CARNEGIE ENDOWMENT FOR INTERNATIONAL PEACE

Cloud Governance Project
Resilience

Data Retrievability and Backup Arrangements

Data retrievability and backup arrangements are critical to enhancing the capacity to restore services functionality following data loss, corruption, or broader disruption to cloud services and digital infrastructure.[1] This issue addresses the challenges associated with enterprise customers setting up, using, and trusting cloud backup and recovery arrangements.

Key Considerations

- ***Customers may not be aware of or be able to configure and enable backup arrangements.*** Enterprise customers share responsibility with cloud providers for configuring backup arrangements. However, customers are not always capable of using the available tools and controls to ensure that the confidentiality and integrity of stored data is not compromised. Moreover, some customers may not enable backups or may rely on a single provider for those services. As a result, they may be unable to recover data in the event that their provider of choice experiences a service disruption.[2]
- ***Data retrievability and backup arrangements are themselves vulnerable to malicious and nonmalicious triggers of failure.*** While they are generally seen as contingency arrangements in the event of attack or disruption, data backup arrangements can themselves be disrupted, preventing or hindering recovery or compromising the confidentiality and integrity of stored data.
- ***Localization requirements may complicate recovery arrangements.*** While the locations of data centers are often carefully selected to minimize the likelihood of disruption (for example, physical disruption and internet outage), data localization requirements imposed by enterprises and/or governments may prevent providers from offering storage arrangements in optimal locations[3] that might be outside a specific jurisdiction.

Stakeholder Perspectives

Government

- May require retrievability and backup arrangements for critical functions or infrastructure that

Providers

- Encourage customers to adopt providers' data retrievability and backup services to avoid data loss or corruption,[4] but may argue that

Customers

- Want to maintain service availability in the event of data loss or service disruption by retrieving backup data.

Others

- *Insurers:* May require customers to enable retrievability and backup arrangements as a condition for

operate on the cloud.

customers are ultimately responsible for appropriately configuring these tools.

extending coverage.

Tensions with Other Cloud Governance Issues

- ***Localization and Routing Requirements:*** Localization mandates may prevent the storage of data in optimal locations outside of the home government's jurisdiction and the relocation of data to such data centers.[5] This can potentially prevent customers from building in geographic and functional resilience in the case of regional outages affecting that jurisdiction.
- ***Equitable Cloud Access:*** Backup strategies need to account for cases where there may be limited bandwidth availability, as this will impact how much data can be backed up with any regularity as well as efforts to access cloud-hosted data in the event of an outage or disruption
- ***Portability and Interoperability:*** The inability to retrieve data may slow or prevent efforts to restore service in the event of cloud outage. A lack of portability and interoperability will also make data retrievability arrangements less useful, if backup data and workloads cannot be transferred to functioning cloud environments

Potential Ways Ahead

Government

- Encourage diversity in routing[6] through additional geographies and geo-redundant deployments (for example, through stringent re-routing requirements and residency requirements).
- Leverage existing encryption protocols and standards for

Providers

- Create multiple retrieval and backup strategies and mechanisms for each service offered and which develop contingency plans for widespread disruptions from diverse sources.[7]
- Provide customers with guidance materials in support of their

Customers

- *Enterprise customers:* Configur e a disaster recovery strategy for each application and service.[10]
- *Enterprise customers:* Regularly carry out drills to test disaster recovery plans.[11] (Shared with cloud providers.)

Others

- *Insurers:* To encourage widespread and effective use of data retrievability and backup arrangements, insurers may consider requiring data retrievability and backup arrangements as

data in transit as baselines.	development of disaster recovery and data restoration strategies. • Regularly carry out drills to test disaster recovery plans. (Shared with enterprise customers.) • Increase transparency of their global dependencies that may impact customers' capacity to retrieve and/or backup data. • Consider adopting practices compliant with international and other key information security standards (such as SOC 2[8] and ISO27001[9]). (Shared with enterprise customers.)	• *Enterprise customers:* Consider alternatives to their primary cloud provider when procuring backup services, and consider the routine use of at least one additional cloud provider for regular operations. • *Enterprise customers:* In addition to cloud-based backups, consider enabling on-premises backups for the most critical workloads. • *Enterprise customers:* Consider incorporating disaster and data recovery strategies into contractual requirements (facilitated by industry's development of guidance materials). • *Enterprise customers:* Consider adopting practices compliant with international and other key information security standards (such as SOC 2[12] and ISO27001[13]).	part of their conditions for coverage. • *Insurers:* Facilitate conversations between governments, customers, and providers on the latter's global dependencies and common mode failures.

(Shared with cloud providers.)

Recent Examples

- Internet blackout in Seoul cut off 210,000 individuals and businesses, effectively halting retailers' commercial transactions as they were unable to process credit card payments or to access their cloud backup systems. For additional information, see: "KT restores most services after fire, but issues linger," Korea JoongAng Daily, November 25, 2018 and "58% of Data backups are Failing, creating Data Protection Challenges and Limiting Digital Transformation Initiatives," Veeam, March 18, 2021.
- OVH fire in France destroyed one of the provider's data centers, rendering various services "unrecoverable." For additional information, see: "OVH data centre destroyed by fire in Strasbourg – all services unavailable," The Register, March 10, 2021 and "OVH says some customer data and configs can't be recovered after fire, some seems to be OK, plenty is safe," The Register, March 15, 2021.
- Capital One's 2019 cloud misconfiguration case. For additional information, see: "Everything We Know About the Capital One Hacking Case So Far," Wired, August 29, 2019.

Notes

[1] A 2021 survey by backup provider Veeam revealed that 58 percent of noncloud backups fail, which is especially concerning against the backdrop of a reported 95 percent of firms experiencing unexpected outages in the year 2020. See: Veeam, "58% of Data backups are Failing, creating Data Protection Challenges and Limiting Digital Transformation Initiatives," Veeam, March 18, 2021, https://www.veeam.com/news/cxo-research-58-percent-of-data-backups-are-failing-creating-data-protection-challenges-and-limiting-digital-transformation-initiatives.html.

[2] "Many tweets from OVH customers said they stored their backups on another server in the same data center that burned, which means their primary and backup data were destroyed by the fire. Others felt it was OVH's responsibility to protect their data from a data-center fire, so they made no provisions at all for backups." See: W. Curtis Preston, "Backup lessons from a cloud-storage disaster," Network World, April 23, 2021, https://www.networkworld.com/article/3615678/backup-lessons-from-a-cloud-storage-disaster.html.

[3] The location of a data center is dependent on several factors including customer requirements and the results of a rigorous risk assessment process. Cloud providers often refer to these

locations as “availability zones.” Some organizations require greater resilience and, therefore, opt for “high-availability zones.” See: Microsoft, “Regions and availability zones,” Microsoft, November 11, 2021, https://docs.microsoft.com/en-us/azure/availability-zones/az-overview; and AWS, “Regions and Availability Zones,” Amazon Web Services (AWS), n.d., https://aws.amazon.com/about-aws/global-infrastructure/regions_az/.

[4] Hyperscale providers offer geo-redundancy, protecting against regional service disruptions and data center failures, through high-availability offerings in independent zones equipped with data centers, independent power, cooling, and networking.

[5] The location of a data center is dependent on several factors including customer requirements and the results of a rigorous risk assessment process. Cloud providers often refer to these locations as “availability zones.” Some organizations require greater resilience and, therefore, opt for “high-availability zones.” See: Microsoft, “Regions and availability zones,” Microsoft, November 11, 2021, https://docs.microsoft.com/en-us/azure/availability-zones/az-overview; and AWS, “Regions and Availability Zones,” Amazon Web Services (AWS), n.d., https://aws.amazon.com/about-aws/global-infrastructure/regions_az/.

[6] Telecoms World, “Diverse Routing,” Telecoms World, n.d., https://www.telecomsworld.com/diverse-routing.

[7] These strategies can include arrangements for failover across regions, load balancers, application gateways, and more, and should account for the people, processes, and applications needed to restore functionality. Moreover, they should be fully and regularly tested through disaster simulations. For example, Microsoft Azure’s locally redundant storage is advertised as providing low-cost single region durability, geo-redundant storage for high durability across regions, and zonal redundant storage for intra-region high durability. See: Microsoft, “Preview of Zonal redundant Storage for Backup data from Azure Backup,” Microsoft Azure, September 22, 2020, https://azure.microsoft.com/en-us/updates/preview-of-zonal-redundant-storage-for-backup-data-from-azure-backup/.

[8] AICPA, “SOC 2 – SOC for Service Organizations: Trust Services Criteria,” AICPA, n.d., https://us.aicpa.org/interestareas/frc/assuranceadvisoryservices/aicpasoc2report.

[9] ISO, “ISO/IEC 27001: Information Security Management,” ISO, n.d., https://www.iso.org/isoiec-27001-information-security.html.

[10] These strategies can include arrangements for failover across regions, load balancers, application gateways, and more, and should as well include a complementary data backup strategy (for example, how frequent should the backup process be, how extensive, should they be simultaneous across all applications, and so on) and a strategy on how to address lost data. A disaster recovery plan should also account for the people, processes, and applications needed to restore functionality, and should be fully and regularly tested through disaster simulations.

[11] Providers may issue guidance to help customers simulate disaster scenarios to test their recovery strategies against. For instance, see: Microsoft, "Performing disaster recovery drills," Microsoft Azure, October 18, 2021, https://docs.microsoft.com/en-us/azure/azure-sql/database/disaster-recovery-drills and Google, "Disaster recovery scenarios for data," Google Cloud, n.d., https://cloud.google.com/architecture/dr-scenarios-for-data.

[12] AICPA, "SOC 2 – SOC for Service Organizations: Trust Services Criteria," AICPA, n.d., https://us.aicpa.org/interestareas/frc/assuranceadvisoryservices/aicpasoc2report.

[13] ISO, "ISO/IEC 27001: Information Security Management," ISO, n.d., https://www.iso.org/isoiec-27001-information-security.html.