Disaster Recovery Plan

|  |  |
| --- | --- |
| Owner |  |
| Approver |  |
| Last Approved |  |

**Scope and applicability**

The procedure applies to [Company Name]ʼs production IT infrastructure, which provides customer services. Other systems can be subject to business impact analysis

BIA which helps to determine mission/business systems and their components required for DR planning.

**Roles and responsibilities**

|  |  |
| --- | --- |
| **Team or role** | **Responsibilities** |
| Service owner | Determine the need and configuration of the application DRP; Execute DRP after its activation; Return the application to its normal operation. |
| Site Reliability team | Determine the need and configuration of fault-tolerant IT infrastructure, including deployment of an alternative site; Execute DRP after its activation; Return primary production infrastructure to its normal state. |
| Director of Engineering, Operations | Approve this procedure and separate DRPs for IT components of the production infrastructure; Manage overall DRP process; Analyze DRPs for improvements and lessons learned. |
| Security team | Review this procedure and other DRPs for compliance with the security requirements. |
| On-call engineer | 24/7/365 reacts to system outages; Informs service owners and other interested (impacted) parties about the system outage; Could |

|  |  |
| --- | --- |
|  | activate this or other system-specific DRP if the service owner is unavailable. |

**Definitions**

|  |  |
| --- | --- |
| **Term** | **Definition** |
| RPO | Recovery point in time, before a disruption or system outage, to which mission/business process data can be recovered (given the most recent backup copy) after an outage. It is a factor of how much data loss the mission/business process can tolerate during the recovery process. |
| RTO | Recovery time objective defines the maximum amount of time that system resources can remain unavailable before an unacceptable impact on other system resources or supported mission/business processes. |

**General companyʼs DRP**

# Requirements for the production availability

[Company Name] commits externally to an RTO of 72 hours. RPO parameter is not officially communicated to the customers.

# Monitoring and notification

Engineers monitor [Company Name]'s service performance via the Datadog tool. In case of service unavailability problems, they receive notifications in Slack. On-call Engineers also obtain an alert about critical failures on the phone via the Alerting tool.

Additional information about possible options and tools that could be used by engineers for gathering metrics at various levels within Cloud Hosting architecture is described in the following article.

An on-call engineer will analyze the issue or outage and inform other team members responsible for the damaged service or dependent service in Slack. The on-call engineer and the responsible team could activate the DRP after the outage analysis.

They also should notify the Marketing team to update [Company Name]ʼs product Status page after the outage analysis and when the system restoration is completed.

# Approach for infrastructure availability

As [Company Name] deploys its infrastructure on [Cloud Hosting Service], possible disasters mostly relate to the unavailability of different cloud resources. So that the scope of infrastructure loss is grouped into three scenarios: a zonal outage, a regional outage, or an outage of multiple regions.

Currently, [Company Name]ʼs services adopt DR solutions within the region and accept the possibility of [Cloud Hosting Service] regional and multiple regions outages. We plan our infrastructure by considering the combined RTO of the [Cloud Hosting Service] products the application utilizes and actions the engineers must take to restart application components.

More information about approaches for DRP of [Cloud Hosting Service] infrastructure and [Cloud Hosting Service] components' RTO abilities is in the following article.

# Approach for application availability

[Company Name] restores its application following the standard way to deploy a new version of the application.

The engineering teams can also use other DR scenarios for recoverability from a disaster if additional measures should be applied to protect their applications.

They can also speed up the application's start-up in the DR environment to meet its RTO requirements.

[Company Name] plans the recoverability of [Source Code Management System], the CI/CD tool, and its content. The application and infrastructure code could be restored from build history. Engineers could also develop and deploy code into production from their laptops in a disaster. Currently, [Company Name] shares the responsibility with the vendor for providing the [Source Code Management System] availability, but if necessary, we can adopt additional DR solutions for protecting the CI/CD tool as described in the vendor's official documentation.

In case of [Source Code Management System] unavailability or for implementation of restoration actions, [Company Name] allows direct changes to the production from the engineersʼ laptops. However, these changes are subject to change management process requirements (e.g.,

peer-review and testing) after the system restoration. Please see more requirements in the Change Management Policy.

# Approach for data availability

[Company Name]'s customer-facing applications have an RPO of an hour, so they take advantage of asynchronous replication. We accept small data loss in the context of zonal and regional outages.

[Company Name] backs up every 30 minutes of databases with customers' data, logging, and billing information to another [Cloud Hosting Service]region. Backups are automatically verified for recoverability. [Company Name] has verified backups for any required data restoration for the previous day, week, month, and quarter.

Monitoring of database backup creation and restoration is performed via the Datadog tool. There are two jobs for tracking:

 Tracker that monitors snapshots being taken

 Tracker that monitors the appropriate amount of snapshots have been taken

Other options for data backup required for [Company Name]ʼs applications can be found in the following article.

Certain databases and types of storage are difficult to scale and make resilient. To ensure that database choices do not constrain application availability and scalability, use the following article.

More information about backup in [Company Name] could be found in the Backu[p](https://getsentry.atlassian.net/wiki/spaces/COM/pages/2183593985)  management procedure.

# Disaster recovery site security

The following requirements are implemented to ensure the same security for disaster recovery sites and data backup storage as for the production:

 Network controls have the same separation and blocking that the source production environment. Shared VPC and firewalls are configured to establish centralized networking and security control of the deployment and subnets and control inbound and outbound traffic.

 IAM policies in the DR environment are applied using infrastructure as code

IAC methods.

 IAM policies are used to grant appropriate permissions to products. For example, access to specific cloud storage buckets is properly restricted.

 Logging from the DR environment is as for the production environment via logging and exported to the main log sink archive.

# Examples of [Company Name].io disaster scenarios

These are possible [Company Name].io disaster scenarios with options for disaster recovery. The teams could use them as a basis to define plans applicable to their systems.

|  |  |
| --- | --- |
| Scenario | DR Approach |
| [Cloud Hosting Service] Disruption | Multi-region [Company Name], or the ability to spin up a new [Company Name] as a service instance in a different region and direct traffic there. |
| [Cloud Hosting Service] Region outage | Impact – total customer outage. No ingestion, no dashboards, no [Company Name].io.  Status.[Company Name].io – should remain online as it is hosted by statuspage.io on AWS and is used for customer communications.  Data loss  Data in flight could be lost, and data persisted to disk but not backed up could also be lost. This is impossible to predict as it depends on what exactly happens at/to [Cloud Hosting Service].  Recovery – wait it out. The amount of work required to move regions is very high. |
| Partial/intermittent [Cloud Hosting Service] outage | Impact – unpredictable customer impact for ingestion and dashboards.  Status.[Company Name].io – remains online and used for communications with customers.  Data loss  Data that reaches us should be safe as the ingestion machines are Highly Available. If a machine that is actively working is shut down, there is a chance of data loss, |

|  |  |
| --- | --- |
|  | depending on which tier the machine is on.  Recovery – the most likely plan is to wait it out. Possibly stop ingestion to guarantee consistency, but that is up to stakeholders to decide and depends on the specific behavior of the outage. |
| SQL outage | Impact – cached data will be returned to customers when possible, ingestion will work but will cause high queues on rabbit/Kafka tiers, which could eventually stop responding. Dashboards are still available but showing potentially stale data.  [Company Name] Status – remains online and used for communications with customers  Data loss  Postgres machines have hourly snapshots of their data drives, so restoring from those snapshots and replaying the queues should result in very little to no data loss.  Recovery – one of the following:  Lost Primary SQL server - promote the active replica to primary status and point all producers at the new primary, build a new replica from a snapshot, and wire it up for replication to regain HA status.  Lost SQL HA Pair - build a new primary from terraform and most recent data drive snapshot and direct all traffic to it.  Build a new replica from [IAS tool] and most recent data drive snapshot and wire it into replication to regain HA status.  Lost SQL replica – build a new replica with terraform from the most recent data drive snapshot and wire it into replication, thus, regaining HA status. |

**Guidelines for creating individual system DRP**

# Contingency Planning and the System Development Life Cycle (SDLC)

The system owners create and maintain their IT system contingency plans throughout the system development life cycle. Below, the necessary steps for the planning process are described:

|  |  |
| --- | --- |
| **SDLC stage** | **Contingency planning actions** |
| Initiation | 1. Conduct risk assessment to understand what system protection against and confidentiality, integrity, and availability objectives should be set.   Evaluate the business processes that support the new information system to determine the systemsʼ recovery time requirements. |

|  |  |
| --- | --- |
| Development/ Acquisition | Determine contingency solutions to be incorporated into the system architecture to provide reliability, maintainability, and availability.  Examples of contingency measures that could be considered in this phase are:  •  Redundant communications paths;  •  Elimination of single points of failure;  •  Enhanced fault tolerance of network components and interfaces;  •  Load balancing;  •  Data mirroring and replication;  •  Alternate site deployment. |
| Implementation/ Assessment | Document recovery strategy in the systemʼs DRP. As the system undergoes initial testing, contingency strategies also should be exercised to resolve any issues with the procedures. |
| Operation/ Maintenance | 1. Maintain a test and training to validate DRP on a scheduled basis to ensure that procedures continue to be effective. Update DRP to reflect changes to procedures based on lessons learned from tests, exercises, and actual disruptions.   Conduct and periodically validate backups. 3.  When the information system undergoes upgrades or other modifications, these modifications should be reflected in the contingency plan. |
| Disposal | 1. Maintain the original systemʼs DRP until the new system is operational and fully tested (including its contingency capabilities).   Consider using a legacy system as a redundant system if a loss or failure of the new information system occurs. |

# Determine recovery requirements

The first step is to define the availability requirements for the application. It is based on Recovery Time Objective RTO and Recovery Point Objective RPO parameters.

**To define the systemʼs RPO** (“How much data can I afford to lose in the event of a disasterˮ), use the following questions:

 How much data do we project to lose should operations be interrupted?  What is the maximum amount of data loss we can tolerate?

 What is the cost of lost data?

 What is the cost to reenter lost data?

 How much will it cost to implement a solution that can meet our requirements?

**To define the systemʼs RTO** (“How long after a disaster before I'm up and runningˮ), use the following questions:

 How much revenue do we project to lose if this system is inaccessible?

 Does this system handle customer data? If yes, what SLAs are in place with customers?

 If X system went offline, does it have dependencies? What other systems would be impacted? What are the RTOs for those systems?

 What customer-facing systems or applications do we have that would result in loss, churn, or customer dissatisfaction if unavailable?

 How much will it cost to implement a solution that can meet our requirements? As a result of answering the above questions, define the system criticality level:

|  |  |  |  |
| --- | --- | --- | --- |
| **Application criticality** | **% of Apps** | **Examples apps** | **RTO & RPO**  **values** |
| Tier 1(most important) | 5% | Global or external customer-facing applications such as real-time payments and eCommerce storefronts. | RPO 4 hours  RTO 72 hours |
| Tier 2 | 35% | Regional or important internal applications such as CRM or ERP. | Up to 1 day |

|  |  |  |  |
| --- | --- | --- | --- |
| Tier 3(least important) | 60% | Team or departmental applications, such as back-office, leave booking, internal travel, accounting, and HR. | More than 1 day |

These levels are important for introducing unified approaches and clear priorities in systems recovery during a disaster. Indicate system level, RTO, and RPO at the top of the DRP document.

# Choose disaster recovery strategy

Disaster recovery DR strategies are considered cold, warm, or hot. These strategies indicate how readily the system is for recovery when something goes wrong, depending on how prepared recovery conditions are.

The system owner can define DR strategy based on the system tier identified in the previous stage.

|  |  |  |
| --- | --- | --- |
| DR strategy | Description | Applicable system tier |
| Cold | In a cold pattern, minimal resources are required for a recovery scenario. For example, it could mirror the production environment in another [Cloud Hosting Service] region deployed after the disaster using standard infrastructure and application deployment tools and available backups. | Tier 3 |
| Warm | Most applications will have an RTO of between an hour and a day, which allows for a warm failover in a disaster scenario, with some components of the application running all the time in a standby mode - such as databases - while others are scaled out in the event of an actual disaster, such as web servers. For these applications, | Tier 1, Tier 2 or Tier 3 |

|  |  |  |
| --- | --- | --- |
|  | automation for the scale-out events should be considered. |  |
| Hot | Achieving RPO zero means using predominantly regional products for the application, which by default are resilient to zone-scale, but not region-scale, outages. It also means designing an application that recovers automatically from a disaster without human intervention or with minimal steps such as pushing a button to failover. | Tier 1 |

# Create and deploy DRP

The DRP should include the following recommended stages.

## Activation and notification stage

Requirements in this stage define initial actions taken once a system disruption or outage has been detected or appears imminent.

Plan activation criteria for system outages or disruptions should be defined, and if they are met, the designated authority (e.g., on-call engineer) should activate the plan.

Criteria may be based on:

 The extent of any damage to the system (e.g., physical, operational, or cost);

 Criticality of the system to the organizationʼs mission (e.g., critical infrastructure protection asset);

 The expected duration of the outage lasted longer than the RTO.

This stage also includes activities to notify recovery personnel about the outage or disruption. The recovery personnel could include members of the engineering team responsible for system development and maintenance. People who can help and are responsible in the DRP for the system recovery form a recovery team.

The recovery team assesses the nature and extent of the disruption.

The outage assessment should be completed as quickly as the given conditions permit.

Outage assessment procedures may be unique for the particular system, but the following minimum areas should be addressed:

 Cause of the outage or disruption;

 Potential for additional disruptions or damage;

 Inventory and functional status of system resources;  Type of damage to system resources or data;

 Dependent systems that are also impacted;  Items to be replaced;

 Estimated time to restore normal services.

Once the impact on the system has been determined, other applicable engineering teams should be notified of updated information and the planned response to the situation. It is also necessary to notify the Marketing team responsible for updating [Company Name]ʼs public Status page for the customers.

## Recovery Phase

Recovery Phase activities focus on implementing recovery strategies to restore system capabilities, repair damage, and resume operational capabilities at the original or new alternate location.

After the Recovery Phase, the information system will be functional and capable of performing the functions identified in the plan. Depending on the recovery strategies defined in the plan, these functions could include temporary manual processing, recovery, and operation at an alternate system or relocation and recovery at an alternate site.

## Reconstitution Phase

The Reconstitution Phase is the third and final phase of the DRP implementation and defines the actions taken to test and validate system capability and functionality.

During Reconstitution, recovery activities are completed, and normal system operations are resumed. If the original facility is unrecoverable, the activities in this phase can also be applied to preparing a new permanent location to support system processing requirements.

This phase consists of two major activities: **validating successful recovery** and

**deactivation the plan**.

**Validation** of recovery typically includes these steps:

 **Concurrent Processing** is the process of running a system at two separate locations concurrently until there is assurance that the recovered system is operating correctly and securely.

 **Validation Data Testing** is the process of testing and validating recovered data to ensure that data files or databases have been recovered completely and are current to the last available backup.

 **Validation Functionality Testing** is a process verifying that all system functionality has been tested and the system is ready to return to normal operations. At the successful completion of the validation testing, the recovery team will declare that reconstitution efforts are complete and that the system is operating normally.

Deactivation of the plan is returning the system to normal operations and finalizing reconstitution activities to prepare the system against another outage or disruption.

**Deactivation** activities include:

 **Notifications**. Upon return to normal operations, users should be notified by the recovery team using predefined notification procedures (if possible).

 **Cleanup** is cleaning up a workspace or dismantling any temporary recovery locations, restocking supplies, returning manuals or other documentation to their original locations, and readying the system for another contingency event.

 **Data Backup**. As soon as reasonable following reconstitution, the system should be fully backed up and a new copy of the current operating system stored for future recovery efforts.

 **Event Documentation**. All recovery and reconstitution events should be well documented, including actions taken and problems encountered during the recovery and reconstitution efforts. An after-action report with lessons learned should be documented and included for updating the DRP. Once all activities and steps have been completed and documentation has been updated, the DRP can be formally deactivated. An announcement with the declaration should be sent to all business and technical contacts.

# Implement system monitoring

Implement controls to prevent disasters and detect issues before they occur.

For example, add a monitor that sends an alert when a data-destructive flow, such as a deletion pipeline, exhibits unexpected spikes or other unusual activity. This monitor could terminate the pipeline processes if a certain deletion threshold is reached, preventing a catastrophic situation.

Use company-available for setting up monitoring of the system operations.

# Periodically test DRP

After creating a DR plan, test it regularly, noting any issues and adjusting the plan accordingly.

|  |  |
| --- | --- |
| **System criticality** | **Required test exercises** |
| Low-impact systems  Tier 3 | N/A |
| Moderate-impact systems Tier 2 | N/A |
| High-impact systems  Tier 1 | A full-scale functional exercise should be conducted annually or after significant changes to the system. The full-scale functional exercise should include a system failover to the alternate location. This could include additional activities such as full notification and response of key personnel to the recovery location, recovery of a server or database from backup, and processing from a server at an alternate location. The test should also include a full recovery and reconstitution of the information system to a known state. |

**Tabletop exercises** are discussion-based exercises where personnel meet in a classroom setting or in breakout groups to discuss their roles during an emergency and their responses to a particular emergency. A facilitator presents a scenario and asks the exercise participants questions related to the scenario, which initiates a discussion among the participants on roles, responsibilities, coordination, and decision-making. A tabletop exercise is discussion-based and does not involve deploying equipment or other resources.

**Functional exercises** allow personnel to validate their operational readiness for emergencies by performing their duties in a simulated operational environment. Functional exercises are designed to exercise the roles and responsibilities of specific team members, procedures, and assets involved in one or more functional aspects of a plan (e.g., communications, emergency notifications, system setup). Functional exercises vary in complexity and scope, from validating specific aspects of a plan to full-scale exercises that address all plan elements.

Functional exercises allow staff to execute their roles and responsibilities as they would in an actual emergency, but in a simulated manner.

For each DRP test activity conducted, results should be documented in an after- action report in a ticket, and Lessons Learned corrective actions captured for updating information on the plan.

In addition to the testing noted earlier, it is recommended to periodically test:

 Permissions and user access work in the DR environment as they do in the production environment.

 Perform penetration testing on your DR environment.

 Perform a test where a usual access path to [Cloud Hosting Service] doesn't work.

 Make sure that while a DR environment is in service, any logs that are backfilled into the log archive of the production environment.

 Ensure the recovery process is auditable, namely, who had access to the backup data and who performed the recovery.

 Ensure users can log in to the DR environment and have appropriate access rights.

# Periodically review DRP

It is essential that the DRP be reviewed and updated annually or whenever significant changes occur to any plan element to ensure that new information is documented and contingency measures are revised if required.

**Revision history**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Description** | **Approved by** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |