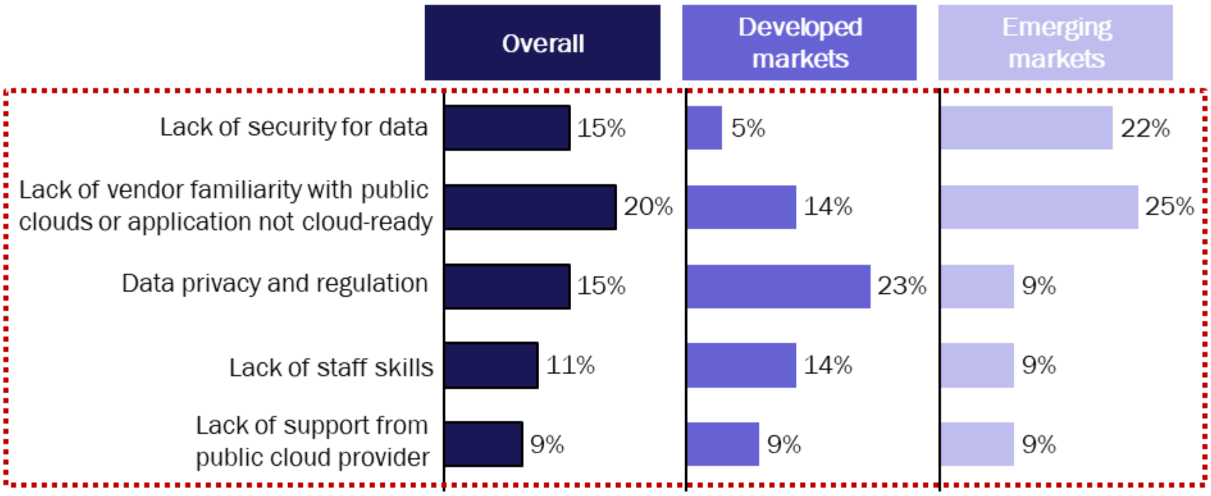
	AZ	R
AWS	32	10
Azure	29	15
Google Cloud	33	11
IBM Cloud	13	7

Source: Hyperscalers' websites, press releases, Analysys Mason

Public clouds have multiple security issues. Although public cloud providers invest heavily in the security of their processes, tools and infrastructure, security remains a top concern for both CSPs and enterprises, frequently cited as a barrier to public cloud adoption in surveys (Figure 3.2). Turkish Airlines subsidiary *Pegasus* and healthcare platform *Doctors Me* both suffered data leaks from their AWS S3 buckets in 2022. This resulted in the exposure of multiple terabytes of data, including highly confidential information such as personal details of flight crews and medical records of patients. Although such security breaches are rare, many companies want sensitive operational data to remain on their physical premises as the ultimate guarantee of the sovereignty and security of that data.

[2] Availability zones can contain one or more data centres and never share data centres for redundancy reasons. Multiple regions can exist within the same country.

Figure 3.2: CSPs' main barriers to public cloud adoption, 2021[3]



Source: Analysys Mason

A more insidious threat to data security is coming from governments around the world and an emerging slew of extraterritorial legislation. Alarmed at the threat that highly mobile digital data stored in unfriendly jurisdictions might pose to national security, governments are proposing, or have passed, legislation that grant extraterritorial jurisdictive power to their countries, enabling them to exercise their country laws outside their borders. The US CLOUD Act (previously Stored Communications Act), Australia's Assistance and Access Bill and the EU's E-evidence Package are examples. These Acts state that companies that are headquartered in their jurisdictions may be required to expose personal information in the event that it is required for a criminal investigation. PCPs are at the epicentre of the controversy over such extraterritorial legislation for two reasons. They own the 'master' cryptographic keys at the root of their infrastructure which can potentially unlock any third-party encrypted data flowing across their clouds. Since PCPs hold so much of the world's data and the keys to unlock it, this makes them a target for governments. The majority of the world's leading PCPs are US-based and are therefore subject to the US Cloud Act. In the *Microsoft v United States (2018)* case, the US Federal Bureau of Investigation (FBI) tried to exercise its power to access information held in one of Microsoft's overseas data centres in 2013. Microsoft refused to provide the information and the case was dismissed by the US. Supreme Court due to the concurrent introduction of the CLOUD Act and a revamped request for data issued under it. However, the case highlights the potential for conflicts of interest between the sovereign laws of one country and the extraterritorial jurisdiction of another.

3.2 Public cloud availability poses a further challenge to regulated enterprises operating across borders

Enterprises often have inflated confidence in the availability of public clouds. Even the largest PCPs are not immune to service outages. In December 2021, an AWS outage affected applications including Amazon Music, Amazon Prime and Netflix. This particular outage affected enterprises and customers along the US East Coast and lasted for 8 hours. Google Cloud experienced a similar outage in March 2022 when a configuration error affected the availability of applications such as Spotify and Discord worldwide for over 2 hours. In June 2022, a power failure disrupted access to Microsoft Azure resources hosted in the Eastern US region for 12 hours. Outages as significant as these rarely occur more than once a year but smaller outages are more common. PCPs put an onus on their enterprise customers to anticipate outages and architect their applications accordingly.

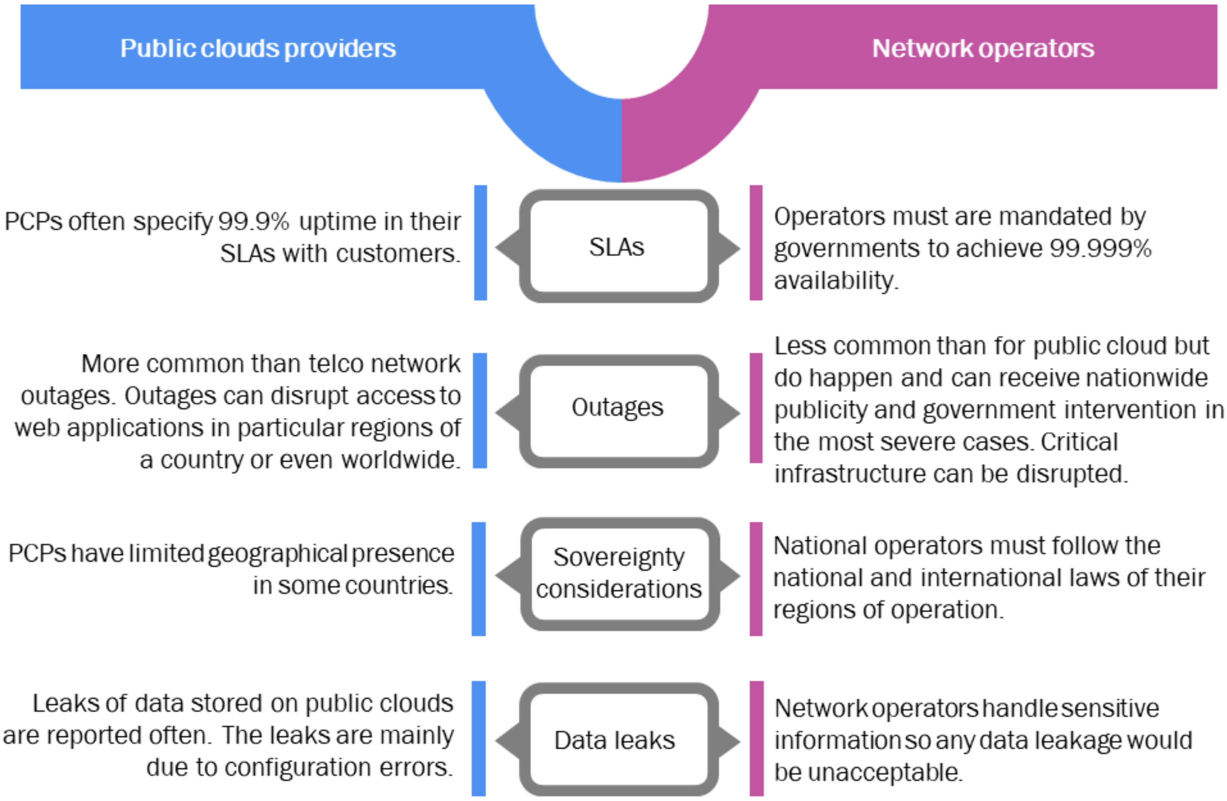
[3] Question: "What are/were the top three barriers to public cloud adoption?"; *n* = 60.

Enterprises affected by public cloud outages may experience a significant loss in revenue during outage periods as well as reputational damage, since they are blamed by their end-users for the lack of service whilst PCPs are somewhat shielded from negative publicity.

Although the availability and resilience of public clouds are not sovereignty issues in themselves, they illustrate a further risk for enterprises of using third-party infrastructure. That risk is particularly high for regulated companies, such as CSPs. CSPs typically have to ensure 99.999% availability, or less than 5.5 minutes of downtime a year on their networks, as outages affect customers and critical national infrastructure. CSPs' obligations are recorded in strict SLAs, setting out their guarantees regarding the quality and availability of their services, as well as the consequences of failing to deliver on those guarantees, In contrast, PCPs typically provide 99.9%, or 'three nines', of availability and do not provide strict SLAs. CSPs frequently cite misalignment between PCPs' interpretation of availability and their own. This means that PCPs are unable to meet CSPs' SLA requirements, a key area of conflict between the parties.

Disaster recovery provision as a result of power outages, floods, lightning strikes or other events is related to availability considerations. Public cloud infrastructure is designed to enable the rapid recreation of workloads and data in a data centre in an unaffected region in case of disaster. However, since in many countries, leading PCPs only have a single data centre, disaster mitigation practices may raise sovereignty issues if workloads and data need to be transferred beyond national borders, For example, customer data that enables calls to be routed correctly is sensitive, so moving it between neighbouring jurisdictions with different data legislations may not be allowed. National CSPs are often required to be served by at least two in-country data centres by the regulatory regimes under which they operate.

Figure 3.3: Comparison of PCPs and network CSPs



Public clouds providers		Network operators
PCPs often specify 99.9% uptime in their SLAs with customers.	SLAs	Operators must are mandated by governments to achieve 99.999% availability.
More common than telco network outages. Outages can disrupt access to web applications in particular regions of a country or even worldwide.	Outages	Less common than for public cloud but do happen and can receive nationwide publicity and government intervention in the most severe cases. Critical infrastructure can be disrupted.
PCPs have limited geographical presence in some countries.	Sovereignty considerations	National operators must follow the national and international laws of their regions of operation.
Leaks of data stored on public clouds are reported often. The leaks are mainly due to configuration errors.	Data leaks	Network operators handle sensitive information so any data leakage would be unacceptable.

Source: Analysys Mason

4. Assessing the benefits and risks of running a 5G core in the public cloud

4.1 General implications of public cloud for CSP networks

Like most other enterprises, CSPs are already moving IT workloads to public clouds. CSPs are attracted by the benefits of public cloud, including perceived cost-benefits as a result of its on-demand consumption model, managed operations and maintenance by the cloud provider which frees operations staff for more strategic activities and ready access to commodity infrastructure. Public clouds obviate the need for CSPs to invest in building their own private clouds by sourcing and integrating complex, fast-moving technologies.

As a result of their experience with IT workloads, advanced CSPs are understandably evaluating the suitability of public clouds to support network workloads. However, since CSP networks provide critical national infrastructure and are therefore highly regulated, CSPs are encountering three main challenges when trying to migrate their networks to the public cloud.

- **Compliance with national and regional data protection legislation**. Networks carry sensitive personal data and CSPs must follow the national and supranational laws that apply in their regions of operation. Depending on the country and its data protection regime, the current operational model and geographical presence of PCPs may not be adequate enough to meet a CSP's sovereignty requirements.

- **Provision of the highest levels of security**. CSPs process, store and transport large amounts of sensitive user data over networks that are considered to be critical national infrastructure, which means that any kind of data leakage, whether of customer data or data that could compromise the network itself, is unacceptable for them. Nor can they afford for this data to be moved to or accessed by an extraterritorial jurisdiction, either inadvertently or to recover from a disaster. A Tier 1 CSP pointed out that it is large enough to negotiate that PCPs surrender their digital encryption keys to any cloud instance in which the CSP hosts workloads. However, the CSP noted that it took years to negotiate this result and meant that cloud source code needed to be reviewed and, in some cases, rewritten to accommodate its requirement. Smaller CSPs may not be in the same position and CSPs that want to migrate sensitive workloads and data quickly may be prevented from doing so by the threat of extraterritorial access if the right PCP safeguards are not in place.

- **Conformance to stringent, carrier-grade SLAs for availability**. CSPs do experience outages, although less frequently than PCPs due to the resiliency of their infrastructure, which is mandated through regulation. A particularly notable outage occurred when Orange experienced a software failure in France, preventing 11 800 calls from connecting to emergency services over 7 hours on 2 July 2021. Due to the regulated nature of telecoms, large outages can result in direct governmental intervention. Smaller outages, such as occurred in O2's UK network in December 2018 as a result of an issue with components of its mobile core, receive nationwide publicity and can shift public perception on network reliability. For these reasons, CSPs are far more sensitive to availability metrics than PCPs. Many CSPs would face regulatory barriers if they tried to use a PCP that only has a single data centre in a country or region, even if that data centre offers multiple AZs and they struggle with the PCPs' ongoing lack of support for carrier-grade SLAs.

4.2 The 5G network needs a cloud platform, but should it be public?

These challenges are particularly pertinent as CSPs evaluate whether or not to use public clouds as the cloud platform for the industry's first cloud-natively designed network function, the 5G mobile core. 5G networks are designed to run on clouds and CSPs want to maximise the benefits that cloud can bring to the network. CSP interest in public cloud deployments of the 5G core has been sparked by two high-profile deals by AT&T with Azure and Dish with AWS. However, CSPs outside the USA, and the 5G core vendors working with them acknowledge that these two examples, which involve US CSPs partnering with US public cloud providers on US soil, represent a different proposition to the one facing non-US CSPs if they were to use the same PCPs in their countries under their regulatory regimes.

It is worth noting that most CSPs are considering deploying the cloud-native 5G SA core initially for enterprise use cases and not to support their highly regulated macro networks that carry consumer traffic. For certain enterprise use cases, the public cloud may provide a viable platform for the 5G SA core. Enterprise networks, which can be delivered as private network instances in the case of the 5G core, are not subject to as stringent regulation as the consumer network and the demand for availability is not as high. After all, enterprises are prepared to tolerate ‘three nines’ of availability for compute, so they may be more willing to accept ‘three nines’ for their private 5G networks as well.

However, enterprises are still subject to national and/or regional data protection legislation and in many cases, they are considering private 5G networks to support mission critical use cases associated with operational transformation. Enterprises in key sectors, such as finance, healthcare and manufacturing, will be as concerned about the compliance, availability and security aspects of their private networks as CSPs are about their own networks.

5G macro network deployments have so far typically been based on 5G extensions to virtualised or non-virtualised 4G Evolved Packet Cores (EPC), known as 5G non-standalone cores. Eventually, CSPs plan to migrate consumer traffic to the 5G SA core. Therefore, CSPs need to understand the drawbacks of running a 5G core in the public cloud, whether that core is destined to support consumer mobile broadband at scale or valuable enterprise use cases and network slices. These drawbacks include the following.

- **Complexity of deployment.** CSPs are unlikely to want to deploy the data-carrying 5G core User Plane Function (UPF) in the public cloud even if they believe they can circumvent sovereignty issues by using the public cloud for 5G core control plane functions. The 5G core has been designed to support the concept of control plane/user plane separation (CUPS) so running the control plane in a public cloud data centre and the user plane on-premises, managing the UPF as an appliance, is a relatively straightforward architecture. However, CSPs may want more flexibility in their deployment architecture, for example, they may want to co-locate the UPF with GNodeBs in base stations that support a 5G New Radio architecture, or to distribute UPF instances across other types of edge cloud platforms, potentially provided by multiple cloud/data centre providers with a local presence. It is much more difficult to realise these deployment scenarios if the 5G core control plane runs in a public cloud. A large Tier-1 European CSP pointed out that even if it ran its control plane in a public cloud, it would not make significant capex savings in the user plane because it would still need to dimension the user plane for peak usage. The data privacy laws to which the CSP is subject would make it too risky to use the public cloud’s ‘bursting’ function.

- **Cost of transporting charging data**. Hauling data out of a public cloud is expensive and CSPs need to consider the high costs that would be involved if they were to run the UPF in the public cloud and had to pull charging data out of it. This consideration is driving many CSPs to deploy the UPF on-premises.

- **Cost of high availability**. As we have seen, PCP outages can affect even the most cloud-natively designed applications, such as Netflix. Netflix invented the idea of chaos testing so that it could keep running no matter what happened in the cloud infrastructure beneath it, but even Netflix cannot plan for every contingency. Cloud-native network functions such as the 5G core can be built in a highly distributed and resilient manner, but the more distributed the components, the higher the processing costs, which may reduce the attractiveness of running a 5G core in the cloud. 5G vendors argue that PCPs need to improve the reliability and availability of their platforms, including offering 'five nines' SLAs, if they are serious about migrating CSP network traffic to their clouds.

- **Cost of encryption.** Encryption is a key feature of the 5G SA core standard, so the function itself can run securely on a public cloud, although the issue of the PCP holding root encryption keys to its cloud infrastructure and any PCP-owned supporting platform services that the network function calls upon still remains. If a CSP intends to run a 5G non-standalone core on a public cloud, it will need to add a layer of encryption that may significantly increase demand for cloud resources. The encryption and decryption of message exchanges between network elements and processes can generate as much traffic again as the core network transactions themselves, adding to operational costs.

- **Cost of supporting lawful interception (LI).** CSPs will need to ensure that they can extract LI data from the 5G core running in the public cloud themselves, because in most cases, their local security agencies will mandate them to process this highly confidential and sensitive data on-premises, under the national laws that govern LI.

- **Cost of failure.** The technical impacts of running a 5G core in the public cloud can be measured in advance. However, the reputational and commercial damage from data leaks, loss of availability and failures of compliance that may result are harder to assess and may only be quantified in retrospect. CSPs need to factor such risks into their assessments of the public cloud as a deployment platform for the 5G core and plan carefully to mitigate them.

4.3 A decision framework for assessing the right cloud environment for the 5G core

Figure 4.1 outlines key questions that CSPs should ask themselves when considering whether to build an on-premises cloud to run their 5G mobile core or to buy infrastructure as a service from a PCP to host this critical network function.

Figure 4.1: Considerations for selecting the public cloud as a hosting environment for the 5G core

Consideration	Decision criteria	Trade-off
Use cases • 5G SA core for consumer network • 5G NSA core for consumer network • 5G SA core for enterprise	• What cost savings (capex and opex) and other benefits does the public cloud bring to my use case? • What are the downsides of using the public cloud to support my use case (for example, deployment complexity, cost of encryption (5G NSA core), cost of failure, lack of SLAs)? • How business-critical/compliance-sensitive are the enterprise customers' use cases that I want to support with a public cloud-based core?	• Public cloud hosting of 5G SA/NSA core in the consumer network if benefits significantly outweigh costs and risks. • Public cloud hosting of 5G SA core to support enterprise use cases with low sensitivity to risk/regulation.
Regulatory environment	• How strict is the national/regional regulation that I need to comply with on issues of data location/residency? • How many different national/supranational jurisdictions do I need to support?	The stricter the legislation on data location/residency that a CSP faces, the stronger the propensity to enforce and the larger the number of jurisdictions a CSP needs to support,

Consideration	Decision criteria	Trade-off
	• How strongly does my nation/region enforce data protection regulation? • Can my nation/region protect me against extraterritorial legislation?	the more difficult it will be to use public cloud for 5G core.
Public cloud provider profile	• Does the public cloud provider have an in-country/in-region presence? • Can the PCP provide more than one data centre in my country/region? • What is the PCP's track record on availability? • What kind of SLAs can the PCP offer? • What guarantees of data sovereignty can the PCP provide?	A CSP may choose to run the 5G core control plane in the public cloud if the PCP has the right level of in-country/region presence to satisfy legislation, is addressing carrier-grade reliability and SLA commitments and is prepared to meet sovereignty demands.

Source: Analysys Mason

5. Conclusion

Legislative activity relating to data protection and privacy has increased significantly during the past couple of years. Countries across the world are introducing new or updating existing national and supranational legislation that governs the use of personal information to prepare their corporations and citizens to play a role in the growing digital economy. The renewed global emphasis on data protection legislation is coinciding with the rise of the PCPs that are supporting an increasing proportion of the compute infrastructure that underpins a global digital economy. Although there are undoubted benefits to using the public cloud for storing and processing data, companies will need to weigh these benefits against the need to conform to increasingly fragmented data protection laws that seek to strengthen the data sovereignty of different countries and territories. In addition, if companies operate in a highly regulated industry, as CSPs do, they will also need to evaluate whether public cloud infrastructure meets their requirements in terms of the availability and resilience it offers.

This is particularly true of network workloads that CSPs are considering migrating to the cloud, and in particular, the public cloud. A 5G network is a mission-critical business asset, as well as a regulated one, and CSPs must fully understand the benefits and the risks of migrating it to a public cloud environment. CSPs will need to adhere to national and supranational data protection regulation, which could add cost and complexity, and CSPs will need to seek commitments from public cloud providers, for example, over the extent of their in-country presence, ownership of cryptographic keys and SLAs. Since the requirements for 5G macro network workloads in each of these areas will be much higher than for their IT stacks, CSPs should consider carefully whether the public cloud is the best environment for such workloads and learn from the experience and decisions of their peers.

6. About the authors

Bence Szeidl (Research Analyst) is based in the London office and is a part of the *Cloud* team. His work focuses on operators' and vendors' activities around data management, AI, analytics and development tools. Bence holds a BSc in international management from the University of Warwick.

Joseph Attwood (Research Analyst) is based in our London office. He studied computer science at the University of Surrey and worked on the feasibility of implementing self-sovereign identity technology in his final-year project.

Caroline Chappell (Research Director) heads Analysys Mason's *Cloud* research practice. Her research focuses on service provider adoption of cloud to deliver business services, support digital transformation and re-architect fixed and mobile networks for the 5G era. She is a leading exponent of the edge computing market and its impact on service provider network deployments and new revenue opportunities. She monitors public cloud provider strategies for the telecoms industry and investigates how key cloud platform services can enhance service provider value. Caroline is a leading authority on the application of cloud-native technologies to the network and helps telecoms customers to devise strategies that exploit the powerful capabilities of cloud while mitigating its disruptive effects.

Analysys Mason Limited. Registered in England and Wales with company number 05177472. Registered office: North West Wing Bush House, Aldwych, London, England, WC2B 4PJ.

We have used reasonable care and skill to prepare this publication and are not responsible for any errors or omissions, or for the results obtained from the use of this publication. The opinions expressed are those of the authors only. All information is provided "as is", with no guarantee of completeness or accuracy, and without warranty of any kind, express or implied, including, but not limited to warranties of performance, merchantability and fitness for a particular purpose. In no event will we be liable to you or any third party for any decision made or action taken in reliance on the information, including but not limited to investment decisions, or for any loss (including consequential, special or similar losses), even if advised of the possibility of such losses.