Google Cloud

Whitepaper
September 2021

Planning for the Worst:

Reliability, Resilience, Exit and Stressed Exit in Financial Services

Table of Contents

Introduction

The use of public Cloud by financial institutions creates a dependency on Cloud Service Providers (CSPs), and a risk that services will be disrupted if a CSP suffers a technical, commercial or compliance failure. Firms need to assess and address this risk on an ongoing basis.

Financial services regulators treat certain uses of public Cloud as a form of outsourcing, and require that firms address their dependency on CSPs through exit plans. These are plans which would allow firms to switch to alternative arrangements if required.

Certain financial services regulators have also recently published or proposed regulations which increase the obligations of firms to ensure that they have the right level of Operational Resilience: the ability to sustain their most important business services through severe but plausible disruption scenarios.

Relevant regulation and guidance includes:

- Recommendations on outsourcing to cloud service providers, published by the European Banking Authority (EBA) in 2017.
- Guidance on outsourcing to cloud service providers, published by the German Federal Financial Supervisory Authority (BaFin) in 2019
- Digital Operational Resilience Act (DORA) published in draft by the European Commission (EC) in 2020
- Outsourcing and third party risk management, published by the UK Prudential Regulation Authority (PRA) in 2021
- Operational resilience: Impact tolerances for important business services, published by the PRA in 2021.
- Sound Practices to Strengthen Operational Resilience, published by the US Federal Reserve System, Office of the Comptroller of the Currency and Federal Deposit Insurance Corporation in 2020
- Principles for Operational Resilience and Revisions to the Principles for the Sound Management of Operational Risk, published by the Basel Committee on Banking Supervision in 2021

In the UK, the PRA has introduced the concept of stressed exit as a form of exit plan to be invoked to achieve Operational Resilience in the event of a sudden, unexpected failure of an outsourcing provider.

This paper summarises Google's recommendations for the ways in which firms can address the risk of dependency on CSPs, and can address their requirements for exit plans and operational resilience.

Overview of Requirements

Although it is important for firms to refer directly to the regulations by which they are governed, the requirements of these different regulations can be summarised as follows:

To manage the risks of cloud outsourcing:

- Firms must assess and manage the risk associated with the use of CSPs, as with any outsourcing provider.
- Firms must identify which business activities they class as 'material'.
- Firms must establish plans to continue business activities in the event of termination of an agreement with a CSP.
- Firms must maintain records of business activities which have been outsourced and make these available to regulators when requested.
- Firms must identify those business activities whose disruption would threaten the stability of the organisation or cause significant customer or market harm.
- Firms must determine the maximum impact to those business activities that could be tolerated: this may be expressed in terms of time for which the activity cannot be performed, but may also be expressed in terms of service degradation, customers impacted, or other dimensions.
- Firms must develop mitigation strategies and plans to restore a viable version of those business activities before reaching that maximum impact.
- Firms must test these plans to ensure that they are viable.

While regulations acknowledge different ways in which public Cloud platforms and other technologies can be used to achieve these goals, they are not prescriptive in how those technologies should be used.

Failure Scenarios

It is important to distinguish between different failure scenarios, as these require different mitigation strategies and responses:

- **Technical failures:** failures of technology components, services, facilities (zones) or collections of facilities (regions). These failures can typically be addressed through the use of the resilience features within the platform of a single CSP and by designing applications to take advantage of those features: this approach is referred to as technical resilience throughout the rest of this paper.

- **Failures in service performance:** failure of a CSP to provide the quality of service expected by the firm. These failures can typically be addressed through remedial action by the CSP or, if this remedial action is not effective, exercise of an exit plan.

- **Commercial failure:** financial instability over time, or the failure of the commercial agreement between the firm and the CSP, resulting in some form of contract termination. These failures are typically addressed through exit plans.

- **Sudden commercial or compliance failure:** unexpected failure of a CSP through insolvency, compliance breach or some other commercial event (including events that effectively prohibit the provision, or consumption, of the service across geopolitical boundaries), within a timeframe too short for an exit plan to be executed. Such a failure must be addressed through stressed exit plans.

- **Catastrophic technical failure:** failure of a CSP's platform which is so widespread and with such limited prospects for recovery that it is necessary to recover services to an alternative platform. Such a failure must be addressed through stressed exit plans.

Given these different scenarios and the different responses each requires, we recommend that firms clearly distinguish between technical resilience, exit plans and stressed exit.

Technical Resilience

Technical failures should be addressed through technical means. On GCP, the technical means of providing resilience include:

- Multiple global facilities with redundant physical infrastructure (networking and utilities).
- Redundant hardware available by default.
- Distribution of services across zones: separate failure domains within the same or separate physical facility.
- Distribution of services across regions: groups of zones in physically separate locations, either different countries or geographically separated areas within the same country.
- Within selected regions, geo-separated zones provide a combination of physical distance and synchronous data replication for applications which need it, and which cannot benefit from multi-region deployment.
- Services which automatically replicate data and allocate hardware between zones and regions.

To make best use of these features, we recommend that firms:

- Define tiers which represent the level of impact caused by the potential disruption of each technology workload.
- Define technical standards which provide the required level of resilience for each tier.
- Distribute workloads and replicate data across zones.
- Distribute workloads and replicate data across regions.
- Distribute workloads and replicate data across geo-separated zones for specialist applications which cannot benefit from multi-region deployment.

By making use of these features, firms can achieve levels of technical resilience and distribute risk more widely than is possible within the traditional on-premise, twin data centre model.

More details on the architecture of our platform and how to use it to protect against outages are provided here.

Exit Plans

We recognise that, whatever level of technical resilience can be achieved on GCP, firms must nonetheless plan for a scenario in which GCP can no longer provide a service. For exit plans, we assume that the exit from GCP takes place over several months.

We support such exit plans through:

- Commitment to Open Source: many of our products and services are available in Open Source versions, meaning that they can be run on other Cloud providers or on-premise.
- Commitment to common standards: our platform supports common standards for hosting applications in virtual machines or containers, which can be replicated by alternative services on other Cloud providers or on-premise.
- Anthos multi-Cloud management: our multi-Cloud management product, Anthos, allows customers to run and manage an increasing range of services in the same way as on GCP across other Cloud providers or on-premise.

In order to benefit from these capabilities we recommend that customers:

- Define standards for exit planning, mapped to the importance of their business activities (for example, low impact business activities might only require a documented plan, while high impact services would require a plan subject to regular rehearsal/ testing.
- Define a multi-Cloud (involving multiple CSPs) or hybrid Cloud (involving GCP and on-premise facilities) strategy capable of providing the capacity or capability needed to support an exit plan.
- Define approaches to address services native to a Cloud platform (for example BigQuery on GCP) which must be migrated to an alternative service.
- Implement sufficient capabilities to make the multi-Cloud or hybrid Cloud strategy viable (for example, deploy production workloads on an alternative Cloud provider or on an internal platform - including security configurations, risk assessments and so on).

- Where warranted by the criticality of business activities, test the viability of the exit plan, rehearsing the migration of workloads to the alternative method of provision.
- Through testing, determine the total time required to execute an exit plan, and ensure that this time is reasonably likely to be available in an exit scenario.

We do not recommend that customers attempt to run the same workloads in production across multiple Cloud providers at the same time, except for the purposes of testing or parallel running: such cross-Cloud production operations increase complexity and cost, require additional organizational oversight, reduce the overall benefits of a cloud-based technology strategy, and potentially introduce new security risks and points of attack.

Stressed Exit

We understand that, however unlikely, firms must also plan for the scenario that a CSP will suddenly become unable to provide services, resulting in a stressed exit. Given the low likelihood of hyperscale platforms and providers failing, and the potential overhead and complexity of implementing stressed exit arrangements, we also understand that these arrangements are only required for the most important and critical business activities.

We support stressed exit in the same way that we support exit plans: through the use of Open Source; through commitment to common standards; and through our multi-Cloud management service, Anthos.

In order to plan for stressed exit, we recommend that customers:

- Define the criteria that would result in a stressed exit rather than an exit, for example:
  - Sudden commercial failure of the CSP to the extent that they were unable to continue operations and support the firm through the duration of an exit clause.
  - Catastrophic technical failure so extreme that executing a stressed exit is preferable to working with the CSP to recover the service.
- Identify the subset of business activities which it would be essential to preserve in the event of a stressed exit.
- As for exit planning, define a multi-Cloud or hybrid Cloud strategy.
- Also as for exit planning, implement sufficient capabilities to prove that the multi-Cloud or hybrid Cloud strategy is viable.
- Using a similar approach to scenario planning for Operational Resilience, define the core business activities that would need to be sustained through a stressed exit, identify the critical and important workloads which support those functions, and define those additional steps which must be taken to enable workloads to be migrated quickly, such as:
  - Enforcing technical standards which make it possible to run the service on an alternative platform.

    - Maintaining a current but dormant copy of the code and workload configuration on the alternative platform.

    - Replicating data to the alternative platform as well as to the live platform.

- Test and rehearse the stressed exit plan on a regular basis.

Concentration Risk

As the use of public Cloud becomes more prevalent across financial services, regulators have an increased focus on concentration risk: the risk that individual firms or the financial system are exposed to a concentration of dependencies on technologies, facilities or providers that may fail.

Similar to the responses to different kinds of failure, different types of concentration risk can be addressed as follows:

- **Dependence of a workload on a technology component or facility:** this can be addressed by distributing workload across zones and regions in line with our best practice and recommendations.

- **Dependence of a firm on a provider:** this can be addressed through exit plans and stressed exit plans.

- **Dependence of the industry on a group of providers:** this is an emerging question best addressed at industry level, by firms maintaining and providing registers of outsourcing arrangements as required by regulation, and by ongoing engagement and consultation with policymakers and regulators.

References

- EBA: Final draft Recommendations on Cloud Outsourcing (EBA-Rec-2017-03)
- BaFin: Guidance on outsourcing to cloud service providers
- EC: Proposal for Digital Operational Resilience Act
- PRA: Operational resilience: Impact tolerances for important business services - SS1/21
- PRA: Outsourcing and third party risk management - SS2/21
- Google: Strengthening Operational Resilience in Financial Services by Migrating to Google Cloud