

# LeMans Filestore LeMans Backup, Archival & Recovery Design

<b>Contents</b>
<ul style="list-style-type: none"> <li>• <a href="#">1.LeMans Backup, Archival &amp; Recovery</a> <ul style="list-style-type: none"> <li>• <a href="#">1.1 Current On Premise Solution Overview</a></li> <li>• <a href="#">1.2 Requirements for GCP LeMans Backup</a> <ul style="list-style-type: none"> <li>• <a href="#">1.2.1 Requirement Workshop Notes</a></li> </ul> </li> <li>• <a href="#">1.3 Solution Options Proposed</a></li> <li>• <a href="#">1.4 Patterns</a> <ul style="list-style-type: none"> <li>• <a href="#">1.4.1 Pattern 1: Regional Filestore</a> <ul style="list-style-type: none"> <li>• <a href="#">1.4.1.1 Cost Regional Filestore</a></li> </ul> </li> <li>• <a href="#">1.4.2 Pattern 2: Filestore Zonal</a> <ul style="list-style-type: none"> <li>• <a href="#">1.4.2.1 Cost Zonal Filestore</a></li> </ul> </li> <li>• <a href="#">1.4.3 Pattern 3: Filestore Snapshots</a> <ul style="list-style-type: none"> <li>• <a href="#">1.4.4.1 Process Flow</a></li> </ul> </li> <li>• <a href="#">1.4.5 Pattern 5: Archival in GCS Buckets</a></li> </ul> </li> <li>• <a href="#">2. RACI Matrix</a></li> <li>• <a href="#">3. Pattern Component Hardening Status</a></li> <li>• <a href="#">4. See Also</a></li> <li>• <a href="#">5. External Links</a></li> </ul> </li></ul>

<b>Pattern ID</b>	LEM-FST-DR
<b>Definition</b>	LeMans is a Low category system. The SLA for recovery, the system must be able to recover the entire data in case of a disaster.
<b>Scope</b>	Filestore Instances
<b>Version</b>	v0.2
<b>Owner</b>	@Hari Lakshmipathi @Fahad Ali (Unlicensed) @Anil Gogia (Unlicensed) @Shweta Dixit (Unlicensed) @Pinky Jain (Unlicensed)
<b>Status</b>	APPROVED
<b>Creation Date</b>	29 Jul 2024
<b>Approvers</b>	@Ubeyde Omer
<b>Recommendation</b>	Option 1 - Regional Snapshots, Management and Archival
<b>Approved Option</b>	Option 2 - Zonal Snapshots, Management and Archival

## Version History

> Version History

Author	Version	Date	Reason for Change
@Hari Lakshmipathi (Deactivated) @Fahad Ali (Unlicensed)	0.1	29 Jul 2024	Initial draft
@Anil Gogia (Unlicensed) @Shweta Dixit (Unlicensed) @Pinky Jain (Unlicensed)	0.2	02 Sep 2024	Added requirements, decision logs and framed the 2 proposed options with the patterns available.

## 1.LeMans Backup, Archival & Recovery

### 1.1 Current On Premise Solution Overview

**BIA Availability:** Low

**LeMans Recovery Point Objective: 24 hrs**

- Restoration of data from Backup = 1 week
- Failover to second data centre < 24 hrs

**LeMans Recovery Time Objective: < 24 hrs**

The backup and archive solution is provided by Tivoli Storage Manager (TSM), which runs on a weekly basis at 23:00 on a Wednesday. The backup takes a snapshot of the delta written to disk since the previous backup, which is then written to tape for archival.

This solution is used to provide a recovery in the event of a total outage and is also used on an ad-hoc basis for end-user restore requests, as a data archival/retrieval solution. Up to 3 iterations of files are retained for up to 12 years. Over time, this has accumulated into over 11.5PB of backup being consumed by LeMans, resulting in the largest backup footprint in the Bank. On 05 July 2022 it was highlighted that the backup is approaching capacity with no option to extend. Based on the current volume of data being backed-up, this is expected to be fully consumed by the end of the year. This issue is being investigated by the BAU teams and will likely introduce new parameters on the backup requirements.

The following requirements are served by the TSM solution on prem:

1. The ability to recover in the event of a disaster, resulting in the total (or partial) loss of the IaaS within GCP.
2. The ability to restore ad-hoc data requests for a specific dataset, or set of datasets.

## 1.2 Requirements for GCP LeMans Backup

Business agreed to follow the on-premise backup, snapshots and file restoration mechanism. Upon analysis we discovered that files are first written on underlying GPFS [wk1] and written to snapshots after the weekly backup [wk2]. Then written off to tape backup from the snapshot [wk2]. Backups are held outside the Live and Snapshot areas and kept until the retention

period lapses. This means data written between weekly snapshots cannot be recovered if the failure happens in that duration.

In GCP, the BIA Availability rating is unchanged, therefore it is necessary to provide a recovery solution that meets the SLAs outlined:

**BIA Availability:** Low

**LeMans Recovery Point Objective:** < 24-hours

**LeMans Recovery Time Objective:** < 24-hours

Following requirements are curated using existing on premise requirements ([LeMans Requirements - Data Migrations - Lloyds Banking Group Confluence](#)) and analysing the on premise as-is processes.

	<b>High level requirement description</b>	<b>Requirement Description</b>	<b>Solution Category</b>
R1	The solution must migrate the processes for archival of unused data and also the functionality to restore data from backups..	To have a safety net process which tidies up certain areas that have missed the housekeeping. 'X' months of data on Online Data(GPFS) based on business process in the MDL layer. Housekeeping of the MDL area is generally controlled by individual jobs.  The data will be moved in different storage class based on their age and will be retained as per the on-prem retention policy. Some data will be kept in the system for 12 years after which it will be deleted.	Archival
R2	The solution must migrate the processes for archival of unused data and also the functionality to restore data from backups..	Some data tables can be automatically archived based on certain archive rules and can be retrieved when needed. To effectively manage the Team folder and save folders' space	Archival

## [Master Requirement Document - LeMans Requirements - Data Migrations - Lloyds Banking Group Confluence](#)

### 1.2.1 Requirement Workshop Notes

We conducted Requirements Workshops with the below listed participants and agreed on the following requirements and baseline configurations:

List of participants:

› [List of participants:](#)

- [@ Alastair Frame](#)
- [@ Simon Dean](#)
- [@ Jenny Yip](#)
- [@ Simon\) Wilson](#)
- [@ Andrew James](#)
- [@ Fahad Ali \(Unlicensed\)](#)
- [@ Rizwan Akhtar](#)
- [@ Adam Baines](#)
- [@ Palapolu, Keerthi \(Data & Machine Learning Platform\) \(Deactivated\)](#)
- [@ Amit Sharma](#)
- [@ Sanjib Dhar \(Unlicensed\)](#)
- [@ Santanu Seth](#)
- [@ Mohan Devarajan](#)
- [@ Pinky Jain \(Unlicensed\)](#)
- [@ Anil Gogia \(Unlicensed\)](#)
- [@ Shweta Dixit \(Unlicensed\)](#)

## 1. Regional Filestore

- On GCP, Regional Filestore instance will be used in case of restoring zonal outage as a solution for High Availability and Resiliency.
- This option provides **Zero RTO and RPO** of Filestore instance as another instance is available in second zone having latest replicated data i.e. you don't need to do anything for data replication.
- For LeMans application to be available after fail-over, **RTO** for restoration of SAS Compute Infrastructure will be 6-8 hrs.

## 1. Snapshots

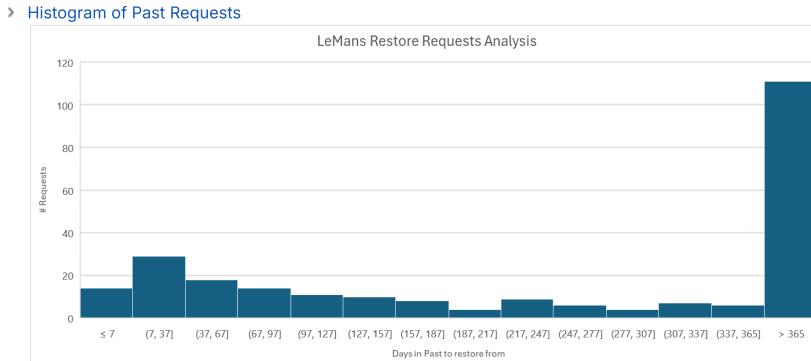
- On GCP, Filestore Snapshot will be used to restore files to the point in time. Snapshot is a preserved state of the file share with incremental changes of data captured in .snapshot directory.
- On GCP, we agreed on the option for taking Snapshots of live data in Filestore instance which is stored in the instance without any ring fenced storage for Snapshots.
- We agreed on **daily** snapshots with retention period of **7 days** as a starting point\* similar to on-premise process. This will provide RTO < 24 hours and RPO will depend on the incremental changes in data.
- We acknowledged the limitation of 240 active snapshots on GCP Filestore.
- On premise backup policy document suggests Snapshots reference the same disk and the actual space consumed is the sum of all of these disks, leading to 3.5x space requirements for GPFS. We are analysing this to provide size estimates for GCP Filestore instances.

## 2. Backup

- On GCP, Managed Filestore Backup will be used to restore entire data in case of an accidental deletion or corruption of Filestore Instance, not recommended for file level restoration.
- We agreed on **weekly** managed instance level backup snapshots with retention period of **12 months** as a starting point\* similar to on-premise process. RTO is >24 and <36 hours. We performed a **POC** where we observed the data restoration time is 3 TB/h. RPO is 1 week.  
*\*Frequency can be slightly altered based on the analysis of cost to store the Snapshots on Filestore, the analysis is done using the snapshot increments sizes from on premise snapshots as this data is yet not available on GCP*

## 3. Archival and Historical File Restoration

- The solution must provide a capability to archive the **data every week**. Use on premise retention policy of data as defined below in Retention Policy table
- The solution must provide a capability to retain the data for 12 years after which the data will be deleted from the system. The data will be retained in different storage class based on the age of the data
- The solution must provide a functionality to restore data from backup. We analysed that users request file restoration frequently(almost daily) and request is fulfilled by Run team. The files are restored from snapshots or tape depending on the retrieval date requested. More details on the process are available [on SharePoint](#).
- Analysing the requests from last year suggests that less than 10% of files are recovered from snapshots on the account of data corruption or data loss. Most of the file restored are from historical data required for Project use cases.



- On GCP, we will follow similar process. If the file requested is not available to recover from snapshots, then we should recover it from GCS.
- Assumption - Storing all files in standard storage as there are no requirements on different retrieval patterns for data. From the analysis of this year's retrieval requests, Files requested are from 2007 to current month.
- We explained different storage classes and cost available in GCS. If there are changes in the storage classes or lifecycle rules, they should be captured as requirements for future enhancements. GCSFuse is proposed as a solution option, to be explored further via POCs.

## 1.3 Solution Options Proposed

The GCP PaaS solution deploys Filestore Instances, which is a Cloud native shared filesystem provided by Google. The LeMans data is stored within the Filestore Instances and surfaced to the SAS clients via NFS mounts.

The patterns describe the change to the backup approach as a result of the migration to GCP and outlines how the backup and recovery requirements will be fulfilled to ensure the SLAs for RTO and RPO can be achieved. The future solution has no requirement to combine the backup and archival requirements into a single solution, instead it can be split into distinct approaches to manage a full restore in the event of a zonal outage, which is separate from the ability to restore single or multiple files.

The two options in this paper use the combination of patterns to meet the requirements for resiliency and high availability of LeMans application on GCP.

Option 1 ( <b>Recommended</b> ) - Regional Filestore, Snapshots, Managed Backups and Archival	Option 2 - Zonal Filestore, Snapshots, Managed Backups and Archival
Pattern 1 : Regional Filestore is the most suitable solution to achieve DR within 24 hour SLA as it will make data highly available between zones. Regional Filestore is ~2x in cost than zonal Filestore.	Pattern 2 : Zonal Filestore is cost effective and high performance, scalable storage solution. We will not be able to achieve DR within 24 hours

Pattern 3 : Filestore Snapshots is for faster ad-hoc file restoration and suitable for ad-hoc file restoration requests with some additional cost overhead as described above with details. However, it does not support retention policies for long term archival.

Pattern 4 : Filestore backups can provide contingency against Filestore deletion. Restoration from Filestore backups are 5 times faster as compared to restoration from GCS bucket using gsutil in Pattern 5. However, restoration from backups cannot meet 24-hour SLA required for DR.

Pattern 5 : is required for archiving data as per the retention policies. It can also be used for ad-hoc file restoration. Pattern is suitable for long term archiving. It can also be used for full restoration, but it is slower than Filestore Backups.

Below is a comparison of all patterns against key requirements:

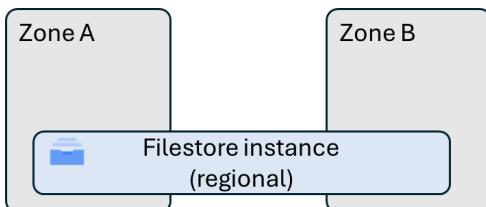
Pattern No	DR backup	DR restoration	Ad-hoc archival	Ad-hoc restoration	Retention Policy
Pattern 1: Regional Filestore	Yes	Yes, meet 24 hrs SLA	No	No	not applicable
Pattern 2: Zonal Filestore	No	No	No	No	not applicable
Pattern 3: Filestore Snapshots	No	No	Yes	Yes	not supported
Pattern 4: Filestore Backups	Yes	Does not meet SLA	No	No	not supported
Pattern 5: Archival in GCS Buckets	Yes	Does not meet SLA	Yes	Yes	supported

As evident, the key difference between the two options is using Zonal vs Regional Filestore instance. Below is a comparison of Regional Filestore and Zonal Filestore across key requirements:

	Zonal Filestore	Regional Filestore
Cost	Lower (30-50% savings)	Higher (1.5-2x cost of Zonal)
Recovery	In a zone failure scenario, recovery could take several hours or days, depending on backups and disaster recovery setups.	High availability (HA) and automatic failover in case of a zone outage.
RPO	RPO is highly dependent on the snapshot and backup schedules, potentially >24 hours if snapshots/backups are infrequent.	RPO: Close to zero because the data is replicated across zones, ensuring minimal data loss.
RTO	RTO can be significantly longer than 24 hours due to the time needed to recover from backups and restore services in a new zone or region.	RTO: Less than 24 hours, typically much lower, as services can fail over to another zone in the same region with minimal disruption.
Risk	Higher risk of Data loss	Significantly reduced risk of data loss due to multi-zone replication.

## 1.4 Patterns

### 1.4.1 Pattern 1: Regional Filestore



Unique Reference	
Option	Regional Filestore
Solution Approach	Provides regional storage resilient to zone outages and meet SLA for DR
Components required	Regional Filestore

<b>Benefits</b>	<ul style="list-style-type: none"> <li>Managed service by google data will be replicated synchronously in Filestore instances across multiple zones.</li> <li>Data replication(backup) will be implicit by google and quick and no manual intervention is required.</li> <li>High availability for mission-critical, high performance computing workloads running within a large namespace</li> <li>Restoration within SLA of 24 hours, as it will require only spinning of new infra and pointing to regional Filestore instance.</li> </ul>
<b>Challenges</b>	<ul style="list-style-type: none"> <li>Cost of keeping replicated data in multi zone will be higher (almost double) comparative to zonal. (cost comparison given below)</li> </ul>
<b>Key Consideration</b>	
<b>Assumptions Made</b>	<ul style="list-style-type: none"> <li>Regional Filestore is available and approved for use within the bank.</li> <li>Cron is used for scheduling and Dynatrace is used for alerting</li> </ul>

#### 1.4.1.1. Cost Regional Filestore LEM-FST-BACKUP-02

Environment	Capacity (GB)	Total Regional Cost (£)	Shared Service Regional (£)	Total Regional Cost incl. VAT (£)	Notes
PROD-PROD	625,408	268,300.032	40,245.0048	368,094.2289	Regional cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.55\$ month as per GCP calculator
PROD-UAT	41,984	18,011.136	2,701.6704	24,710.37804	Regional cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.55\$ month as per GCP calculator
PROD-CIT	41,984	18,011.136	2,701.6704	24,710.37804	Regional cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.55\$ month as per GCP calculator
PROD-DEV	41,984	18,011.136	2,701.6704	24,710.37804	Regional cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.55\$ month as per GCP calculator
PRE-DEV	21,504	9,225.216	1,383.7824	12,656.53509	Regional cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.55\$ month as per GCP calculator

#### 1.4.2 Pattern 2: Filestore Zonal

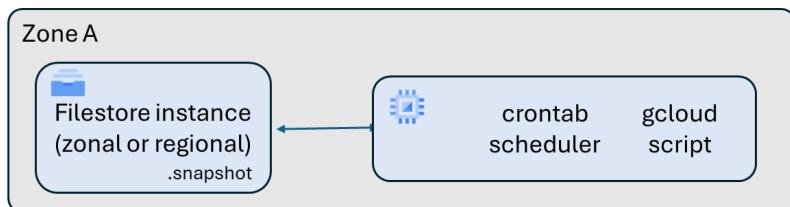


<b>Unique Reference</b>	LEM-FST-BACKUP-02.ZFS
<b>Option</b>	Zonal Filestore
<b>Solution Approach</b>	Provides cost effective, high performance, scalable and secure network storage solution.
<b>Components required</b>	Zonal Filestore, TWS
<b>Benefits</b>	<ul style="list-style-type: none"> <li>Cost effective managed storage service by google which is easy to manage.</li> <li>High performance computing workloads running within a large namespace</li> <li>In case of zonal outage, requires restoration of data from backup option or from archival storage followed by spinning of new infra and pointing to newly provisioned zonal instance.</li> </ul>
<b>Challenges</b>	<ul style="list-style-type: none"> <li>DR will not be achieved in 24 hours considering the amount of data we have ( 600-800TB) in LeMans application</li> </ul>
<b>Key Consideration</b>	<ul style="list-style-type: none"> <li>Zonal Filestore cost is almost half of regional FileStore</li> </ul>
<b>Assumptions Made</b>	<ul style="list-style-type: none"> <li>Zonal Filestore is available and approved for use within the bank</li> <li>Cron is used for scheduling and Dynatrace is used for alerting</li> </ul>

#### 1.4.2.1 Cost Zonal Filestore

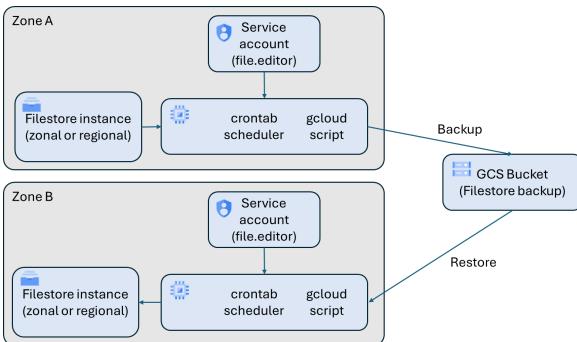
Environment	Capacity (GB)	Total Zonal Cost (£)	Shared Service Zonal (£)	Total Zonal Cost incl. VAT (£)	Notes
PROD-PROD	625,408	146,345.472	21,951.8208	200,778.6703	Zonal cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.30\$ month as per GCP calculator
PROD-UAT	41,984	9,824.256	1,473.6384	13,478.38802	Zonal cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.30\$ month as per GCP calculator
PROD-CIT	41,984	9,824.256	1,473.6384	13,478.38802	Zonal cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.30\$ month as per GCP calculator
PROD-DEV	41,984	9,824.256	1,473.6384	13,478.38802	Zonal cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.30\$ month as per GCP calculator
PRE-DEV	21,504	5,031.936	754.7904	6,903.56459	Zonal cost includes 15% shared service charge and 19.3% VAT and cost per GB is 0.30\$ month as per GCP calculator

#### 1.4.3 Pattern 3: Filestore Snapshots



Unique Reference	LEM-FST-BACKUP-03
Option	Filestore Snapshots
Solution Approach	Snapshot feature will be enabled on the Filestore which will allow us to recover the latest copy of a dataset if it is deleted.
Components required	Filestore, TWS
Benefits	<ul style="list-style-type: none"> <li>A Filestore Snapshot is preserved state of the file share data at the time of snapshot is created.</li> <li>Whenever a snapshot is triggered Filestore instance only captures modified state of files and maintains under .snapshot directory.</li> <li>Each .snapshot directory contains the snapshots of its parent directory.</li> <li>Can be configured to maintain the required number of iterations.</li> <li>Individual files can be instantly recovered. Not impacted by utf8 characters in filenames.</li> <li>Snapshot creation usually takes no longer than two minutes to complete because it doesn't involve the copying of data, and it doesn't affect instance performance.</li> </ul>
Challenges	<ul style="list-style-type: none"> <li>Snapshots takes up space within the Filestore Instance. Additional capacity needs to be planned if snapshots are being configured.</li> <li>script need to be scheduled using crontab.</li> </ul>
Key Consideration	<ul style="list-style-type: none"> <li>Quicker recovery of individual files with snapshot feature</li> </ul>
Assumptions Made	<ul style="list-style-type: none"> <li>Filestore snapshot feature is available as GA within the Bank</li> <li>Cron is used for scheduling and Dynatrace is used for alerting</li> </ul>
Additional Notes	<b>Additional Storage requirements</b> - Initial Snapshot Size: Minimal (metadata and pointers). after that only data changes are added to snapshot, example if 2 TB data changed in a day out of 8 TB total data, after taking snapshot, Filestore size will be 8 TB + 2 TB = 10 TB. Additional storage depends on the volume of data changes.

#### 1.4.4 Pattern 4: Filestore Managed Backups



#### 1.4.4.1 Process Flow

##### Backup

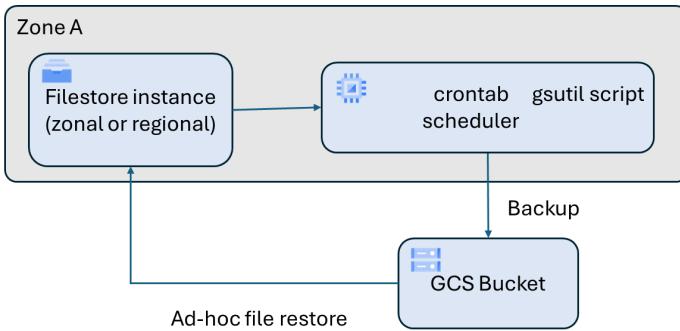
1. Create a new FST service account with file.editor permissions only.
2. Create a KMS key for the new FST service account and attach it to service account.
3. Set up a new server for Filestore backup, attach the new FST service account to it, and grant access to the run team.
4. Create three gcloud Filestore backup scripts: one for /team area, one for MDL spds, and one for sasshare01.
5. Schedule these scripts in crontab for automated runs.

##### Restore

1. create a new VM in zone B with FST service account.
2. get the backup ID.
3. create new Filestore instance using backup ID.

<b>Unique Reference</b>	LEM-FST-BACKUP-04									
<b>Option</b>	Filestore Backups,									
<b>Solution Approach</b>	Utilizing Filestore backup schedule to backup the Filestore metadata and data at regular intervals so that the Filestore instance can be launched from the specific backup in another zone in case of zonal failure									
<b>Components required</b>	Filestore, GCS Buckets, TWS									
<b>Benefits</b>	<ul style="list-style-type: none"> <li>• Integrated solution that is designed specifically to facilitate backups within the Filestore Instance</li> <li>• Provides incremental backups of changed artefacts</li> <li>• No outage requirement for creating backups</li> </ul>									
<b>Challenges</b>	<ul style="list-style-type: none"> <li>• No file level backup/recovery is possible via managed backup, need snapshots for that</li> <li>• Entire Filestore Instance needs to be created from the required backup iteration to recover the data.</li> <li>• Time to recover data is high</li> <li>• <del>For Schedule Filestore Backups, Cloud Functions, Cloud Scheduler are required which are currently not available in LBG environment.</del></li> <li>• <del>In absence of Cloud Scheduler and Cloud Functions, script need to be scheduled using crontab.</del></li> <li>• <del>for recovery during DR, gcloud command need to be run for restoring Filestore from backups because terraform module is not available to do restoration using automated pipeline.</del></li> </ul>									
<b>Key Consideration</b>	<ul style="list-style-type: none"> <li>• Individual schedules can be created for each Filestore Type and Data type to avoid disruption.</li> <li>• Filestore Backups can be quickly recovered across the region in case of zonal failure.</li> </ul>									
<b>Assumptions Made</b>	<ul style="list-style-type: none"> <li>• Filestore Backup feature is available as GA within the Bank.</li> <li>• Cron is used for scheduling and Dynatrace is used for alerting</li> </ul>									
<b>Additional Notes</b>	<p>Link to POC - <a href="#">POC - gcloud Filestore backup and recovery - Data Migrations - Lloyds Banking Group Confluence</a></p> <p>GCS Bucket size - Filestore backups will be stored in GCS buckets, GCS bucket size will equal to volume of data backed up.</p> <p>Data Volumes and Throughput</p> <table border="1"> <thead> <tr> <th>Data</th> <th>Backup Time</th> <th>Restore Time</th> </tr> </thead> <tbody> <tr> <td>4 X 47TB</td> <td>17 hours for first backup, subsequent backups will be faster as it is incremental.</td> <td>15 hours</td> </tr> <tr> <td></td> <td></td> <td>3 TB per hour</td> </tr> </tbody> </table>	Data	Backup Time	Restore Time	4 X 47TB	17 hours for first backup, subsequent backups will be faster as it is incremental.	15 hours			3 TB per hour
Data	Backup Time	Restore Time								
4 X 47TB	17 hours for first backup, subsequent backups will be faster as it is incremental.	15 hours								
		3 TB per hour								

#### 1.4.5 Pattern 5: Archival in GCS Buckets



<b>Unique Reference</b>	LEM-FST-BACKUP-05
<b>Option</b>	Schedule gsutil scripts using crontab to sync the data stored on Filestore Instances to GCS buckets.
<b>Solution Approach</b>	Schedule gsutil rsync scripts using crontab which can sync the data from the Filestore Instances to the corresponding GCS buckets where Lifecycle Policy Management is implemented which controls the duration and no of iterations needs to be backed up.
<b>Components required</b>	GCE Instances, Filestore Instances, GCS Buckets, TWS
<b>Benefits</b>	<ul style="list-style-type: none"> <li>Individual Files can be recovered using gsutil copy.</li> <li>Provides incremental backups of changed artefacts with versioning at file level. Note, incremental changes are at file level. so a small change in the file will result in copy of full file.</li> <li>No outage requirement</li> <li>Supports Retention Policy on the data backed up to GCS buckets.</li> <li>can support backup within SLA as it takes incremental backup.</li> </ul>
<b>Challenges</b>	<ul style="list-style-type: none"> <li>Custom schedules and scripts are required as part of development.</li> <li>May take longer to recover to a new Filestore Instance launched across the region in case of zonal failure. cannot meet SLA.</li> <li>Script would fail if any user created file has utf8 or special characters in filename.</li> <li>only owner of the folder can take backup of the folder and underlying files. correct permission needs to be setup for files and folders.</li> </ul>
<b>Key Consideration</b>	<ul style="list-style-type: none"> <li>Individual files can be backed up and recovered at any time.</li> </ul>
<b>Assumptions Made</b>	<ul style="list-style-type: none"> <li>No security compliance issue when scheduling scripts to sync the data to the GCS buckets</li> <li>Cron is used for scheduling and Dynatrace is used for alerting</li> </ul>
<b>Additional Notes</b>	<p><b>GCS Bucket size -</b> Filestore backups will be stored in GCS buckets, GCS bucket size will equal to volume of data backed up.</p> <p><b>Throughput</b></p> <p><b>Backup -</b> First backup will be a full backup, after there will be incremental backups for changes artifacts.</p> <p><b>Recovery -</b> 54TB in 81 hours, 0.6 TB in 1 hours approximately.</p>

## 2. RACI Matrix

	Responsible (R)	Accountable (A)	Consulted (C)	Informed (I)
approval for regional Filestore		Ubeyde, Vikki, Neevash	Google	
approve for Filestore snapshots		Ubeyde, Vikki, Neevash	Google	
approve for Filestore backups		Ubeyde, Vikki, Neevash	Google	
approve for rsync to GCS buckets		Ubeyde, Vikki, Neevash	Google	
addition to ITSD	Anil, Pinky			
low level design on approved solution and sign-off	Fahad, Anil, Pinky	Ubeyde, Vikki, Neevash	Google, Alastair Frame	
availability of regional Filestore	DCX	Amit, Santanu, Pinky		
availability of snapshot feature	DCX	Amit, Santanu, Pinky		

availability of backup feature	DCX	Amit, Santanu, Pinky		
availability of gsutil feature	DCX	Amit, Santanu, Pinky		
provision regional Filestore	Amit, Santanu		DCX	Roshan, Varun, Alastair
copy data to regional Filestore	Roshan, Varun, Alastair		Amit, Santanu	Alastair Frame
start and validate SAS application	Keerthi, Sadashiv, Mohan			
<b>sign-off on regional Filestore</b>	Fahad, Anil, Pinky	Ubeyde, Vikki, Neevash	Alastair Frame	
configure snapshot	Amit, Santanu		DCX	
sign-off on Filestore snapshot	Fahad, Anil, Pinky	Ubeyde, Vikki, Neevash	Alastair Frame	
configure backups	Amit, Santanu		DCX	
sign-off on Filestore backups	Fahad, Anil, Pinky	Ubeyde, Vikki, Neevash	Alastair Frame	
configure rsync to GCS buckets	Amit, Santanu		DCX	
sign-off rsync to GCS buckets	Fahad, Anil, Pinky	Ubeyde, Vikki, Neevash	Alastair Frame	
security sign-off	Fahad, Anil, Pinky	Keith		
availability of TWS	DCX	Amit, Santanu, Pinky		
configuration of TWS	Amit, Santanu		DCX	
configure TWS alerts	Amit, Santanu		DCX	
sign-off on TWS alerts	Fahad, Anil, Pinky	Ubeyde, Vikki, Neevash	Alastair Frame	

### 3. Pattern Component Hardening Status

GCP Component	Status
GCS Buckets	GA
Filestore	GA
GCE Instances	GA

### 4. See Also

[Google Cloud Filestore - Data Migrations - Lloyds Banking Group Confluence](#)

### 5. External Links

[Snapshots overview](#) | [Filestore](#) | [Google Cloud](#)

[Backups overview](#) | [Filestore](#) | [Google Cloud](#)

[Copy data to or from instances](#) | [Filestore](#) | [Google Cloud](#)