Emerging Tech: Leveraging Hyperscale Clouds to Address Compliance and Network Latency Requirements for SaaS and PaaS

Published 22 November 2022 - By Evan Zeng, Michael Warrilow

https://www.gartner.com/document/4021447?ref=follow-dashboard

Hyperscale cloud providers continue to build global infrastructure and are extending distributed cloud offerings. Product leaders can leverage such infrastructure to address residency and locality mandates and network latency requirements for SaaS and PaaS delivery.

Overview - Key Findings

- Few hyperscale cloud providers have in-country cloud data center footprints (cloud regions and/or availability zones) in all countries with residency and locality mandates. This challenges SaaS and PaaS providers to add additional hosting infrastructure for compliance reasons.

- Global leaders of hyperscale cloud platforms do not always have the best cloud networking infrastructure in all geographical regions and/or countries that product leaders need. But cloud network performance is a key factor for SaaS and PaaS delivery and to enable low-latency services.
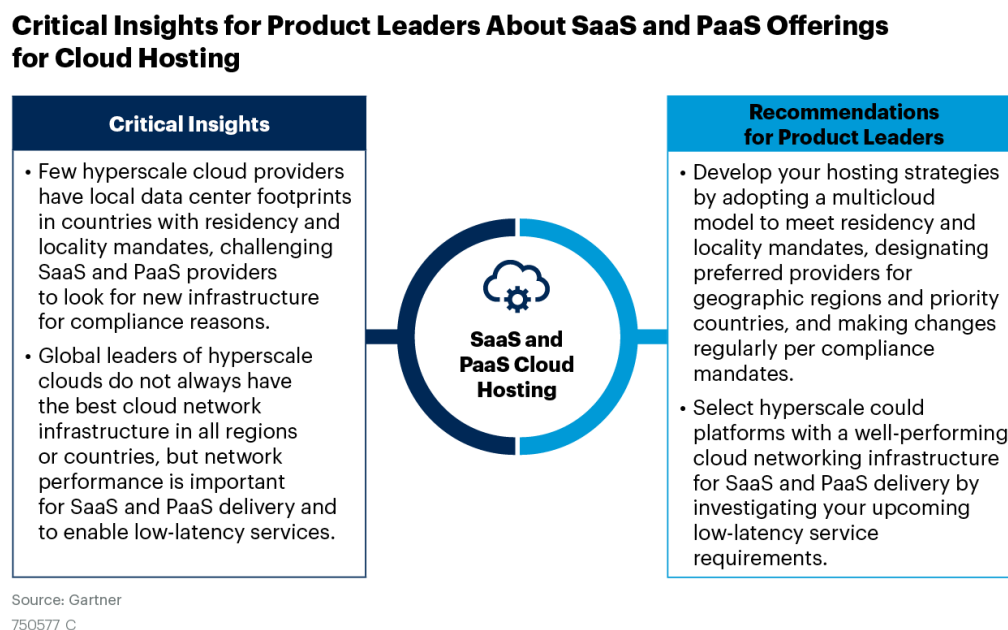
Recommendations

Product leaders focusing on emerging technologies and trends related to SaaS and PaaS offerings should take these steps:

- Develop your hosting strategies by adopting a multicloud model to meet residency and locality mandates of different jurisdictions. Designate a preferred provider for each geographical region and priority country, and regularly make changes based on refreshed regulatory mandates.

- Select hyperscale cloud platforms with well-performing cloud networking infrastructure for SaaS and PaaS delivery by investigating your low-latency service requirements for the next few years. Include cloud networking performance, especially network latency, as a key criteria for cloud platform selection, and regularly test the platform's cloud network performance.

Analysis

Figure 1 is a summary of the critical insights we discuss in this document.

Figure 1: Critical Insights for Product Leaders About SaaS and PaaS Offerings for Cloud Hosting

Critical Insights for Product Leaders About SaaS and PaaS Offerings for Cloud Hosting

Critical Insights

- Few hyperscale cloud providers have local data center footprints in countries with residency and locality mandates, challenging SaaS and PaaS providers to look for new infrastructure for compliance reasons.
- Global leaders of hyperscale clouds do not always have the best cloud network infrastructure in all regions or countries, but network performance is important for SaaS and PaaS delivery and to enable low-latency services.

SaaS and PaaS Cloud Hosting

Recommendations for Product Leaders

- Develop your hosting strategies by adopting a multicloud model to meet residency and locality mandates, designating preferred providers for geographic regions and priority countries, and making changes regularly per compliance mandates.
- Select hyperscale could platforms with a well-performing cloud networking infrastructure for SaaS and PaaS delivery by investigating your upcoming low-latency service requirements.

Source: Gartner
750577_C

Gartner

Description

SaaS and PaaS providers can self-build an infrastructure platform to host their services. Examples of self-hosted giants include ByteDance, eBay and Salesforce. Alternatively, they can leverage public cloud infrastructure to host their services. Gartner sees the latter

approach being increasingly adopted by SaaS and PaaS providers. For example, although Salesforce was self-hosted, it announced Hyperforce in 2020, which unifies the foundation of various public clouds to scale Salesforce rapidly. Such an approach requires less upfront investment and is more affordable for most SaaS and PaaS providers. The pay-as-you-go infrastructure from public cloud providers can offer better pricing flexibility. Also, using existing public cloud infrastructure instead of self-building it can offer much better time to market for providers to build global coverage.

Furthermore, hosting SaaS and PaaS services on public clouds can better leverage the trend of enterprises' IT spending shifting to cloud. Product leaders who fail to adapt to the shift to cloud face increasing risk of suboptimal hosting and less intimacy with cloud users. Gartner forecasts that nearly 60% of IT spending on application software will be directed toward cloud technologies by 2024 (see Market Impact: Cloud Shift —
2022 Through 2025).

SaaS and PaaS providers can neither rely on a few cloud data centers and the internet to deliver services worldwide, nor rely on one single hyperscale cloud platform to meet all compliance mandates. The increasing volume of regulatory and legislative requirements, along with emerging latency-critical service requirements, is driving such providers' workload placement decisions. As these drivers increase, product leaders must increasingly leverage a cloud infrastructure platform with in-country and/or in-region footprints for service delivery. For example, SaaS providers need such a platform to host their SaaS applications and customer data.

A sample of providers leveraging multicloud infrastructure to deliver various SaaS and PaaS services include (in alphabetical order):

- Alkira
- Cisco (Nexus Cloud)
- Snowflake
- Teridion
- VMware (NSX Cloud)

- Salesforce (Hyperforce)
- Tencent (Games)

The sample providers shown above are cited as examples and do not imply an exhaustive list.

Critical Insight: The Availability of Local Cloud Footprints Will Drive SaaS and PaaS Providers to Evaluate Multiple Hyperscale Cloud Platforms for Residency and Locality Compliance

As society becomes more dependent on digital information and globally interconnected, more regulation and legislation — hereafter called compliance mandates — are being created to control and protect citizens, consumers and companies across regions and nations. Fundamental to these drivers is a set of interconnected information requirements (namely, data residency and data localization) to protect sensitive, personal and/or confidential electronic data. For more information, see Emerging Technologies: Compliance Regimes Will Break Public Cloud Scale and Innovation.

Local Cloud Footprints Are Foundational for Residency and Locality Compliance

Gartner breaks compliance regimes into six layers of cloud sovereignty actions for technology and service providers (TSPs): protection, data privacy, residency (movement and rest), locality (physical location), authority (decision rights) and ownership (legal rights and independence). (See Quick Answer: How Can Hyperscale Cloud Providers Address the Growing Dilemma of Cloud Sovereignty?) Although all hyperscale cloud providers focus on these six layers to improve their cloud platforms' sovereignty capabilities, the cloud sovereignty layers of residency and locality are foundational for hosting SaaS and PaaS

offerings in the cloud. This is because these two layers require the hyperscale cloud platform to have a local data center footprint — a cloud region or an availability zone — within the physical boundary of a jurisdiction or an approved substitute location. Without meeting this, SaaS and PaaS providers cannot leverage their selected hyperscale cloud platforms to meet the layers of sovereignty requirements in a country that has demanded that SaaS and PaaS providers do so.

Figure 2 and Gartner's Tool: Public Cloud Services Location Map, 3Q22 visualizes the cloud regions for seven of the cloud infrastructure and platform services (CIPS) providers listed in Magic Quadrant for Cloud Infrastructure and Platform Services.

Figure 2: Worldwide Cloud Region Footprints of Major Hyperscale Clouds

Worldwide Cloud Region Footprints of Major Hyperscale Clouds
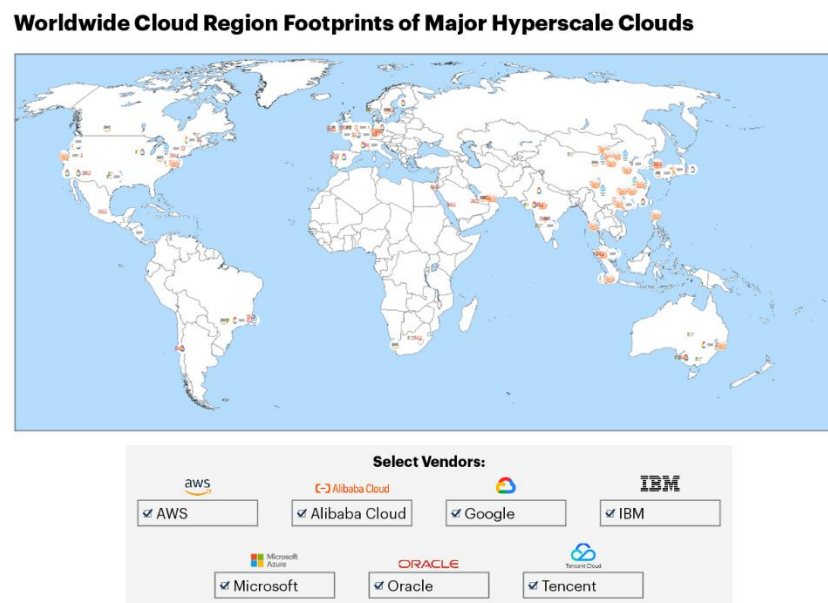
Source: Gartner
779247_C

Cloud infrastructure can meet one of the following requirements for SaaS and PaaS hosting, if it involves sensitive data subject to compliance mandates within multiple jurisdictions:

- Scenario 1: Must have in-country data center footprints (cloud regions and/or availability zones) within the physical boundary of a jurisdiction

- Scenario 2: Must have in-region data center footprints (cloud regions and/or availability zones) in neighboring countries where a jurisdiction has been approved as a substitute location satisfying the home country's regulatory requirements

In selecting cloud infrastructure for SaaS and PaaS hosting, Scenario 1 is better than Scenario 2. This is because the former can be broadly, consistently applicable to all data, including data categorized as very sensitive, such as personally identifiable information. The latter, meanwhile, is still subject to the scope of data being approved to reside at substitute locations based on relevant compliance mandates.

Few hyperscale cloud providers have in-country cloud data center footprints (cloud regions and/or availability zones) in all countries that have data residency and locality mandates.

To meet data residency and locality mandates worldwide, SaaS and PaaS providers should select a combination of different hyperscale clouds, depending on their business opportunities and compliance mandates in target markets. Today, many SaaS and PaaS providers select one or two global market share leaders as their strategic cloud hosting providers. However, they don't always consider whether those providers have cloud data center footprints in all required markets. Examples include China, Southeast Asia and Latin America, where the global market share leaders like AWS and Microsoft have fewer in-country cloud region footprints than some of their peers (see Figure 2 above).

For countries where the selected hyperscale cloud providers currently do not have local data center footprints, a different approach will be needed. If compliance mandates allow the Scenario 2 infrastructure as a substitute, SaaS and PaaS providers might have a choice. They might keep using existing providers if such providers have Scenario 2 substitutes at

targeted countries; they might also involve a new provider that has local data center footprints there. This is because adding a new provider will make the cloud hosting architecture more complex. But it brings benefits to hosting services' compliance because the Scenario 2 infrastructure has limitations on the scope of data being allowed to reside outside.

Some people might argue that, in countries where the selected hyperscale cloud providers do not have local data center footprints, SaaS and PaaS providers can leverage the hyperscalers' distributed cloud solutions to implement a local footprint. For example, such providers can adopt AWS Outposts to install a local footprint within a jurisdiction to meet the data residency requirement. This might satisfy the residence and locality compliance requirements but might not provide sufficient cloud services. Gartner's Tool: Public Cloud Services Location Map, 3Q22 identifies a cloud region as a physical and logical grouping of data centers designed to deliver public cloud services and located within a broad metropolitan area. Distributed cloud footprints, such as metro area community cloud and on-premises public cloud, are hyperscale cloud substations that do not meet the criteria of a cloud region.

From a resilience and high-availability perspective, a cloud region is a better choice than a distributed cloud substation to host mission-critical SaaS and PaaS services and data. Further, a cloud region generally has much more diversity of cloud services, such as security and networking services, than a distributed cloud substation, offering better choices for such providers.

Near-Term Implications for Product Leaders

- Expect the continuous growth of compliance mandates for SaaS and PaaS business. A regular review of compliance requirements is required for all jurisdictions that involve SaaS and PaaS business. It's also necessary to check whether current cloud hosting infrastructure can meet all compliance requirements.

- There will be a need to collaborate across a diverse range of internal stakeholders, including various technology teams, legal and compliance, and business units. The goal of the collaboration will be to identify compliance requirements in each

jurisdiction that has SaaS and PaaS business, the impacts to current service delivery and cloud hosting, and how to address these potential impacts.

Recommended Actions for the Next Six to 18 Months

- Build a decision framework to determine possible options if the existing provider does not have the necessary cloud data center footprints for compliance needs. The best options might be to add a new cloud provider with cloud regions in targeted countries or a colocation provider with data centers there, or to install existing cloud providers' distributed cloud substations. This requires case-by-case assessment, such as the analysis of detailed business growth projections, currently available skills of the targeted cloud providers, and available budget and financial support model.

- Develop your SaaS and PaaS multicloud hosting strategy by identifying one or more strategic hyperscale cloud providers that can meet your cloud hosting location requirements, and develop a cloud-agnostic architecture to unify the foundation of your selected public clouds. Because few providers have cloud hosting infrastructure in all locations that will meet compliance requirements, product leaders must work across multiple stakeholders to identify the right combination of locations and providers.

- Regularly revisit compliance requirement updates for each geographical region and priority country to see whether you need to make changes to your cloud hosting infrastructure according to location requirements.

Critical Insight: Increased In-Region and/or In-Country Cloud Footprints Will Improve Network Peering and Performance

Network performance, especially network latency, is always a spotlight metric that SaaS and PaaS providers measure on their cloud hosting infrastructure platforms. This is critical

because reducing network latency between a service instance and its consumers will enhance user experience and enable latency-critical service offerings.

Table 1 shows three types of in-country and/or in-region cloud infrastructure that can offer on-demand cloud services and are currently available. It also shows how they can reduce network latency to service consumers and what major business benefits they can deliver to hosting services.

Table 1: In-country and/or In-Region Cloud Infrastructures to Reduce Network Latency for Hosting services

Cloud Infrastructure	How It Reduces Network Latency	Benefits to Cloud Hosting Services
Cloud Region	Optimizes network path to service consumers by improving service proximity, enabling private peering with local partners (such as carriers) and optimizing local region routes to service consumers	Enhances user experience by reducing response time
Metro Area Community Cloud	Improves service proximity to service consumers by hosting service instances at the same metro network with its consumers	Enables the support of low-latency services, such as those requiring sub-10 ms roundtrip latency (e.g., AR/VR services)
5G Mobile Edge Cloud	Improves mobile service proximity to service consumers within the same provider's mobile network in one metropolitan area	Enables the support of low-latency *mobile* services, such as those requiring sub-10 ms roundtrip latency (e.g., cloud gaming and AR/VR)

Note: Distributed cloud software or hardware appliance solutions

Source: Gartner (November 2022)

Leading cloud providers do not always have more cloud regions in any country or geographic region than other players. Thus, they don't always have the best-performing cloud networks in the locations you need. For example, among the seven selected providers shown in Figure 2 above, Oracle has the most cloud regions in the Middle East and North Africa, and Alibaba Cloud has the most cloud regions in Southeast Asia. For more information on how to address this situation if/when necessary, see Differentiate

CIPS Offerings to Compete Against Infrastructure-Led Hyperscale Cloud Providers Outside Their Home Markets.

SaaS and PaaS product leaders should include network performance as a key selection criteria for cloud hosting platforms. They should look at:

- The number of cloud regions that cloud platforms have locally (in-region and/or in-country)
- Local private and public peering partner ecosystems
- Ability to broadcast cloud region and SaaS and/or PaaS segmented routes to the internet
- Regular cloud network performance testing between cloud data centers and the service consumers

When designing a cloud hosting strategy, product leaders should shortlist cloud providers with a superior cloud network infrastructure for targeted countries or geographic regions. They should also develop a guiding framework based on the latency requirements of their cloud hosting workloads. Then, they can decide whether to add a new cloud provider for a targeted market if the current shortlisted providers do not have well-performing local cloud network infrastructure.

Furthermore, the emergence of low-latency workloads will increasingly require cloud hosting infrastructure to be able to deliver low-latency network services such as those with roundtrip latency within 10 ms to service consumers in a metropolitan area. Examples include robot control, highly interactive AR/VR video and cloud gaming. Such cloud infrastructures are often hyperscale clouds' distributed cloud substations that are located at the network edge and have connections back to their centralized cloud regions. Currently, hyperscale cloud providers are racing to deploy their distributed cloud infrastructure worldwide to meet such demands. For example, AWS recently announced it would launch 32 local zones in 26 countries with availability starting in 2022.

If involved with low-latency service delivery, product leaders must design a network infrastructure strategy to support such services. Very often, the inclusion of in-metro cloud infrastructure can be the extension of a current cloud hosting service strategy. First, they

should investigate what metropolitan areas have demand for low-latency service delivery and which cloud providers can offer in-metro cloud hosting infrastructure. They can also find such information in Tool: Public Cloud Services Location Map, 3Q22.

Distributed cloud infrastructure is still a nascent area for hyperscale cloud providers. So product leaders should discuss with their current strategic cloud providers to understand their roadmap to deploying metro area community cloud and/or 5G mobile edge cloud at targeted markets. Hyperscale cloud providers, if they have not had a rollout plan for your targeted countries, will often respond with distributed cloud software or hardware appliance solutions that SaaS and PaaS providers can deploy at locations of their choice. Product leaders should include such an option from their shortlisted cloud providers into their cloud hosting strategy.

Near-Term Implications for Product Leaders

- Improved latency for cloud hosting platforms is a must-have to improve SaaS and PaaS customer experience and enable latency-critical services.

- Cloud network performance can be improved by carefully choosing different categories of cloud infrastructure as stated in Table 1 and by selecting the in-country or in-region cloud regions with the best cloud network performance and peering partners.

Recommended Actions for the Next Six to 18 Months

- Include cloud network performance, especially network latency, among the key criteria for cloud platform selection; always shortlist cloud providers with well-performing cloud networks in your targeted markets.

- Investigate your low-latency service delivery demands for the next three years. Discuss with your shortlisted cloud providers whether they have a distributed cloud deployment roadmap for your targeted markets.

- Start to build cloud hosting SaaS and PaaS architecture with shortlisted providers' distributed cloud software or hardware appliance solutions. This will be useful when

you need in-metro or in-country cloud infrastructure for service compliance and/or hosting network performance. This is applicable when your shortlisted providers have not had such cloud footprints and you don't want to add new cloud providers.

- Regularly revisit whether multicloud hosting service infrastructure can meet the growing demands of hosting network performance requirements, including low-latency services at different geographical regions and priority countries, by continuously testing network performance between cloud hosting service instances and their consumers.