databricks

Disaster Recovery Impact Assessment

This assessment is for any customer, partner or professional services organization that has been tasked with implementing a Disaster Recovery (DR) solution.

Although some of the knowledge will become an important input into the decision-making process for a DR strategy and implementation pattern, the rest will serve as an analysis of the ability to maintain the DR solution and onboard new applications onto it.

This assessment focuses on questions at the workload level. However, they are equally valid for applications, use cases and services.

Please contact your Databricks representative to understand what assistance is available for understanding and implementing DR for a Databricks workspace.

Business Priorities

1 What is the main purpose of implementing a DR solution?

2 What risks exist if mission-critical workloads are unavailable for an hour, a day, a week, a month or longer?

a. What are the financial costs for every hour, day, week, month, etc. it is unavailable?

- **Business costs:** The loss of revenue that would have been made if fully operational.
- **Productivity costs:** The amount paid to employees without any productivity in return.
- **Direct costs:** Expenses incurred to return to full operational capacity.
- **Other financial costs:** Additional expenses that are required to keep operating after the disaster but do not contribute to recovering capacity. For example, fines and/or penalties related to a specific industry would be considered other costs.

b. What are the nonfinancial costs? As there is no clear formula to determine the cost for most nonfinancial scenarios, an ordinal scale of 1–5 is recommended, where 1 is the most costly.

- **Reputation costs:** Although no specific formula exists to calculate reputation loss, it can range from a negligible difference to a total lack of trust from customers and the general public.
- **Other nonfinancial costs:** Delayed projects, shifted priorities, etc.

c. Has a Recovery Time Objective (RTO) been defined in terms of measurable time (minutes, hours or days)?

3 Which data assets are mission-critical, meaning that function cannot be sustained without this data, to keep the workload and downstream business units operating?

a. What are all the services needed to use these data assets?

4 How much potential data loss or inconsistency, can the business tolerate?

a. Can a data asset be recreated from the upstream source?

- Would those sources be made available as part of the DR solution?

b. Has a Recovery Point Objective (RPO) been defined in terms of measurable time (minutes, hours or days)?

5 Which business units (i.e., a division/department) are downstream consumers of the workload?

a. Do these business units have varying degrees of tolerance?

b. Do the business units have their own workspaces for development, staging and production?

6 Is there an agreed-upon, measurable definition for a Disaster Event that would trigger the DR solution?

a. Which systems/services are mission-critical to the business?

b. How quickly does a response need to occur (trigger the DR solution) when a Databricks service (i.e., create a cluster) or required infrastructure (i.e., provision a VM for the cluster) is unavailable?

7 Is resilience to outages at the availability zone or region level required?

8 Is there a plan for what should be done if the DR site is also down?

a. Is more than one Disaster Recovery site needed?

Technology Capabilities

1 Is an inventory of any existing DR assets available? If not, can one be made?

2 Is a monitoring tool, such as Datadog, in place that can be used to detect and monitor the progress of a Disaster Event?

a. What metrics need to be monitored as part of the DR solution?

3 Is an alerting system, such as PagerDuty, available to notify stakeholders, decision-makers and executors of the DR solution?

4 Is an Infrastructure as Code (IaC) tool, such as Terraform, available?

a. Databricks has a provider for Terraform. If the IaC tool is not Terraform, can Terraform be used for the DR solution?

5 Do existing personnel have sufficient experience and time to implement and maintain the DR solution, including developing and supporting code to implement the solution?

a. The minimum skill set required in the case of a Databricks workspace is:

- Cloud security — IAM, roles, policies
- Terraform
- Monitoring — Datadog, audit logging (AWS | Azure | GCP)
- Delta Time Travel
- Delta DEEP CLONE
- Python and/or Scala

6 Are there CI/CD processes in place to deploy code and configurations to the various sites?

7 Does documentation exist detailing the cloud services, configuration and artifacts that are part of the workload structure?

8 How will the team responsible for the DR solution communicate with business and technical resources?

9 For scheduled batch processing, does processing need to resume at the point of the disaster event or can the whole pipeline be re-run without overlapping with the next period?

10 For stream processing:

a. Is the data source a single point of failure?

b. Is idempotency built into the processing logic?

c. How is missing data detected?

d. In the scenario of an active-active deployment, how will the streams be kept in sync?

11 How will an end-of-day reconciliation between the primary and DR sites be implemented?

a. How will cut-over between sites be communicated and handled?

b. Can the DR site be synchronized and reconciled with the primary site when the primary site returns back to normal operation?

12 Are existing data processes decoupled into isolated units?

a. Are table definitions, orchestration and transformation logic contained within the same object (source code file) or defined in distinct objects (one file per process)?

Technology Limitations

1 What is the subscription level (AWS | Azure | GCP) of the Databricks workspace?

2 What existing assets and tools can be leveraged in the implementation?

a. Is there a current, in-place DR strategy or mitigation plan that can be used as a starting point for the current state?

3 Have the financial costs for implementing a DR solution been quantified and assessed against the risks of a disaster-caused outage?

a. **People costs:** The amount paid to employees for time spent creating the DR strategy, designing the solution, and for the actual implementation. In addition, contracting of additional persons required and training for existing employees.

b. **IT costs:** Costs to provision and maintain the infrastructure at the DR site, as well as tooling (potentially licensing, additional cloud provider services, etc.) required to switch between primary and DR sites. When selecting a DR site, also commonly referred to as a secondary site, factors such as data transfer requirements, data volume, service availability, and scope of the impact of a disaster (Availability Zone vs. Regional) should be considered.

c. **Maintenance and ongoing costs:** DR cannot be a one-off exercise. The whole strategy and implementation must be maintained and updated. Updates can include revisions based on lessons learned or new tooling that becomes available in the market.

d. **Other costs:** Any additional one-time or ongoing costs.

4 What is the budget for setting up and maintaining a DR site?

a. Is an active/active strategy required or could an active/passive strategy be used?

b. How much data staleness is permissible based on the RPO?

c. Can the DR site run a subset of the production processes?

5 Is the full universe of data in scope or can critical data assets be identified?

6 Are there any contractual SLA agreements that should be considered when designing the solution?

7 What are the security requirements for data at rest and in transit as part of company policy, government regulations and/or contractual obligations?

a. Are there additional requirements for data tagged as Personal Identifiable Information (PII)?

b. Will the tools and processes need to be HIPAA and/or PCI compliant?

c. Does the data need to be encrypted for egress/ingress?

8 In case of an active-passive deployment, will the DR site remain active and become the primary site after a disaster event? If not, will there be a dedicated active site to which there will be a cut-over when it is available again?

9 Has an inventory of data assets been gathered, including storage service (Object storage, OLAP or OLTP system, etc.), estimated size of the data, schema, and partitioning strategy?

Other Important Considerations

1 Are there users on the system who have to be notified before, during or after a fail-over to the DR site?

a. Will those same users need to be notified if failing back to the primary site, when available, from the DR site?

2 What is the role of individual users on the system?

a. Do those roles change in a development, staging or production environment?

b. Do the users own and create their own Databricks objects (clusters, SQL endpoint, jobs, notebooks, secrets, etc.)?

c. Will the users be connecting to Databricks from an external system? For example, a BI tool that connects to a SQL endpoint.

d. Do users invoke the Databricks REST API (AWS | Azure | GCP) using their own Personal Access Token(s) (AWS | Azure | GCP)?

e. Do users create, update, maintain and/or delete their own databases and tables?

3 Is a comprehensive list of stakeholders, decision-makers and executors for DR compiled?

4 What is the current dev-loop for Databricks users?

5 Are there any developers using Databricks Connect (AWS | Azure | GCP)?

6 Is it possible to use the Databricks Repos (AWS | Azure | GCP) feature to integrate with a git service provider, such as GitHub?

7 Are there any upstream or downstream systems with hard-coded URLs?

8 Can some of the pipelines be redesigned to be DR-site aware? This is a broad question since DR generally includes site awareness, site statuses, site monitoring and the automation of a cut-over.

9 Do any workloads (interactive or automated) contain an on-premises system as part of a contiguous process for which Databricks is also contained?

10 Is there a list of specific scenarios that a DR solution must handle?

e. Within this list, are compelling edge cases included?