Internet Way of Networking Use Case
Data Localization

September 2020

How mandatory data localization impacts the Internet Way of Networking

This use case analyzes the effect that government policies regarding data localization may have on the Internet Way of Networking. To understand how such policies could undermine the Internet's broader benefits such as innovation and socioeconomic growth, we view them through the lens of the Internet's critical properties.

What is mandatory data localization?

Mandatory data localization refers to government requirements that control the storage and flow of data to keep it within a particular jurisdiction. Data localization laws - sometimes called "data residency" or "data sovereignty" - are typically intended to keep personal or financial transaction data in-country where they are subject to access and local regulation. Mandatory data localization measures range from obligations to physically locate data in the country where it originates, to restricting or even forbidding its transfer to other countries. What does mandatory data localization mean for the Internet's critical properties, and what would happen if more countries imposed these restrictions?

Current trends

In the past few years, India, Indonesia, and Vietnam have considered or introduced laws requiring personal or business data to be kept within national borders and not processed in other countries.[1] While India's 2019 Personal Data Protection Act ultimately discarded measures to keep all personal data-processing geographically located in India, it still forces the localization of an undefined set of "critical personal data". Indonesia has had mandatory data localization measures since 2012, although they were somewhat relaxed in 2019. Vietnam's 2019 Law on Cybersecurity initially required all non-resident Internet services firms that processed Vietnamese personal data to create a physical presence in the country, but this requirement was targeted more narrowly in subsequent legislation.

[1] https://thediplomat.com/2020/01/the-retreat-of-the-data-localization-brigade-india-indonesia-and-vietnam/ and https://www.fticonsulting-asia.com/~/media/Files/apac-files/insights/articles/localization-to-fragment-data-flows-asia.pdf

But while some countries considered, and then at least partly stepped back, from forcing businesses to keep personal and commercial data within their borders[2], still "in the last few years, over seventy countries have passed new or updated data privacy laws that include some form of data localization".[3]

Recent laws in Russia and China forbid companies from sending citizens' personal data outside the country. A 2019 Russian law imposes fines on companies and employees that fail to comply with the country's 2015 data localization law (which itself resulted in the blocking of the LinkedIn website in Russia). China's 2017 Cybersecurity Law requires critical infrastructure operators and network operators to store "important data" – both personal and commercial - in China, or complete a broad and stringent "security assessment" to request the ability to export the data. These laws have resulted in companies carrying higher burdens and risks, restricting the availability of value-added services. Many companies have exited those markets altogether.

Data localization measures typically focus on personal or commercial data, and so are mostly targeted at companies that process that data, e.g. business-to-consumer firms, banks, and technology platforms dealing predominantly with third-party content.

Current policies are also focused on data "at rest", i.e. data that is not actively moving from device to device or network to network, such as data stored on a hard drive, a laptop or archived/stored in some other way. Therefore, they do not usually target Internet infrastructure services that carry this data without knowing its content. The current laws – stretching to different extremes – tend to confine such data within national borders, reflecting broader geopolitical realities. As nations shift towards more nationalistic and sovereignty-based approaches that question globalization and international cooperation,[4] data localization becomes a powerful means of allowing state actors to erect 'digital boundaries' in the Internet.[5] Examples of this trend include Europe's recent cloud federated architecture project, GAIA-x, which is premised around the notion of "data sovereignty";[6] additionally, Brunei, China, Indonesia, Nigeria and Russia already have strict data localization laws, specifically requiring storage on servers within the country itself.[7]

But there are already instances where countries impose even more extreme data localization measures, including data "in transit". For example, Russia and China are planning to introduce measures to centralize, control, and restrict Internet infrastructure services, driving Internet fragmentation at all levels [ref to Russia's "Sovereign Internet" law and Chinese "Draft Data Security Management Measures of 2020, Article 29"].

If such a trend towards data localization continues, it will create a more constricted and less resilient network with suboptimal performance, retrofitted to comply with national borders. Businesses will have to narrow their choices and capabilities, and network operators may be forced to use costly and less resilient ways to route traffic. Cybersecurity may suffer as organizations are less able to store data outside borders with the aim of increasing reliability and mitigating a wide variety of risks including cyber-attacks and national disasters.

Countries trying to forcibly localize data will impede the openness and accessibility of the global Internet. Data will not be able to flow uninterrupted on the basis of network efficiency; rather, special arrangements will need to be put in place in order for that data to stay within the confines of a jurisdiction. The result will be increased barriers to entry, to the detriment of users, businesses and governments seeking to access the Internet. Ultimately, forced data localization makes the Internet less resilient, less global, more costly, and less valuable.

[2] https://thediplomat.com/2020/01/the-retreat-of-the-data-localization-brigade-india-indonesia-and-vietnam/

[3] https://foreignpolicy.com/2020/05/13/data-governance-privacy-internet-regulation-localization-global-technology-power-map/

[4] https://www.theatlantic.com/ideas/archive/2020/03/dont-abandon-globalizationmake-it-better/608872/

[5] https://www.centerforfinancialinclusion.org/data-globalization-vs-data-localization

[6] https://www.bmwi.de/Redaktion/EN/Publikationen/Digitale-Welt/das-projekt-gaia-x-executive-summary.pdf?__blob=publicationFile&v=6

[7] https://bigbang360.com/data-localization-laws/

Which critical properties does forced data localization affect?

Critical Property 1 – An open and accessible infrastructure with a common protocol

The only essential condition for a network or node to access the Internet is to adopt its common protocols, the Internet Protocol (IP) at the minimum. This "permissionless" model of the lowest possible technical barrier to entry is the basis of the Internet's rapid growth and global reach. It does not require network operators to operate in ways that match national borders as they exchange traffic from network to network.

One of the negative effects of data localization policies is the high barriers to entry for new service providers, which, in turn, affect global connectivity and growth. Under a data localization regime, a service provider, such as a web portal, would not only be required to set up a web server (or host it in a CDN) to ensure basic services, but also provide separate storage facilities in the countries where data localization laws exist.

Data localization laws, such as those considered in India or Vietnam, typically target the processing and use of specific categories of personal and business information at the application layer of the Internet, like a cloud computing application. They do not target the Internet's infrastructure providers directly by requiring traffic passing through networks to conform to national borders. However, we tend to observe the unintended consequences they produce, especially in relation to the increased costs associated with running the associated services. Data localization policies require service providers to create additional hosting facilities, which in turn means they need reliable infrastructure to keep the data physically in the country. Because not all countries are capable or equipped to facilitate such infrastructure, service providers could encounter additional costs and possible infrastructure vulnerabilities, or simply choose to not let their services in those countries.

Critical Property 3 – Decentralized management and a common distributed routing system

The Internet is a "network of networks", made up of almost 70,000 independent networks that use the same technical protocols and choose to collaborate and connect together. Each network makes independent decisions on how to route traffic to its neighbours, based on its own needs, business model, and local requirements. There is no centralized control or coordination.

One of the negative consequences of data localization is that data may not be stored optimally, both in terms of resilience and connectivity. The topologically closest (and therefore fastest) place in the network to put data may not be in the same country. Data is stored where it makes most sense – and this involves dynamic considerations of efficiency and performance reliability rather than location. Even if data is located in one country, the transmission path may cross national borders for resilience or performance reasons. Data localization measures may either directly or indirectly force Internet data to follow national borders at the expense of efficiency.

Some policies go even further and also consider data in transit. Although there is a range of approaches to data localization, one of the ways it may be enforced is by requiring Internet intermediaries to impose additional requirements on routing policy. For instance, article 29 of China's draft Data Security Management Measures stipulates that when users in China access local websites, their traffic should not be routed to servers outside of China.[8] Such policies could impact how information is transmitted between networks, potentially undermining efforts to reduce latency, provide redundancy and replication to distribute data closer to its destination, and other basic traffic-engineering and traffic-optimization goals. This would reduce network operators' routing autonomy and their ability to optimize connectivity. Overall, aligning routing policy with the requirements of different jurisdictions creates needless complexity and inefficiency, as routing would no longer serve the technical requirements of connectivity, resilience, and optimized flow.

[8] Original law: http://www.moj.gov.cn/news/content/2019-05/28/zlk_235861.html; unofficial translation, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinas-new-draft-data-security-management-measures/

If current trends continue, forced data localization would interfere with the autonomous and agile distributed routing of the Internet, reducing the ability to collaborate with other networks and ultimately constraining the Internet's global reach.

Critical Property 5 – A Technology Neutral, General-Purpose network

The Internet is a 'technology neutral, general-purpose network' because there is no defined limit to the uses its infrastructure can support. A general-purpose network requires operators of network services to perform only very basic functions: passing data packets on to its next destination without caring about their content.

Forced data localization would require limits to the services that can be offered in specific countries if those services involve sending personal or commercial data across borders. While current laws are unlikely to immediately require direct changes by network providers, these requirements may filter down over time. Harsher data localization regimes would bring a greater need for coordination between companies and governments to determine what data networks are carrying, and between networks to ensure specified traffic flows follow national borders. Any additional requirements based on all operators understanding the nature of the data/content would make the network more specialized and less general-purpose, needing additional functionalities such as traffic inspection, and would more narrowly prescribe the functions of networks overall.

The loss of simplicity and basic functionality at the Internet's transit layers caused by data localization measures would make networks more complex and less efficient, with an increased need for coordination. This would undermine the Internet's model of permissionless innovation and create barriers to entry for new network operators and Internet infrastructure providers.

Conclusion

While some countries in South Asia have recently stepped back from imposing strict data localization laws, in other regions such as the European Union new measures to boost "data sovereignty" are under consideration.[9] If the data localization trend continues, it will restrict services like cloud computing that can be offered to Internet users in different countries, shaping the Internet as many people use it today into a more nationally based experience. Data localization measures designed to change business practices also risk shaping and constraining the unimpeded flow of traffic in the Internet's infrastructure. The impact of forced data localization laws will ultimately trickle down to the Internet's infrastructure and undermine the critical properties of the Internet Way of Networking.

This likely impact on the critical properties will lower the value of the Internet to all users around the world as it is no longer an 'end-to-end' network offering people everywhere the widest range of opportunities.

[9] https://www.bmwi.de/Redaktion/EN/Dossier/gaia-x.html