

Data Archiving and Cloud Replication Using Igneous

Table of contents

Introduction	3
Features	3
Benefits	3
Value	3
Architecture	4
Data movement engine	4
Cloud integration	4
Performance	5
Costs	5
Use cases	5
Try it for yourself	6
What's next	9

Introduction

Igneous integrates directly with Google Cloud Platform (GCP) to replicate data and to archive cold data for long-term storage. Igneous provides policy-based data movement and deletion, reducing management overhead and keeping total storage costs down, while helping to ensure regulatory compliance.

Set up a policy in Igneous that automatically creates a second copy of any on-premises backup data to store in Google Cloud Storage. Use Igneous to search through files stored in Cloud Storage and quickly restore them to Igneous.

Features

Igneous is a scale-out secondary storage platform installed in your data center that is designed for unstructured file data. Igneous does the following:

- Integrates directly with your primary-tier network-attached storage (NAS) device, providing round-the-clock replication, backup, and archiving.
- Manages unstructured data across your primary-tier NAS, Igneous, and any tier of Cloud Storage.
- Provides visibility into files that are stored across all storage sources through an intuitive web-based user interface (UI).
- Searches across your primary-tier NAS, Igneous, and any tier of Cloud Storage and restores data with a single click.
- Efficiently move both small and large files using transfer compression and object grouping.

Benefits

Igneous is designed to move enterprise-scale unstructured data to Cloud Storage. Igneous does the following:

- Reduces the number of Cloud Storage operations through transfer compression and object grouping.
- Saves IT time by running backups automatically in the background without affecting performance.
- Boosts the visibility of your data wherever it lives.
- Maintain continuous control whether you store data in native format or in an open, self-describing format.

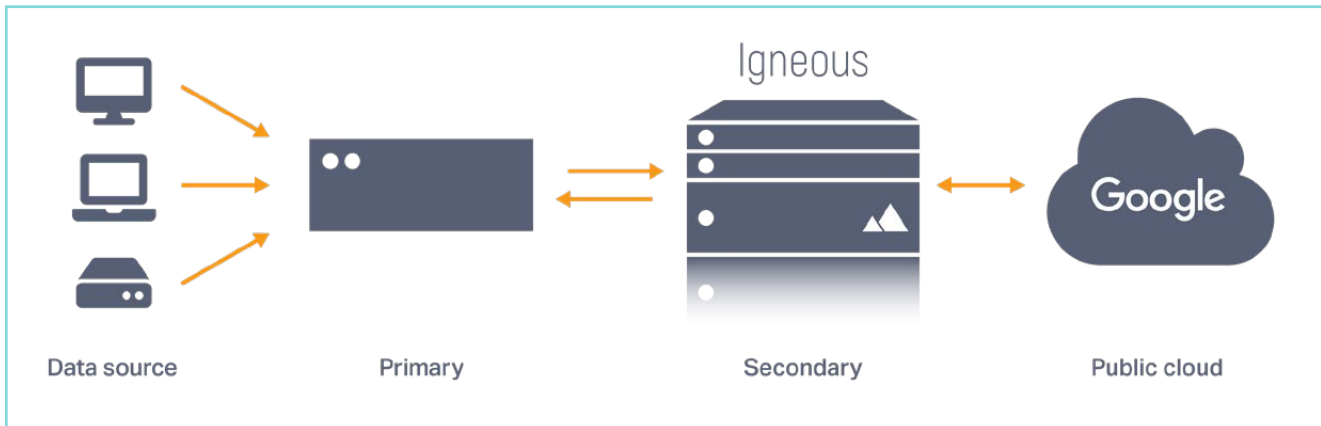
Value

Cloud Storage Nearline and Cloud Storage Coldline are valuable for long-term archiving and replication of file data, but it can be difficult, costly, and time-consuming to identify and move the data from your data center to Cloud Storage. These issues are exacerbated when dealing with billions of files across petabytes of NAS.

Igneous features a UI that provides visibility, search, and data movement functionality across all connected storage sources. Cross-tier keyword search and a consolidated secondary tier help you find and move your data.

Architecture

Igneous is an on-premises secondary storage tier that provides backup for primary NAS storage and simplifies moving data to Cloud Storage, as illustrated in the figure below.



Igneous offers the benefits of cloud services along with hardware that is installed in your data center. With its on-premises infrastructure, Igneous is designed to be your secondary storage tier. In addition, the cloud-native architecture means you can use Cloud Storage as a tertiary tier.

Igneous provides native API integration for Dell EMC Isilon, NetApp, Pure Storage FlashBlade, and Qumulo QF2. It can also back up generic NAS file systems by using parallel, latency-aware data movement that is capable of handling environments with billions of files and petabytes of data.

Data movement engine

Igneous moves data from primary storage in parallel streams. By accessing NAS systems through frontend protocols, such as Network File System (NFS) and Server Message Block (SMB), Igneous opens many parallel streams. In addition, the scanning and data movement algorithms are engineered to maximize throughput and minimize impact on primary NAS systems.

The data movement engine behind Igneous is a policy-driven system of data-replication microservices that provides integrated backup and archiving for unstructured file data. It is highly scalable, resilient, and efficient.

Cloud integration

To constrain your data center footprint and provide long-term retention, Igneous automates tiering and data replication to Cloud Storage. Set policies to replicate or group data into tiers for archiving. Backups run automatically in the background without affecting performance. Igneous integrates with and manages data across all Cloud Storage tiers (Regional, Nearline, or Coldline), allowing data to be where it needs to be when it needs to be there.

Performance

Igneous quickly scans and copies data from NAS systems at scale. The exact speed of data movement depends on the size and density of your primary NAS.

In addition, the round-the-clock backups are latency aware: when the primary tier is experiencing a lot of other traffic, the backup job slows down to avoid affecting the performance of the primary tier. Whether or not Igneous is currently moving data to and from Cloud Storage, the performance of the primary tier isn't affected.

Igneous is meant to be secondary storage. Its performance level falls between modern primary NAS that is sold by vendors like NetApp, Dell EMC Isilon, Pure Storage, and Qumulo, and modern public cloud archive tiers like Cloud Storage Nearline or Coldline storage classes. Regardless of location, data that is stored with Igneous is readily available to restore.

For most people, cold data copy speed isn't paramount. Igneous compresses and groups data to minimize time and costs.

Costs

Igneous is on-premises secondary storage that integrates directly with on-premises primary NAS systems and Cloud Storage. Igneous charges an annual subscription that is based on the capacity of data under management. Your overall cost consists of two parts: your Igneous subscription fee and your Cloud Storage fees.

The automatic policy-based tiering of Igneous means that you keep cool data on-premises for a given period of time. Only store cold data in Cloud Storage Nearline and Cloud Storage Coldline to make it less likely to pay fees for egress from Cloud Storage. In addition, Igneous throttles data movement to and from Cloud Storage to reduce total costs and manage network capacity.

Use cases

Integrate Igneous and Cloud Storage to replicate backups for disaster recovery and to store cold data for archiving.

- For extra protection of on-premises primary data, replicate each completed backup to the public cloud. Then, as short-term backups age out of Igneous, the replicated copy remains in Cloud Storage.
- For long-term storage of unstructured data, move your cold data to Cloud Storage. Using this approach, control growth of primary NAS, retain data economically, maintain access, and comply with industry-specific regulations.

Try it for yourself

Set up a data retention policy and replicate to Cloud Storage.

1. In Igneous, create a policy and enable replication.
2. In the **Add an account** drop-down list, select **Google Cloud Platform**.
3. To create a service account for Igneous, click **Go to the GCP Service Accounts page**.

Create a new policy

Backup frequency
Daily >
Enqueue from 7pm to 6am

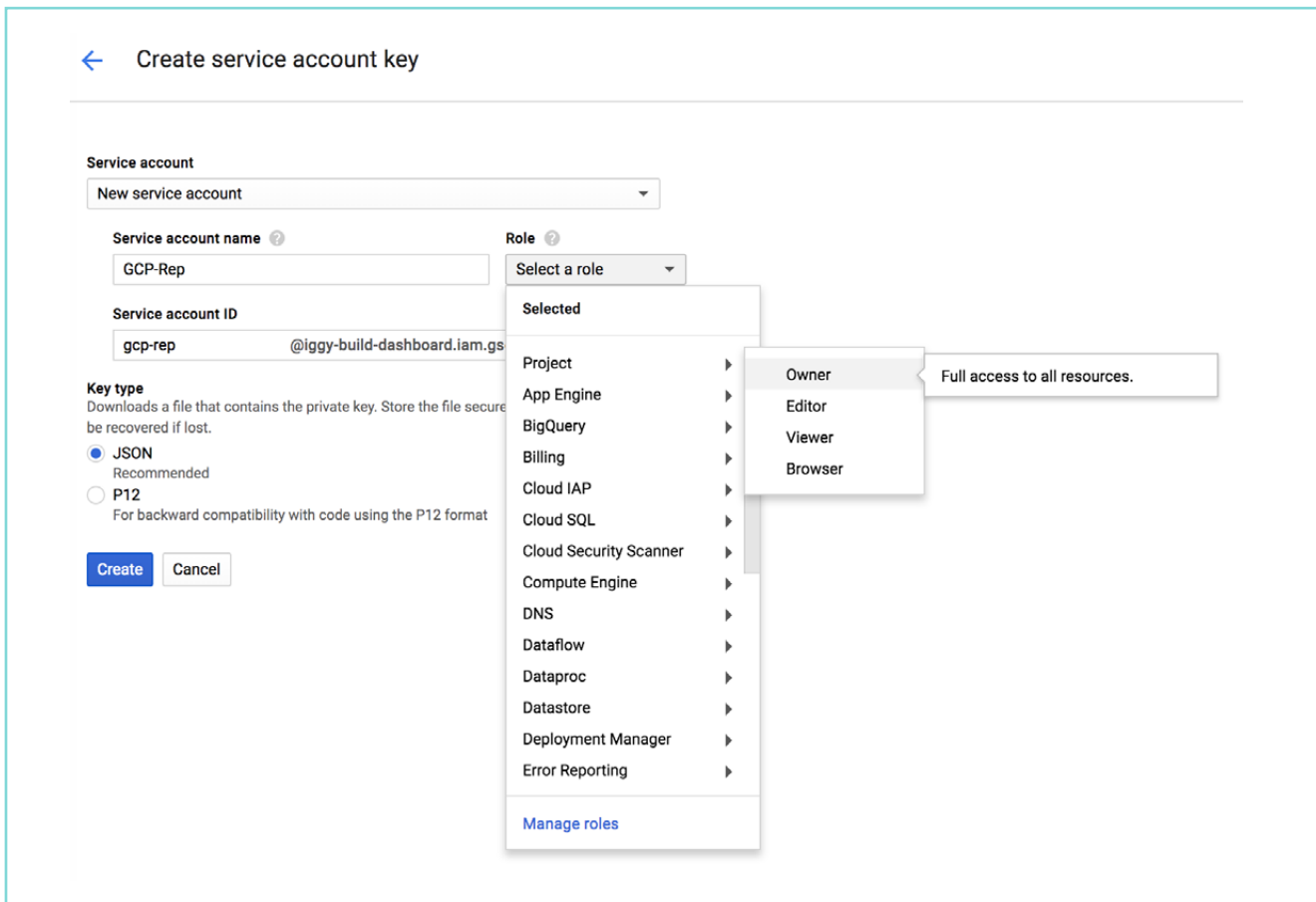
Retention
Keep data on Igneous forever
1d 30d 60d 90d 180d 1y 2y 3y 5y 7y Forever
 Enable WORM

Replication
+ Add an account

Policy name
Gold-Replication-GCP

Google Cloud Platform
Igneous will connect to Google Cloud Platform (GCP) using a service account.
Upload a JSON key generated from the service account in the GCP console. **This service account must have the Owner role.**
[Go to the GCP Service Accounts page](#)
Select JSON key Cancel

4. To create a new service account, follow these steps:
 - a. Create a unique name for the account in the **Service account name** box.
 - b. To provide the Igneous service account with project owner privileges, from the **Role** drop-down list, select **Project > Owner**.
 - c. To download a file with a private key, click **JSON**.
 - d. Click **Create** and download the JSON key.



- To upload the JSON key for the service account to Igneous, click **Select JSON key**. Locate the downloaded JSON key, and click **Open**.
- Select the storage tier that you want to use on Cloud Storage, and then click **Create policy**.

Create a new policy

Backup frequency

Daily >

Enqueue from 7pm to 6am 🕒

Retention

Keep data on Igneous forever

1d 30d 60d 90d 180d 1y 2y 3y 5y 7y Forever

Enable WORM ⓘ

Replication

iggy-build-dashboard ▾

coldline ▾

Policy name

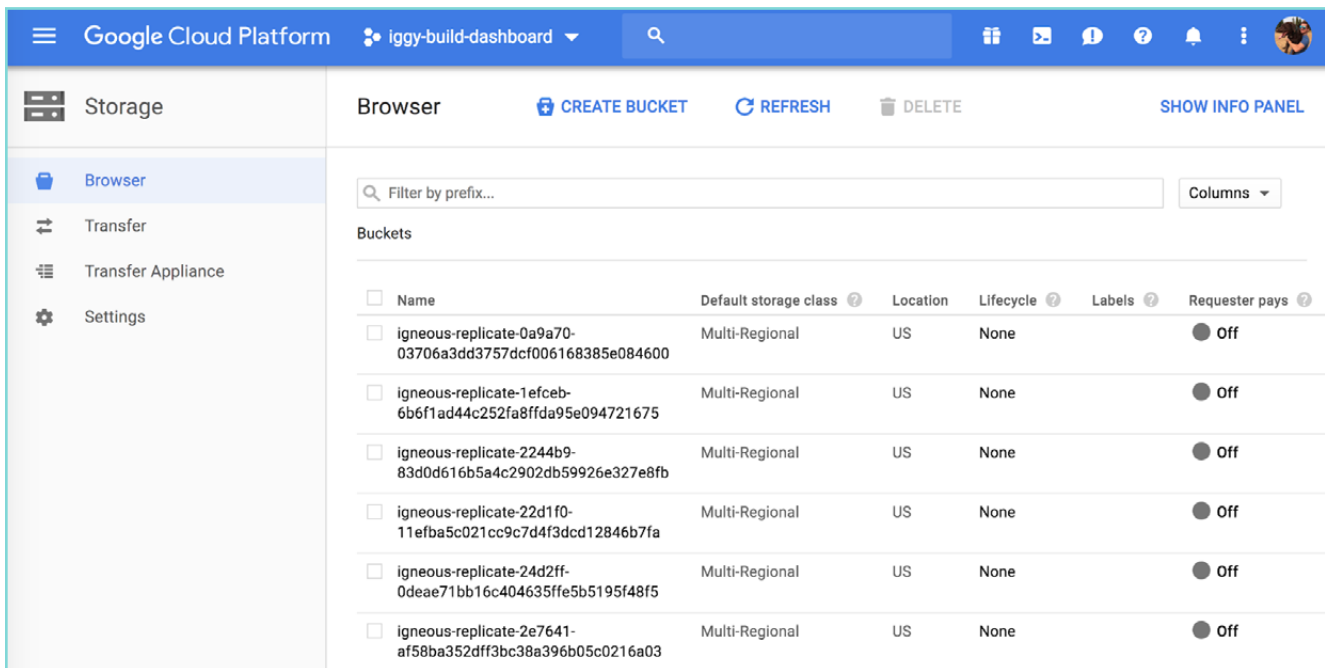
Gold-Replication-GCP

Create policy

- Apply the policy to the systems and exports that you want replicated to Cloud Storage. Select your policy from the drop-down list.
- To replicate your data to your Cloud Storage bucket immediately, click **Run now** in the **Actions** drop-down list to the right of your export.

The screenshot shows the Igneous interface with a sidebar on the left containing 'Systems', 'Policies', 'Activities', 'Data Search', and 'Reports'. The main area displays a table of systems and exports. A dropdown menu is open over the 'Gold-Replication-GCP' policy, showing options: 'NoPolicy', 'Unprotected', 'Bronze', 'CLEARME', 'Gold', 'Gold-Replication-GCP' (highlighted), 'Monthly', 'Oracle-Backup-Retention', 'Silver', and 'TwiceADay-2Y'. The table lists systems like '/dense1' through '/pipeline_molly_destination_nfs' with columns for policy, size, and NFS status.

Igneous creates the buckets, which appear in your Cloud Storage bucket.



The screenshot shows the Google Cloud Platform Storage interface. The top navigation bar includes the Google Cloud Platform logo, the user's profile, and various utility icons. The left sidebar shows the 'Storage' menu with options for 'Browser', 'Transfer', 'Transfer Appliance', and 'Settings'. The main content area is titled 'Browser' and includes buttons for 'CREATE BUCKET', 'REFRESH', 'DELETE', and 'SHOW INFO PANEL'. A search bar is present with the placeholder text 'Filter by prefix...'. Below the search bar, a table lists several buckets created by Igneous. Each row includes a checkbox, the bucket name, the default storage class (Multi-Regional), the location (US), the lifecycle (None), labels, and the requester pays status (Off).

<input type="checkbox"/>	Name	Default storage class ?	Location	Lifecycle ?	Labels ?	Requester pays ?
<input type="checkbox"/>	igneous-replicate-0a9a70-03706a3dd3757dcf006168385e084600	Multi-Regional	US	None		Off
<input type="checkbox"/>	igneous-replicate-1efceb-6b6f1ad44c252fa8ffda95e094721675	Multi-Regional	US	None		Off
<input type="checkbox"/>	igneous-replicate-2244b9-83d0d616b5a4c2902db59926e327e8fb	Multi-Regional	US	None		Off
<input type="checkbox"/>	igneous-replicate-22d1f0-11efba5c021cc9c7d4f3dcd12846b7fa	Multi-Regional	US	None		Off
<input type="checkbox"/>	igneous-replicate-24d2ff-0deae71bb16c404635ffe5b5195f48f5	Multi-Regional	US	None		Off
<input type="checkbox"/>	igneous-replicate-2e7641-af58ba352dff3bc38a396b05c0216a03	Multi-Regional	US	None		Off

What's next

- Check out [Integration with Google Cloud Platform](#).
- Watch the [Igneous Protect](#) video to see how Igneous fits into your environment.
- Read about how [cloud architecture](#) is vital to your data center.
- Try out other Google Cloud Platform features for yourself. Have a look at our [tutorials](#).