Data Protection and Recovery Solutions in Spanner

Winston Douglas, Lakshmi Palikila

Google Cloud

White paper
March 2024

Table of Contents

Table of Contents

Google Cloud

Introduction

The purpose of this white paper is to provide a guide to Google Cloud technical practitioners on configuring Spanner for data protection and recovery to meet backup and recovery, business continuity, and compliance needs.

Spanner is a highly scalable database that combines unlimited scalability with relational semantics, such as strong consistency, secondary indexes, schemas, and SQL providing 99.999% availability. It removes the stress from managing databases by providing zero-touch maintenance.

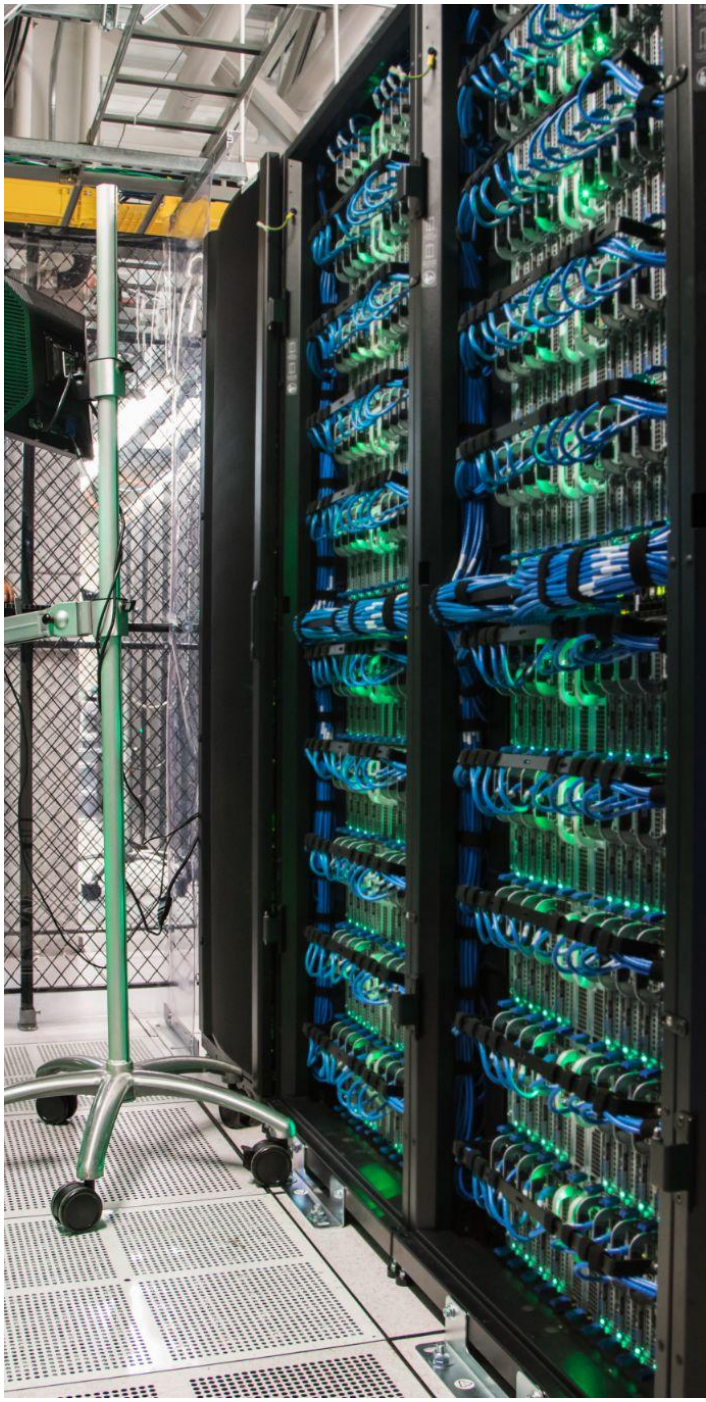
Replication and availability in Spanner is fully managed. Spanner automatically replicates your data across multiple failure domains and with multi-region configurations your data is replicated across multiple geographic regions all while providing strong consistency for transactions. Therefore, customers can easily address the tough challenges of handling regional failures and meet the most demanding Disaster Recovery (DR) requirements.

This whitepaper will cover all the features of Spanner that you can choose from to meet your data resilience requirements and weave into your enterprise strategy. It will highlight native and zero touch data protection features and how to configure backup and recovery features according to your needs. WIth Point-in-Time-Recovery , Backup and Restore and other data protection features , Spanner meets the needs of customers to protect their data with minimal effort and low cost.

Google Cloud

Chapter 1

Data Protection and Availability



What is Data Protection ?

Data protection is the practice of safeguarding important information from unauthorized access, misuse, corruption, or loss. It involves a combination of strategies like encryption, strict access controls, regular backups, and even employee training. The goal of data protection is to ensure the privacy, confidentiality, and integrity of sensitive data, whether it's personal information belonging to customers, financial records of a company, or intellectual property. Data protection is crucial for both ethical reasons and to comply with regulations like the GDPR and CCPA. One important area of data protection is availability and ensuring that your data is always available when needed.

Google has engineered the Spanner database service to provide the high availability that customers require for business critical applications and it provides the features that are needed for you to protect your data according to your business needs.

Data Protection Focus Areas

There are many topics that are covered in the practice of Data Protection . The main focus areas include:

Data Security: This focuses on the integrity of your data and protects it from malicious or accidental destruction by external or internal actors.
Access Control: Ensuring that users accessing data are who they say they are and that they have access to only the data they are authorized to.
Data Availability: The data is always available to support business operations and can be quickly restored in the event of loss or corruption. **Backup and Recovery** is an important pillar that this stands on.

Google Cloud

This document will address key concepts within the realm of data availability. The primary focus will be on backup and recovery strategies, alongside in-region and data center availability, replication, and disaster recovery. Additionally, it will touch upon data security and access control topics, such as encryption.

Backup and Recovery

Backup and recovery is your safety net against data disasters. Think of it as insurance for your valuable data, databases, and systems. By making a small investment upfront you safeguard against significant loss in value to your business and even the business itself. By regularly creating and securely storing copies of your information, you proactively defend against the devastating impact of infrastructure failures, cyberattacks, human error, or natural disasters. A well-planned backup and recovery strategy is the foundation of a strong disaster recovery (DR) plan, it guarantees you can quickly get your business back on its feet when the unexpected strikes.

Think of it as insurance for your valuable data

A well-planned backup and recovery strategy guarantees you can quickly get back on your feet when the unexpected strikes.

Google Cloud

Why Is Data Protection Important?

Minimizes Downtime: Data protection enables businesses to minimize downtime and maintain operations, even in the face of disruptions. This translates to reduced financial losses and preserved productivity.

Protects Reputation: Data breaches and extended downtime can severely damage a company's reputation. Data protection helps protect your brand and maintain customer trust.

Ensures Compliance: In many industries, there are strict data protection and recovery regulations. Data protection strategies help businesses comply with these requirements and avoid fines or penalties.

Fosters Competitive Advantage: Data protection provides a competitive edge by ensuring that your business is always prepared and able to respond quickly to disruptions, positioning you ahead of competitors who may be less prepared.

Google Cloud

Data Protection in the Cloud

Data protection in the cloud hinges on strategies that transcend traditional on-premises approaches. Google Cloud offers built-in redundancy, with data stored across multiple locations, minimizing the impact of localized hardware failures or disasters. Proactive internal backup systems, and versioning further protect data integrity. Additionally, the cloud's scalability and availability of resources enables rapid restoration if needed. A strong data resilience plan in the cloud means embracing these tools alongside security measures to combat the ever-present risk of cyber attacks.

Fully managed services like Spanner offer robust, self-service database protection features, minimizing maintenance effort and delivering significant cost savings.

By simply adopting database services available in the cloud customers can meet most of their requirements and obligations towards protecting their valuable data. However, this is a shared responsibility and customers also need to configure some resources to ensure their business needs are met.

Prioritize Your Data Protection Strategy

Data protection is not an option, but a necessity in the modern business landscape. By investing in this critical area, you safeguard your organization's most valuable asset – data – and ensure the continuity of your business operations.

Google Cloud offers built-in redundancy

Data protection is not an option, but a necessity in the modern business landscape

Google Cloud

Chapter 2

Requirements

The following is a high level breakdown of the core requirements of customers when building a data protection and backup & recovery strategy:

Core Requirements:

Recovery Time Objective (RTO): Define how fast critical applications and data must be restored after disruption. Short timeline (minutes or hours) for mission-critical applications.
Recovery Point Objective (RPO): The acceptable data loss measured in time, often needing near real-time to ensure minimal loss.
Availability: Customers will demand high availability guarantees, often expressed as a percentage like "99.999%" (five nines) uptime, allowing only minimal downtime per year.
Reliability: The data resilience solution itself needs to be highly reliable. This means consistent backups, minimal failures, and thoroughly tested recovery procedures.
Security: Cybersecurity protection is essential. It includes data encryption, strong access controls, and measures against insider risks, unauthorized access or modification.
Compliance: Strict data regulations exist in some industries. Adhering to standards like HIPAA, PCI DSS, or GDPR is essential for customers' solutions.
Industry Best Practice: A sound backup strategy starts with following industry best practices, including:
a) Using cloud backup technologies such as Spanner backup and restore
b) Scheduling and automating the process
c) Regular backups
d) Adopting the 3-2-1 backup rule
e) Monitoring and alerting for failures and errors
f) Cataloging and logging of backup execution
g) Encrypting backups and restricting access with proper access controls
h) Regular testing of backups as well as restore workflows.

Chapter 3

Failures and Risks

Companies should perform their own risk assessment to identify risks for database protection. They would thoroughly analyze their databases to identify potential failure points, vulnerabilities, and the impact of various disaster scenarios. This assessment would provide critical insights to help them develop proactive strategies to protect their data, minimize downtime, and ensure business continuity in the event of failures.

Unplanned database downtime can be incredibly costly for businesses. It can lead to lost revenue, damaged customer relationships, productivity disruptions, and even legal repercussions in some cases.**Companies that can attach a specific monetary value to data loss and downtime related to identified risks are normally more successful in prioritizing and operationalizing the most suitable data protection and backup & recovery solutions.**

The following failure types and risks will be covered in the next sections::

- **Zonal Failure**
- **Regional Failure**
- **Logical corruption**
- **Major geographical disaster**
- **Unintended database removal**
- **Unauthorized access to backup media**
- **Long term data retention (Legal/Compliance risks)**

Note that these are some of the likely types of failures. Sometimes failures occurs that are “unknown” until they do occur. Each application and databases should be assessed separately to determine what the potential risks are.

The next section describes these failure modes and how Spanner can easily mitigate the risks either natively or by configuring available features.

Unplanned database downtime can be incredibly costly for businesses.

Each application and databases should be assessed separately to determine what the potential risks are.

Google Cloud

Chapter 4

Spanner Data Protection Features

Spanner offers several features designed to give customers the option to backup their data and to protect against different types of failures depending on their needs. The following is an overview of some of these features:

Regional configurations for HA: Spanner maintains replicas of your instances compute and data in separate failure domains (zones) within a single region. If a single zone fails the Spanner instance continues to provide service without any downtime.
Multi-regional configurations for HA+DR: Similar to regional configuration but multiple replicas are maintained across more than one region. If a single region fails, Spanner continues to provide service without any downtime.
Point-in-time Recovery: Spanner lets you retain all the versions of your data up to a certain configured retention period. This lets you recover from accidental data loss or corruption for up to 7 days.
Database backup and restore: Spanner makes a consistent copy your database to highly available and redundant storage separate from your database. The backup can be restored to a new database at a later date.
Cross-region copy: This allows you to copy your backups to different region to meet DR or compliance requirements.
Export and import: Export your data to Cloud Storage in CSV or Avro file format. The data can be imported into any Spanner database or other service.
Database deletion protection: This allows you to protect your database from accidental deletion due to operator or automation errors.

Google Cloud

Failure modes and Data protection (Quick Glance)

Failure mode	Spanner Feature Recommendation
Single zone failure	No additional action required, spanner provides this out-of-box.
Regional failure	Multi-region configs OR Regional config + Cross-region Backups
Logical corruption	PITR Managed Backups
Accidental database deletion	Drop database protection
Longer retention backups OR **Offline backups (outside spanner system)**	Export + Import. Export using Databoost to avoid impact to existing workloads.
Security	CMEK + IAM

Google Cloud

Chapter 5

Protect with Spanner

Zonal Failure

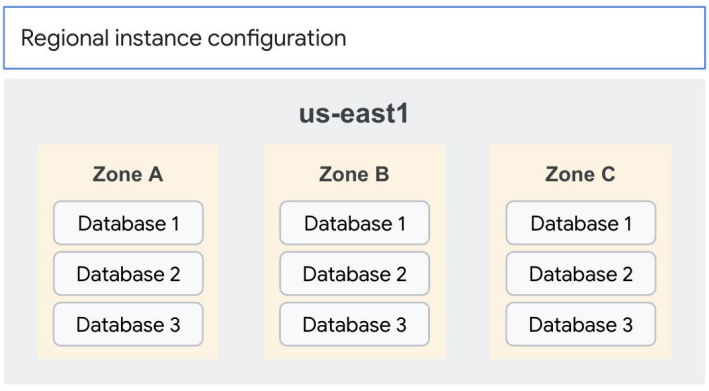
A zone is a deployment area within a region. Each zone is designed to be an independent single failure domain. To deploy fault-tolerant applications within a region with high availability and help protect against unexpected failure of a zone, create your database in instances with regional or multi-regional configuration.

Regional configuration

Spanner instances can be either **regional** or **multi-region configurations**. A regional instance maintain replicas of your database across three separate zones within the same geographical region as shown in the diagram below. Each zone (replica) has a full complement of compute resources for the instance and a full copy of your data.

A single zonal failure will not affect the availability of your Spanner database in a regional configuration. Spanner in a regional configuration provides **99.99%** availability. This blog details Spanner's handling of zonal and regional failures.

Regional configurations provide **zero RTO** and **zero RPO** for failure of a single zone.



Spanner in a regional configuration provides 99.99% availability.

A regional instance maintain replicas of your database across three separate zones within the same geographical region.

Recommendations:

- Choosing to run an application on Spanner automatically provides protection against a single zone failure and 99.99% or higher availability.
- Chose regional configurations for applications that require 99.99% availability and do not require protection against regional failures.

Google Cloud

Spanner provides you with 99.999% availability in a multi-region configuration.

Regional Failure

A region is an independent geographical area consisting of multiple zones. To deploy fault-tolerant applications with high availability and help protect against unexpected failure of a region, deploy your applications across multiple regions.

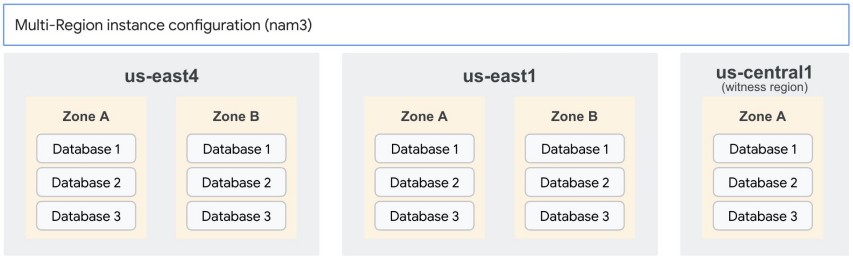
Multi-Region configuration

Spanner **Multi-region** instance configurations maintain replicas of your database in multiple zones across more than one region as shown in the example below of the nam3 multi-region configuration. In this configuration a single zonal or a single regional failure will not affect the availability of your Spanner database. Spanner provides you with **99.999%** availability in a multi-region configuration.

Multi-Region configurations provide **zero RTO** and **zero RPO** in the event of a regional failure (or failure of a zone).

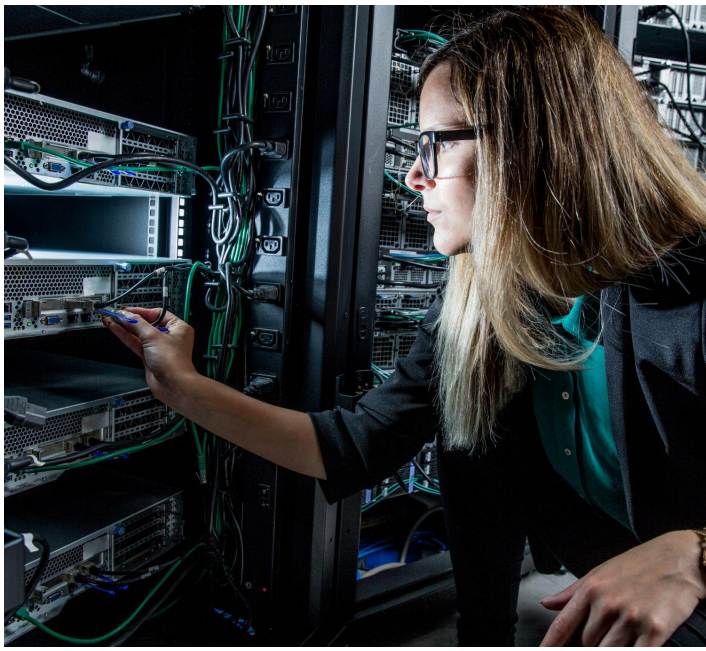
Multi-region configurations incurs higher latency and costs but meets the needs of many applications that require the higher availability and zero RTO and RPO in the case of a regional outage.

The Google Cloud blog "Demystifying Cloud Spanner multi-region configurations" provides a detailed technical description of Spanner's multi-region configurations.



Recommendations:

- Use one of the available Spanner multi-region configurations to protect against a regional failure. This will automatically provide built-in DR for the database and 99.999% availability.
- *Applications that require some protection against regional failures (DR) but can operate with a higher RTO and RPO (ie > 0),* ***consider copying backups to a different region****. (this will be covered later). Also* ***consider exporting data to Cloud storage*** *in another region or by using a multi-region bucket.*



Logical corruption

A database can become logically corrupt if an operator or application accidentally deletes from or writes to the database . This risk is often higher when application changes or updates are rolled out or if the application contains a bug.

For rapid recovery from database failures such as this, you need the ability to swiftly restore your data to the most recent valid version.

Spanner point-in-time recovery (PITR)

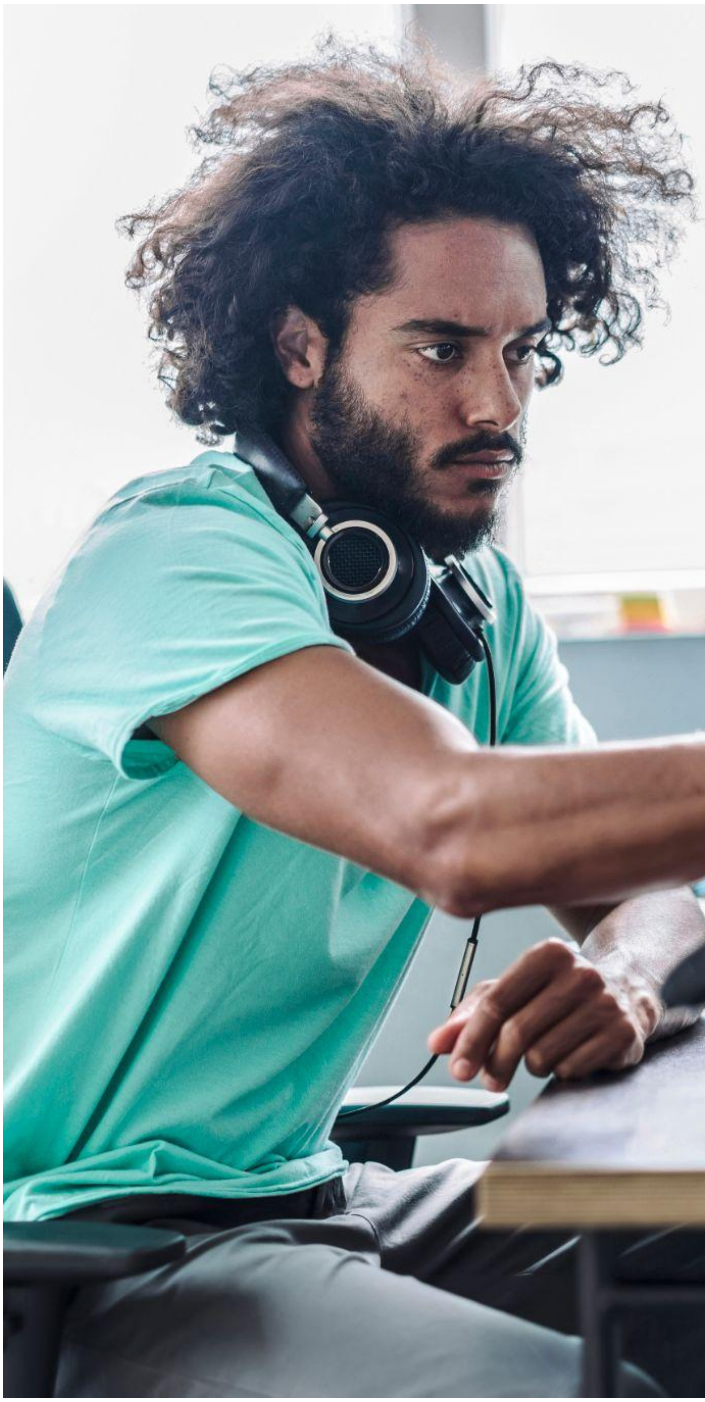
Spanner point-in-time recovery (PITR) provides protection against accidental deletion or writes. For example, if an operator inadvertently writes data or an application rollout corrupts the database, with PITR you can recover the data from a point-in-time in the past (up to a maximum of seven days) seamlessly.

With PITR, you can recover to any point with microsecond precision (within the last 7 days). However, RTO in the event of a logical corruption will depend on the extent and nature of the corruption. The restore operation delivers rapid time-to-first-byte results because the database accesses the backup directly, bypassing the need for data copying. However, in the event that only a small subset of data or a few rows need to be restored it will require a longer manual process and hence longer RTO.

you need the ability to swiftly restore your data to the most recent valid version.

Spanner point-in-time recovery (PITR) provides protection against accidental deletion or writes

Google Cloud



Managed Backup and Restore

Spanner's **managed backup and restore** offers seamless, on-demand database protection with just a few clicks. The backup creates a transactionally consistent copy of the database. You can then restore from the backups as needed in the event of an operator error, logical corruption, or major disaster. The backups are replicated and encrypted and can be retained for up to 1 year. Additionally backups can be copied to other regions and GCP projects.

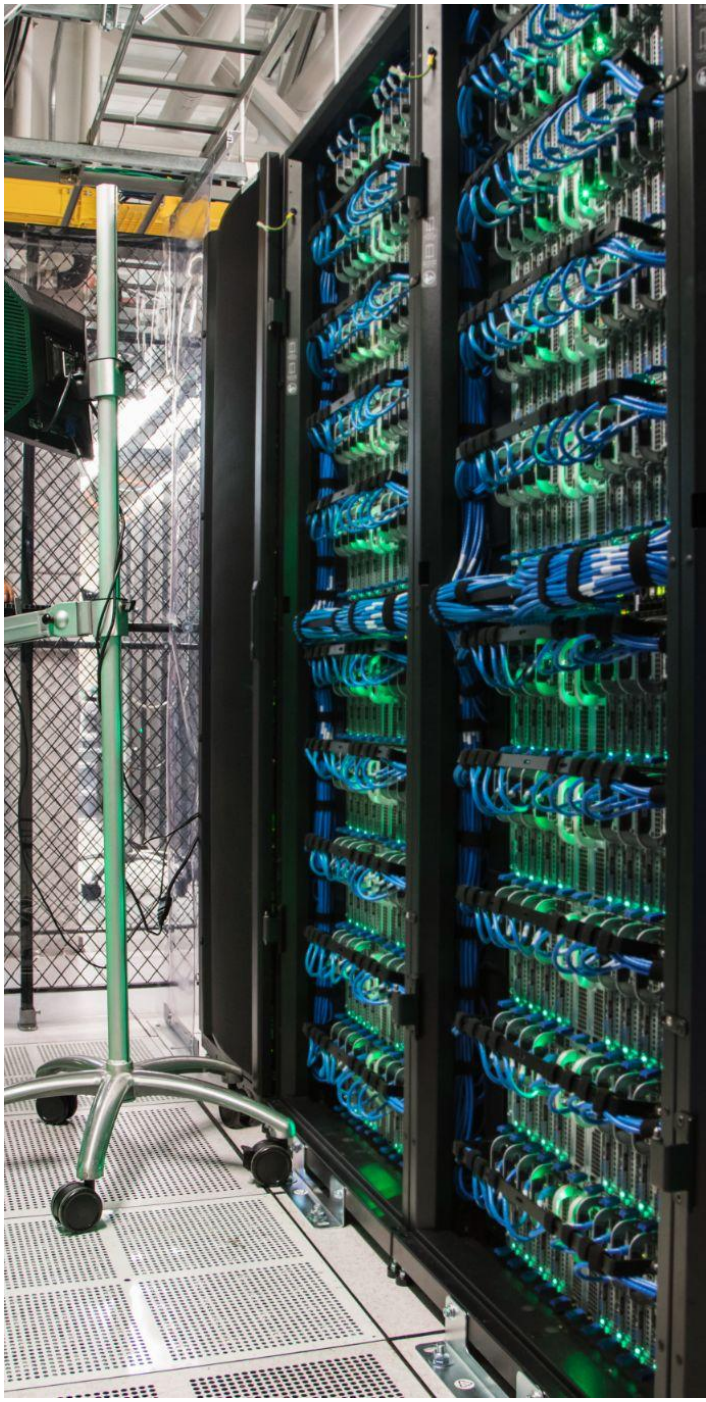
There is an additional cost for storing backups. More frequent backups offer better protection but increases storage cost.

In cases of major database corruption, restoring from a backup can lead to a fast recovery time (**RTO**). Restore operation is often completed within an hour, ensuring minimal disruption. More details around the restore process can be found here.

The **RPO** will depend on the frequency of backups and when the database failure occurs. More frequent backups will provide potentially lower RPO. As mentioned above this will also incur a higher storage cost.

This also adds a layer of defense against today's prevalent ransomware attacks.

Google Cloud



Recommendations:

- **Configure Spanner PITR and Backup and Restore features to protect against and recover from logical corruption of a small subset of rows or larger database wide corruption.**

  Recover data using point-in-time recovery.
  Spanner to Cloud Storage Avro template - Dataflow

  PITR:
- **Set the PITR retention period (version_retention_period).** The default is 1 hour and the maximum is 7 days. Increasing retention period results in retaining more data, this can have slight performance implication on the serving system. So increase this setting gradually, starting with 1-3 days and ensuring database performance is acceptable. You may need to increase the resources to account for increased data retention.
- Before changing the default in a production system, **test** how increasing the retention period would impact your application performance and storage usage in a test environment.
- Consider **increasing the retention period incrementally** in production (eg: 6 hours 12 hours 24 hours etc).
- Increasing the PITR **retention period** provides the following benefits and potential trade offs for consideration:
  - ***Greater protection of your data.*** *Facilitates data recovery and analysis over a longer period of time.*
  - ***Increased compute cost.*** *Additional compute may be required to process multiple versions of your data.*
  - ***Increased storage cost.*** *Additional storage will be required to store multiple versions of your data.*
  - ***Increased time to process schema updates.*** *More schema versions must be retained for a longer period of time leading to additional processing time for schema changes.*

Google Cloud

Backup and Restore:

- **Schedule regular backups.** Consider scheduling at a frequency (daily) that meets your recovery needs and budget. This wIll not affect serving performance as backups are executed in the background.
- **Spanner backup schedules**: Spanner natively supports backup scheduling. The service allows for the scheduling of a backup of a specific database with configuration for the type of backups (eg: full or incremental), the time for the backup, frequency, and retention. All new instances will have a default set for the databases created within them.
- **Set a retention period of at least 7 days** for backups in order to manage storage costs and provide multiple options when it is time to restore. For regulatory or business purposes you may decide to set the retention period for some of your backups to retain them for as long as 1 year.
- **Optionally: Copy the backups to a different region** if you wish to protect against a complete failure or disaster in the production regions. Multi-Region configs will store a copy of the backup in all the regions that are a part of the multi-region config. This is covered in detail later.
- Backups typically take 1-4 hours but could be longer depending on your database size.
- Spanner charges separately for the amount of backup storage used. See the pricing documentation for additional pricing details.
- **Spanner incremental backups:** incremental backups is available for Spanner and is a very efficient way to backup your database. Unlike full backups, which copies the entire database, only data that has changed since the last incremental or full backup is copied. As a result, the size of backups is significantly lower, especially for databases with low change rates. Consequently, storage costs will be lower for storing the number of backups required to meet your RPO.
- Additional Details on Spanner backup and restore functionality can be found in the documentation.

.

Google Cloud

Major geographical disaster

The 3-2-1 backup strategy is a tried-and-true method for protecting your database. Three (3) copies of your data (production, online backup , offline secondary backup copy), two (2) storage mediums (online disk, cloud storage), one (1) offsite. With the 3-2-1 backup strategy some businesses also require an extended distance between the online database and the “offsite” backup in the event of a major or catastrophic disaster (eg: hurricane) affecting access to multiple regions.

This is also a compliance requirement in some regulated industries. FINRA requires that backup copies are stored “off-network” from the online system.

Cross-region copy of backups

Spanner's cross-region backup copy feature lets you duplicate backups across different geographic regions. This provides enhanced disaster recovery capabilities, ensuring your data is safe even if an entire region experiences an outage.

The **RPO** in this recovery scenario will be relatively higher . However, with frequent backups and copies automated this can be managed according to risk. The restore operation delivers rapid time-to-first-byte results because the database accesses the backup directly to meet a low **RTO**.

Recommendations:

- **Copy backups to another region** using the Spanner cross-region backup feature to meet DR and compliance requirements to protect against the possibility of a major disaster affecting access to multiple regions.
- **Automate** and included as a step in the overall backup automation process.

Google Cloud

Accidental database deletion

Although it does not occur frequently, it is possible that a database could be deleted accidentally due to operator error or by executing a script (eg: terraform) against the wrong environment.

Spanner database deletion protection

Spanner database deletion protection prevents the accidental deletion of existing databases by users or service accounts that have the necessary IAM permissions to delete the database. By enabling database deletion protection, you can safeguard databases that are important to your application and services.

Recommendations:

- Deletion protection is disabled by default. **Enable deletion protection** for production databases immediately after they have been created.
- **Routinely verify** the setting to maintain compliance with requirements.
- The deletion protection feature is preventative. In the absence of deletion protection, if database is accidentally deleted the Spanner restore feature can be used to restore the database. **Ensure that Spanner scheduled backups are enabled** to run regularly in order to recover from such a failure and meet your RPO. New instances will automatically have a default backup schedule set, which will be applied to each database that is created on the instance.

It is possible that a database could be deleted accidentally.

Spanner database deletion protection prevents the accidental deletion of existing databases .

Google Cloud

Access to backup media

Cyberattacks and data breaches dominate the headlines, underscoring the critical need for robust data protection. Unencrypted backups represent a significant vulnerability for databases, leaving sensitive information exposed if attackers gain access to backup files.

Additionally, encryption of backups is a requirement in industries such as banking and healthcare and to meet compliance standards such as ISO 27001, FINRA, and HIPAA. FINRA requires that backup copies are encrypted.

FINRA requires that backup copies are encrypted.

Spanner backup encryption

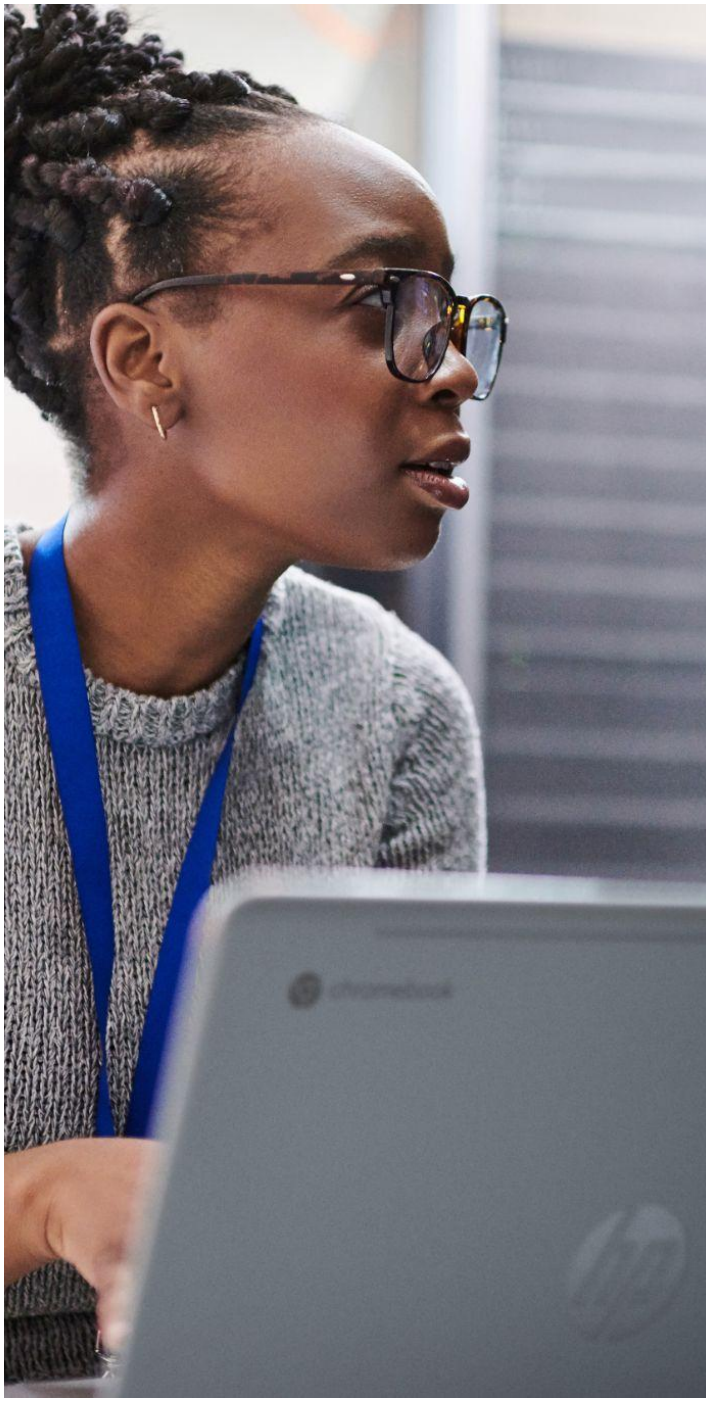
GCP prioritizes security, encrypting all customer data at rest by default. This includes your Spanner database and its backups, providing robust protection without additional configuration. Customers can choose to use the default Google managed encryption keys (GMEK) or choose their own Customer managed encryption keys (CMEK) for database and backups. Whichever is chosen for the database will be applied to the backups . However, this behaviour can be overridden by specifying a key using CMEK during backup creation.

GCP prioritizes security, encrypting all customer data at rest by default.

Recommendations:

- Encryption using GMEK will meet the requirements of most business because it provides encryption of all data at rest and uses encryption algorithms of a high strength (ie AES-256 or better). In this case **nothing needs to be configured in Spanner**.
- For organizations that are in a regulated industry or simply need to manage their own encryption keys to meet their internal security requirements **use CMEK** . Depending on the level of separation of duty (for example separate backup administrator and database administrator) ,consider using a CMEK key for backup operations.

Google Cloud



Long term data retention

Most businesses have a data retention policy that describes how long data needs to be retained and how accessible it should be. This is often required in the event of legal cases, audits, and possibly analysis of historic data for trends or patterns. The policies inherently will cover important business data stored within databases. The processes involved to retain the data in databases over a long term period is also referred to as data archiving.

Data should only be kept for how long it is needed, whether it is 1 minute or 7 years. Industries like FSI, healthcare, and publicly traded companies often have strict data retention policies driven by regulations. For example, SOX mandates 7-year retention for certain data, while HIPAA requires 6 years.

The data could be needed for analytics, reporting, or across sandbox, dev, and test environments.

Spanner import and export

The Spanner export functionality gives you the option to export your database to Cloud Storage in either CSV or Avro format. The data can be imported into any database service supporting these formats.

The data can be imported into other Spanner instances across regions or projects and for dev, test, and sandbox environments.

Your exported data can be retained indefinitely on Cloud Storage. Alternatively, the object lifecycle management feature of Cloud Storage can be used to manage the retention, cost and destruction of the data. Also, bucket locks can be used to ensure that data cannot be deleted before it expires.

Google Cloud

Exporting data to Cloud storage regularly will also safeguard against the rising threat of ransomware attacks.

In many cases where data is not required or needs to be retained for more than a year, using the export feature is not required. Spanner backup and restore feature would be sufficient to meet your requirements.

Spanner Data Boost for export

Performing a data export , on most database systems, requires increased CPU and IO resources . This would normally impact online transaction processing on the systems. In the past the only solution to avoid that was to perform exports very selectively and at precise times when the transactional workload on the database is at its lowest. Even then there could be a risk of affecting really critical users and transactions. Scaling up or out is another option that is available on systems that can scale but this can potentially be costly.

Spanner Data Boost will offload the CPU resources required for your export job to internal Google Managed CPU nodes and avoid negatively impacting your transactional operations, cause resource contention, or require you to scale up the nodes on your system.

Export data with Data Boost

Performing a data export , on most database systems, requires increased CPU and IO resources.

Spanner Data Boost will offload the CPU resources required for your export job.

Recommendations:

- **Use Spanner import export** feature to meet long term data retention and archiving requirements. It is also recommended as another layer of defence to safeguard against ransomware attacks.
- **Export the tables and data needed to Cloud Storage r**egularly (weekly, monthly, etc) .
- **Use Spanner Data Boost to offload resource requirements for your export job** and avoid negative impact to your transactional operations from resource contention.
- **Use Cloud Scheduler** to schedule exports. This can be weekly or monthly for more frequent data snapshots. Or Quarterly or yearly at more strategic points in time.
- Use a **bucket naming and folder structure** that is easy to understand and organize the exports.
- **Use object lifecycle management** polices on Cloud Storage to manage retention of the exported data files on Cloud Storage.
- Consider using bucket locks to ensure that data cannot be deleted before it expires.
- Additionally, the exported data may be used for:
    - copying data to and refreshing Dev/Test systems
    - AI/ML or Analytics in other services
    - migrating across environments
    - migrating across storage systems

Chapter 6

Test, Test, Test

Though Spanner ensures that each individual feature (backup, restore, copy, import, export) works, It is import to test end-to-end application recovery process to ensure it meets customer's RTO & RPO expectation.

Always test your backup and restore configuration and processes. It is very important to test your backup, configuration, and recovery processes before deploying them into production. It is also important to optimize and test them on an regular basis. All possible failure and disaster scenarios needs to be tested and documented for consistency and efficiency.

This is also a compliance and regulatory requirement in some industries such as FSI and Utilities,

Consider the following when developing a backup and recovery test plan:

- The most likely incidents and disasters that you would need to recover from. Also consider unlikely events including failed backups.
- The steps involved to restore the data, building your playbook, and ensuring that it is easily understood.
- The time to restore the database (RTO).
- Data that is lost after restoration if any (RPO).
- The total time the database will be unavailable (RTO).
- Recovery to different points in time.
- TIme to complete the backup.
- Performance impact when the backup is running.
- Impact on storage utilization as a result of data protection policies (backup size, PITR retention, TTL).
- Overall impact on transactions when backups are not running (PITR retention time, TTL, restoring another database within the same instance).
- Failing over and recovering applications in the event of a zonal or regional failure. This can be simulated in Spanner by blocking the network traffic for the application in a specific region or zone.
- Consider the costs associated with each test, the configuration and the procedures and optimize for costs as needed. For example, maybe the PITR time does not need to be for as long a period as initially set to meet the business requirements.

It is very important to test your end-to-end backup, configuration, and recovery processes.

This is also a compliance and regulatory requirement in some industries.

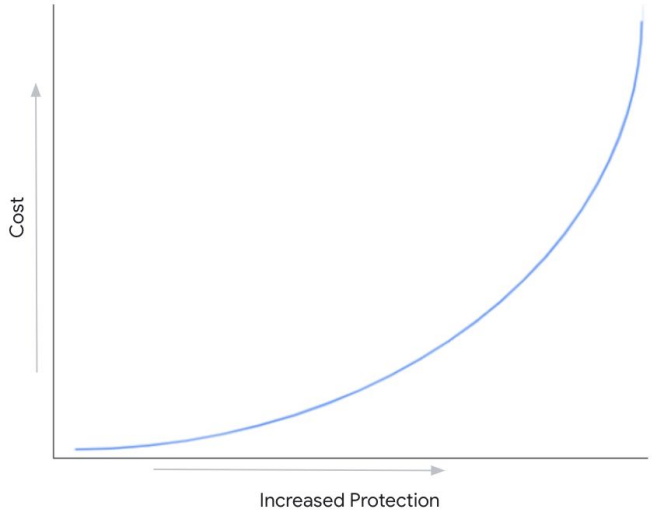
Google Cloud

Chapter 7

Cost Optimization

Increased database protection typically means higher costs. Spanner solves this, offering robust protection as a fully-managed, serverless solution.

The diagram below shows how cost typically increase for databases as improved protection (RTO, RPO and SLA) is sought out.



Choosing and configuring Spanner features for your instance depends on your application's use case, business requirements, and operational needs. This selection process will involve trade-offs between cost and performance, as is common with most database services. Careful consideration of these factors will help you optimize your Spanner instance for your specific needs.

Google Cloud

Cost Optimization Options for Data Protection and Recovery

Consider these options when seeking to optimize your cost for protecting your database :

- Chose regional configurations only for 99.99% availability. Read only latencies can be optimized by using read replicas in other regions. Backups and exports can be copied to other regions.
- Execute Spanner backups no more frequently than is required to meet your requirements. **Spanner backup schedules** is available to schedule full or incremental backups of databases at the required time, and frequency, and to specify the optimal retention period for backup copies.
- To optimize backup storage costs and capacity, leverage Spanner's incremental backup functionality. Incremental backups only capture changes made since the previous backup, whether full or incremental. Consequently, storage needs are minimized, particularly for data with infrequent modifications. This efficiency enables more frequent backups at a much lower cost, ensuring that your RPO can be met.
- Set backup retention to retain your backups for only as long as they are needed for recovery and/or compliance
- Use cloud Spanner TTL to remove unwanted table rows and maintain the size of the database and storage utilization to only what you need. This also translates to smaller backup storage in multiples.
- Configure PITR retention to retain data only for as long as is needed to meet business requirements.
- Use Data Boost for Exports
- Use Cloud Storage Object LIfecycle Management to keep only the exported files that are needed.

Summary

Protect your database with Spanner

Spanner provides all the features you need to implement a strategy to meet your requirements for database resiliency, data and disaster recovery, long term retention, and archiving.

Spanner provides industry leading five nines availability and you can backup and restore your data at anytime with a few clicks on the Google Cloud console or with a single API call,

Refer to the **Spanner documentation** to learn more about Spanner database protection features and for technical implementation details

Google Cloud

References

- Spanner documentation
- Architecting disaster recovery for cloud infrastructure outages
- Backup and restore overview | Spanner
- Failure Scenarios and Resiliency with Spanner
- Database Backup and Recovery Best Practices
- The Business Case for Data Backup and Recovery
- Choose between backup and restore or import and export
- Monitor instances with Cloud Monitoring | Spanner | Google Cloud

Google Cloud