stonebranch

Universal Task Documentation

Universal Automation Center support for scheduling Google Cloud Storage (GCS) file Transfers

ut-google-cloud-storage-linux

Associated Activities:

Date: 28 November 2019

Author: Nils Buer

Revision: 00

CONFIDENTIALITY INFORMATION

Distribution list:

Revision	Date	Author	Changes
00	20191129	Nils Buer	Initial Document (WIP)

Abstract:

The here described Universal Tasks allow to Transfer and retrieve files from Google Blob Storage in the Cloud. As a result, you can integrate any Google Blob Storage file transfer into you existing or new scheduling workflows, providing a true hybrid cloud (on-premise and cloud computer) file transfer solution.

Contents

1 Disclaimer 3

2 Scope 3

3 Introduction 3

4 Installation 4

4.1 Software Requirements 4

4.2 Installation Steps 5

5 Universal Task Configuration 6

6 Universal Tasks for Google Blob Storage 7

6.1 create-bucket 7

6.2 monitor-blob-in-bucket 8

6.3 copy-file-to-bucket 9

6.4 list-blobs-in-bucket 11

6.5 download-file-from-bucket 12

6.6 delete-blob-from-bucket 14

6.7 delete-empty-bucket 15

6.8 list-buckets 16

7 Test cases 17

8 Document references 18

1 Disclaimer

No support and no warranty are provided by Stonebranch GmbH for this document and the related Universal Task. The use of this document and the related Universal Task is on your own risk.

Before using this task in a production system, please perform extensive testing.

Stonebranch GmbH assumes no liability for damage caused by the performance of the Universal Tasks

2 Scope

This document provides a documentation how to install and use the Universal Tasks for Google Blob Storage File Transfers. If more Task will be created in the future this document will be updated accordingly.

3 Introduction

Storing data in the cloud becomes an integral part of most modern IT landscapes. With Universal Automation Center you can securely automate your AWS, Azure, Google or any other Cloud File Transfer and integrate them into your existing scheduling flows.

The here described Series of Universal Tasks focuses on the Google Blob Storage file transfer. A similar solution as for Google is also available for Amazon S3 or Azure Blob storage.

Some details about the universal tasks for Google Blob Storage:

- The Universal Tasks are calling the Microsoft Google “google-cloud-storage” API
- The python *google-cloud-storage* module is called by a Universal Agent running on a Linux Server or Windows Server – Note: This document focuses on the Linux Version
- The Server Running the Universal Agent needs to have the Python 3.6.x option installed
- All Universal Task support encrypted connections via a Proxy Server
- You can configure all connection Parameters for the Proxy and Google via the Universal Task
- You can select different log-levels e.g. Info and debug

The following Universal Task for Google Blob Storage have been implemented:

Command	UT Name	Description
create-bucket	ut-google-blobstorage-create-bucket-linux	Creates a bucket in Google Blob Storage
monitor-blob-in-bucket	ut-google-blobstorage-monitor-blob-in-bucket-linux	Monitors at a given interval for a blob in a bucket
copy-file-to-bucket	ut-google-blobstorage-copy-file-to-bucket-linux	Copies a file to a bucket
list-blobs-in-bucket	ut-google-blobstorage-list-blobs-in-bucket-linux	Lists all blobs in a bucket
download-file-from-bucket	ut-google-blobstorage-download-file-from-bucket-linux	Downloads a Google Blob to a local file

Command	UT Name	Description
delete-blob-from-bucket	ut-google-blobstorage-delete-blob-from-bucket-linux	Deletes a blob from a bucket
delete-empty-bucket	ut-google-blobstorage-delete-empty-bucket-linux	Deletes an empty bucket
list-buckets	ut-google-blobstorage-list-buckets-linux	List all buckets of a Google account

4 Installation

4.1 Software Requirements

Universal Task name: *ut-google-cloud-storage*

Related UAC XML Files for template and task: *Github repository*

Software used:

For the set-up you need:

1. Python 3.6 – installed as part of the Universal Agent
2. For Python the following modules are required:

    - ***Re,*** to support regular expression matching operations
    - ***glob,*** *to* find Unix pathnames matching a specified pattern
    - ***os,*** to support operating system dependent commands
    - ***sys***, for output re-direct processing
    - ***datetime, date*** and time stamps for messages
    - ***logging***, to provide logging capabilities for debug, info etc.
    - ***argparse***, to allow testing of the Universal TPL. script on the command line
    - ***google-cloud-storage***, The Google libraries for Python to use Google services and manage Google resources
    - ***fnmatch,*** module provides support for Unix shell-style wildcards
    - from google.cloud import storage, exceptions

    Note: Only the module ***google-cloud-storage*** *need to be added to python 3.6.x. e.g. using pip.*
    - *sudo /opt/universal/python3.6/bin/python3 -m pip install google-cloud-storage*

3. Universal Controller 6.4.5.x or higher
4. Universal Agent 6.4.2.2 or higher installed on a Linux Server
5. A Google account to try it out
6. Google keyfile to authenticate with your google account

4.2 Installation Steps

The following describes the installation steps:

1. **Check the current Python Version**

   python -V (Note: Captial "V")

   If your Version is Python 3.6 or later all is fine. If a no python or a lower Version has been installed upgrade your python Version or install the Universal Agent with the Python binding option (--python yes). This option will install python 3.6. along with your universal agent.

   e.g.

   sudo sh ./unvinst --network_provider oms --oms_servers 7878@192.168.88.12 --oms_port 7878 --oms_autostart no --ac_netname OPSAUTOCONF --opscli yes --python yes

   NOTE: the above install string does not work in case you did a user mode install

   Official Download link: https://www.python.org/downloads/

   Note: To check the current Python Version type in a shell: python -V (capital V)

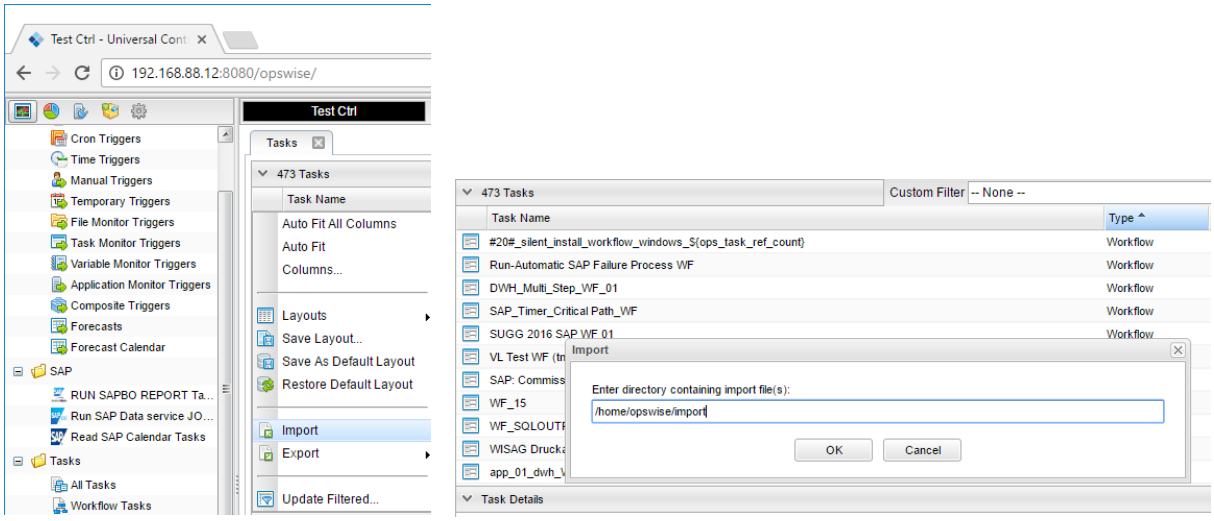
2. **Add the *google-cloud-storage* modules to your python installation**

   In a command shell run as root or sudo:

   - *sudo /opt/universal/python3.6/bin/python3 -m pip install google-cloud-storage*

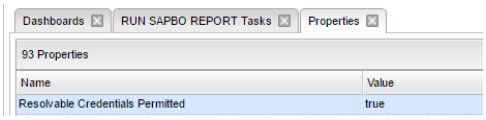
3. **Import each Google Blob Storage Universal Task including the Universal Template to your Controller**

Go to "All Tasks" and load via the Import functionality the Universal Task configuration into the Controller.



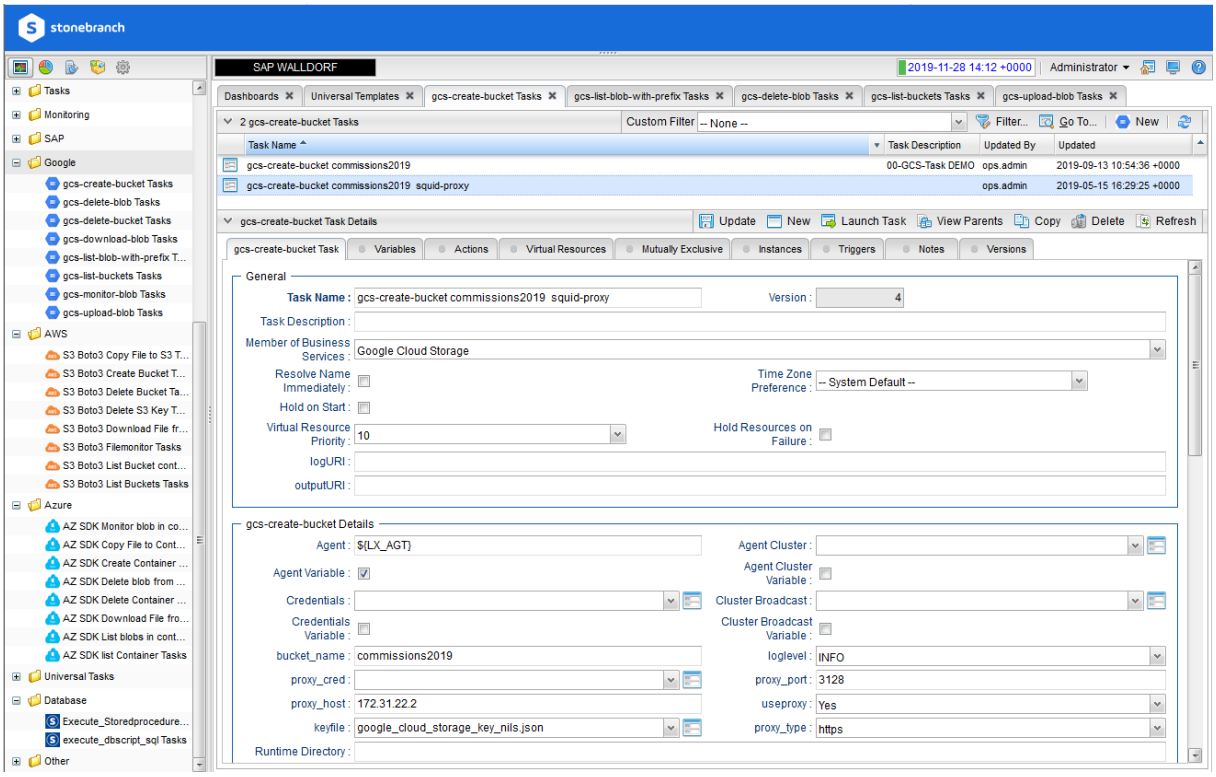
5 Universal Task Configuration

1. Activate: Resolvable Credentials in Universal Automation Center:



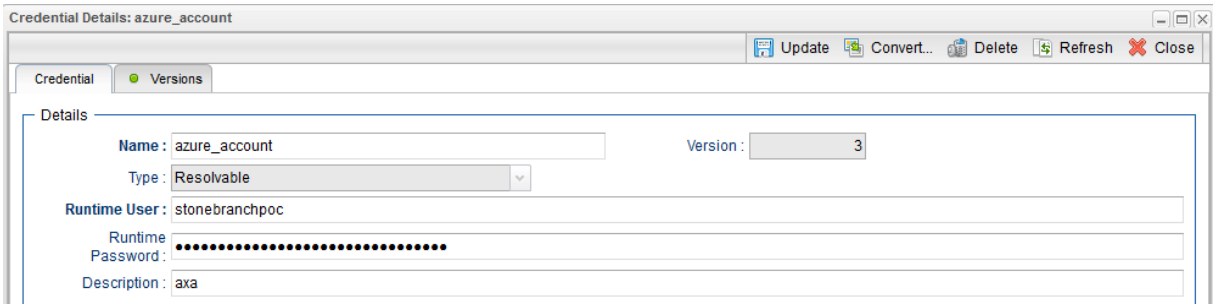
2. Fill Out the Universal Task for each Google Blob Storage command, which you want to execute:

In the example below the Google *Create bucket Task* was selected



Fill out or select the required Credentials for Google and optionally a Proxy Server

In the example below the Google_account credentials are shown:



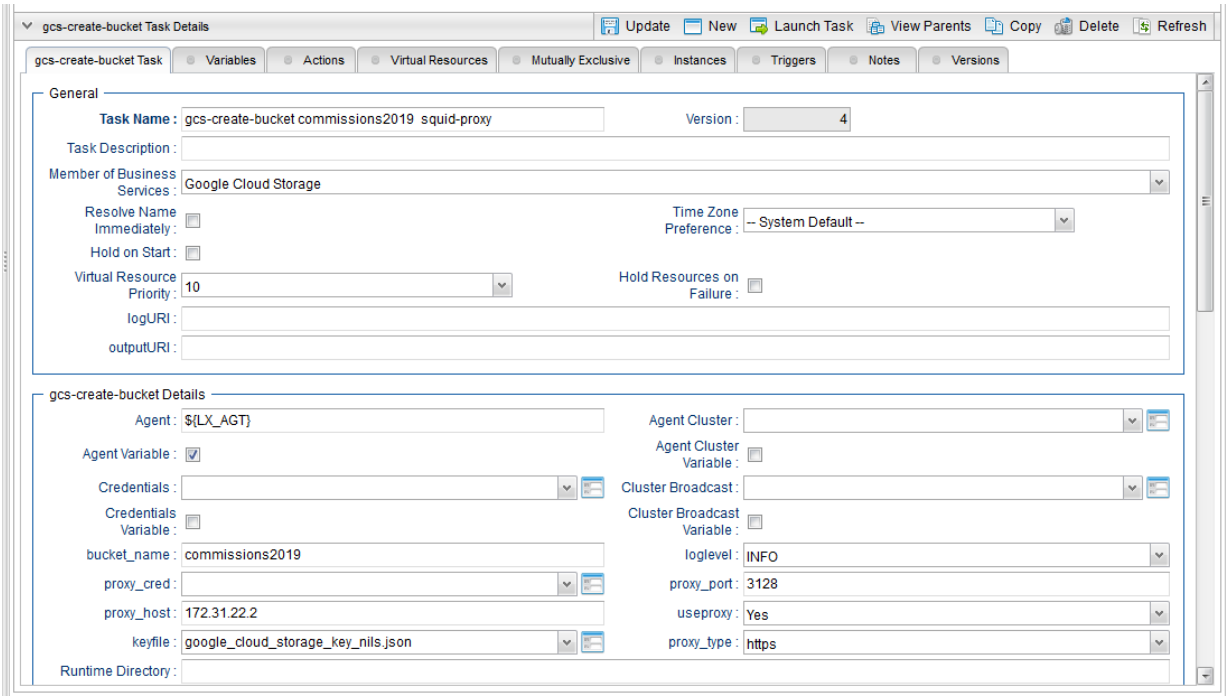
6 Universal Tasks for Google Blob Storage

The following chapter describes the provided Google Blob Storage Universal Tasks.

6.1 Create_Bucket

Command	UT Name	Description
create-bucket	ut-google-blobstorage-create-bucket-linux	Creates a bucket in GOOGLE BLOB STORAGE

Task Screenshot:



Field Description:

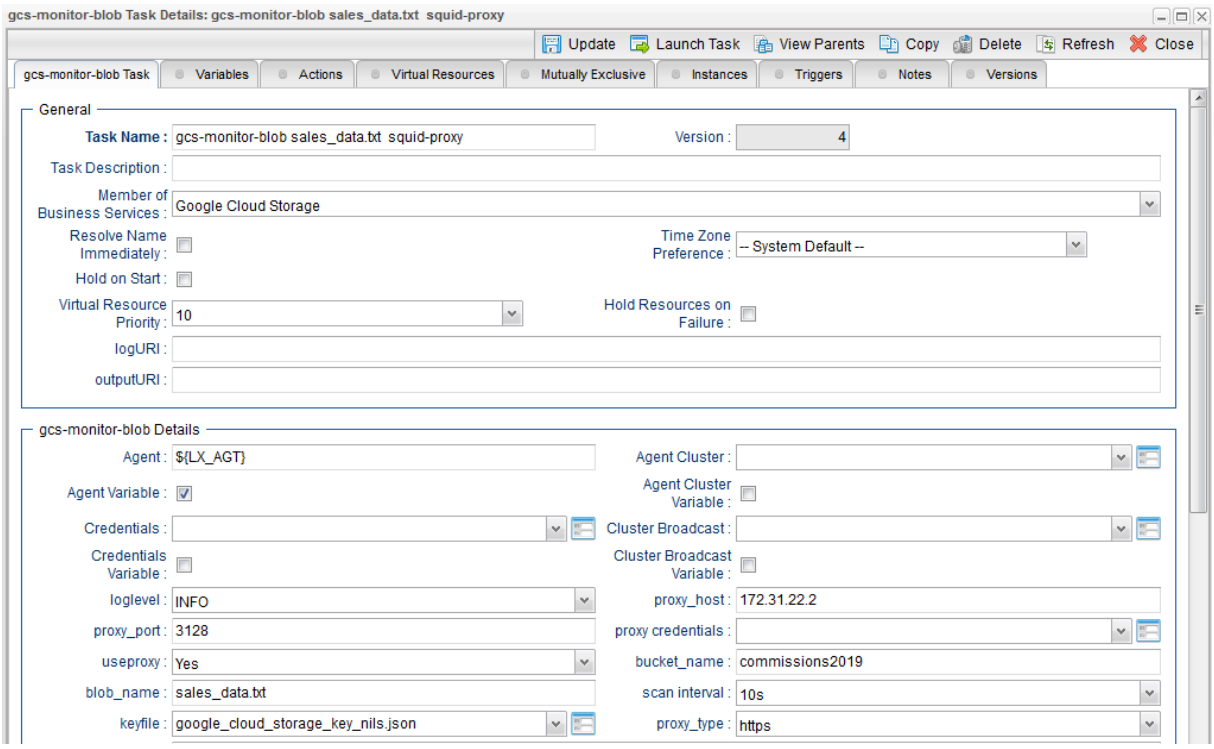
Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server
Use Proxy (yes/no)	Mandatory	Decide If a Proxy Server should be used yes or no (If “no” is selected the fields Proxy,Port and proxycred are ignored)
Proxy_host	Optional	Proxy Server IP or hostname (only used in case Use Proxy = yes)
Proxy credentials	Optional	Proxy Server Credentials (only used in case Use Proxy = “yes” is selected)

Field	Required	Description
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
Bucket-name	Mandatory	Bucket to be created. If the bucket already exists, the task will fail.
Keyfile	Mandatory	Link to a script in Universal Controller containing the keyfile

6.2 Monitor-blob-in-bucket

Command	UT Name	Description
monitor-blob-in-bucket	ut-google-blobstorage-monitor-blob-in-bucket-linux	Monitors at a given interval for a blob in a bucket

Task Screenshot:



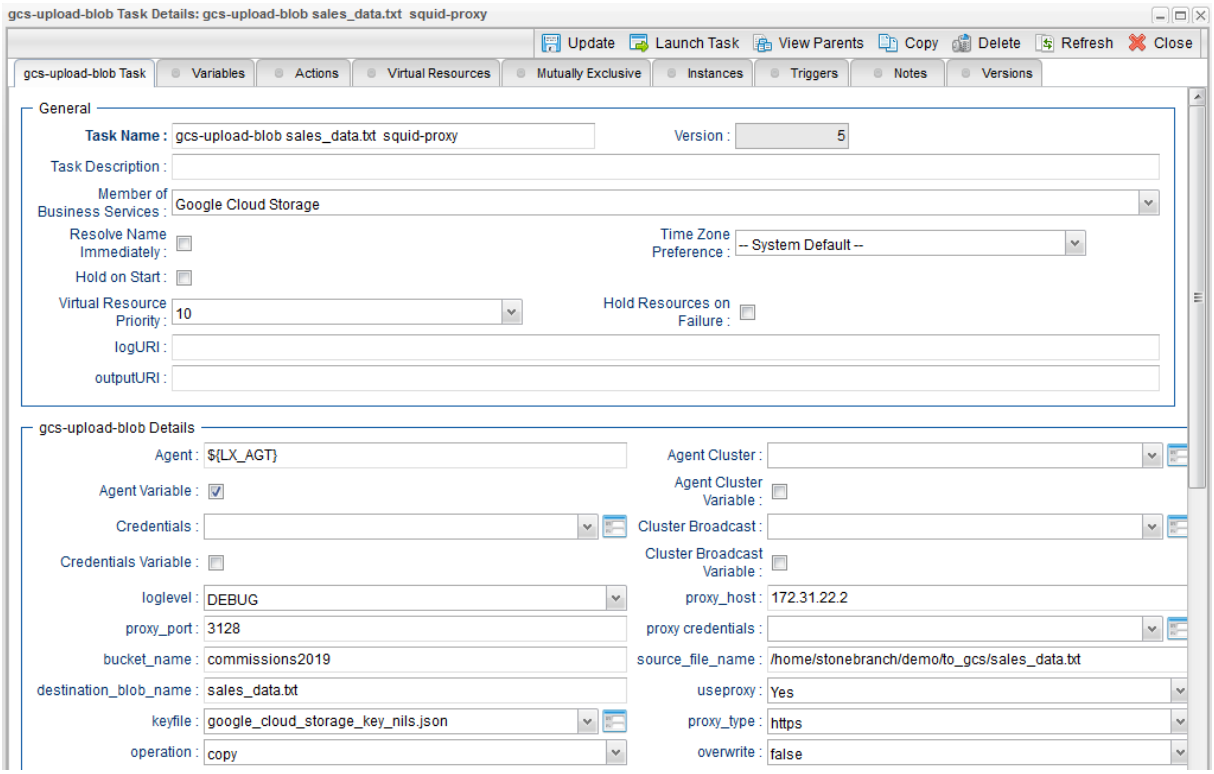
Field Description:

Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server
Use Proxy (yes/no)	Mandatory	Decide If a Proxy Server should be used yes or no (If “no” is selected the fields Proxy,Port and proxycred are ignored)
Proxy_host	Optional	Proxy Server IP or hostname (only used in case Use Proxy = yes)
Proxy credentials	Optional	Proxy Server Credentials (only used in case Use Proxy = “yes” is selected)
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
File to scan	Mandatory	GOOGLE blob (=file) to scan in the given bucket
Bucket_name	Mandatory	GOOGLE Bucket to scan for the given GOOGLE blob (=file)
Scan Interval	Mandatory	Scan Interval e.g. every 10s
Keyfile	Mandatory	Link to a script in Universal Controller containing the keyfile

6.3 Copy-file-to-bucket

Command	UT Name	Description
copy-file-to-bucket	ut-google-blobstorage-copy-file-to-bucket-linux	Copies a file to a bucket

Task Screenshot:



Field Description:

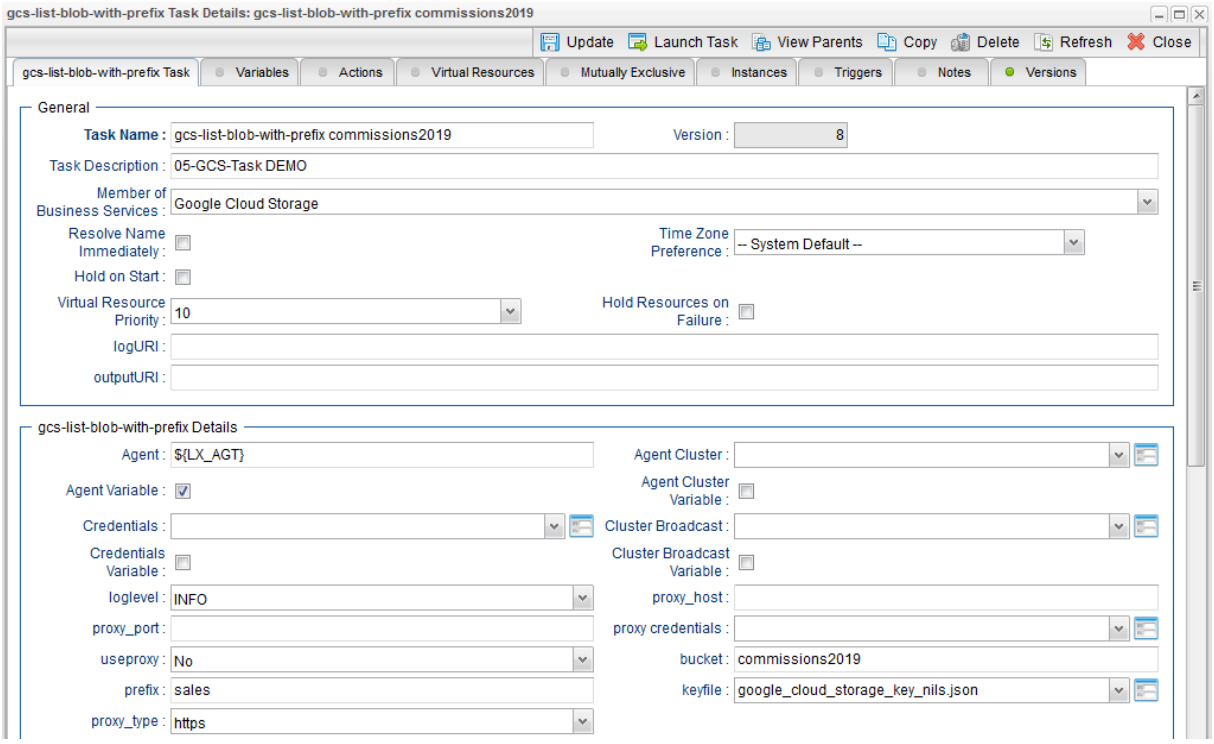
Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server
Use Proxy (yes/no)	Mandatory	Decide If a Proxy Server should be used yes or no (If “no” is selected the fields Proxy,Port and proxycred are ignored)
Proxy_host	Optional	Proxy Server IP or hostname (only used in case Use Proxy = yes)
Proxy credentials	Optional	Proxy Server Credentials (only used in case Use Proxy = “yes” is selected)
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
Bucket_name	Optional	Name of the target bucket
Overwrite	Mandatory	If checked, allow to overwrite an existing blob in the given bucket
Source_Blob_name	Mandatory	file to copy to the given bucket

Field	Required	Description
Destination_Blob_name	Optional	blob name in the given bucket (if empty the source file name will be kept)
Overwrite	Mandatory	• Timestamp => if file exists add a copy with a timestamp at the end • False => do not overwrite • True => overwrite
Operation	Mandatory	• Copy => keep file on AWS after download • Move => delete file on AWS after download
Keyfile	Mandatory	• Link to a script in Universal Controller containing the keyfile

6.4 List-blobs-in-bucket

Command	UT Name	Description
list-blobs-in-bucket	ut-google-blobstorage-list-blobs-in-bucket-linux	Lists all blobs in a bucket

Task Screenshot:



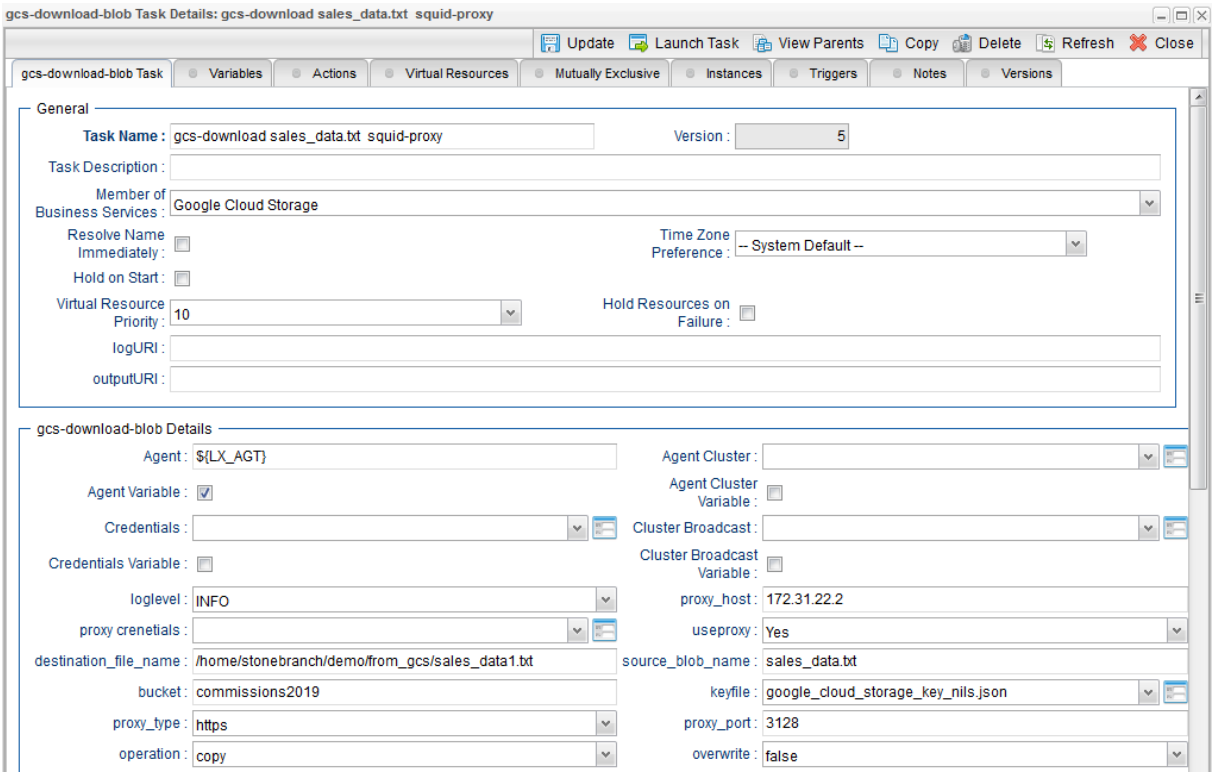
Field Description:

Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server
Use Proxy (yes/no)	Mandatory	Decide If a Proxy Server should be used yes or no (If “no” is selected the fields Proxy, Port and proxycred are ignored)
Proxy_host	Optional	Proxy Server IP or hostname (only used in case Use Proxy = yes)
Proxy credentials	Optional	Proxy Server Credentials (only used in case Use Proxy = “yes” is selected)
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
Bucket	Mandatory	List all blobs in the bucket
Keyfile	Mandatory	Link to a script in Universal Controller containing the keyfile

6.5 Download-file-from-bucket

Command	UT Name	Description
download-file-from-bucket	ut-google-blobstorage-download-file-from-bucket-linux	Downloads an GOOGLE BLOB STORAGE blob to a local file

Task Screenshot:



Field Description:

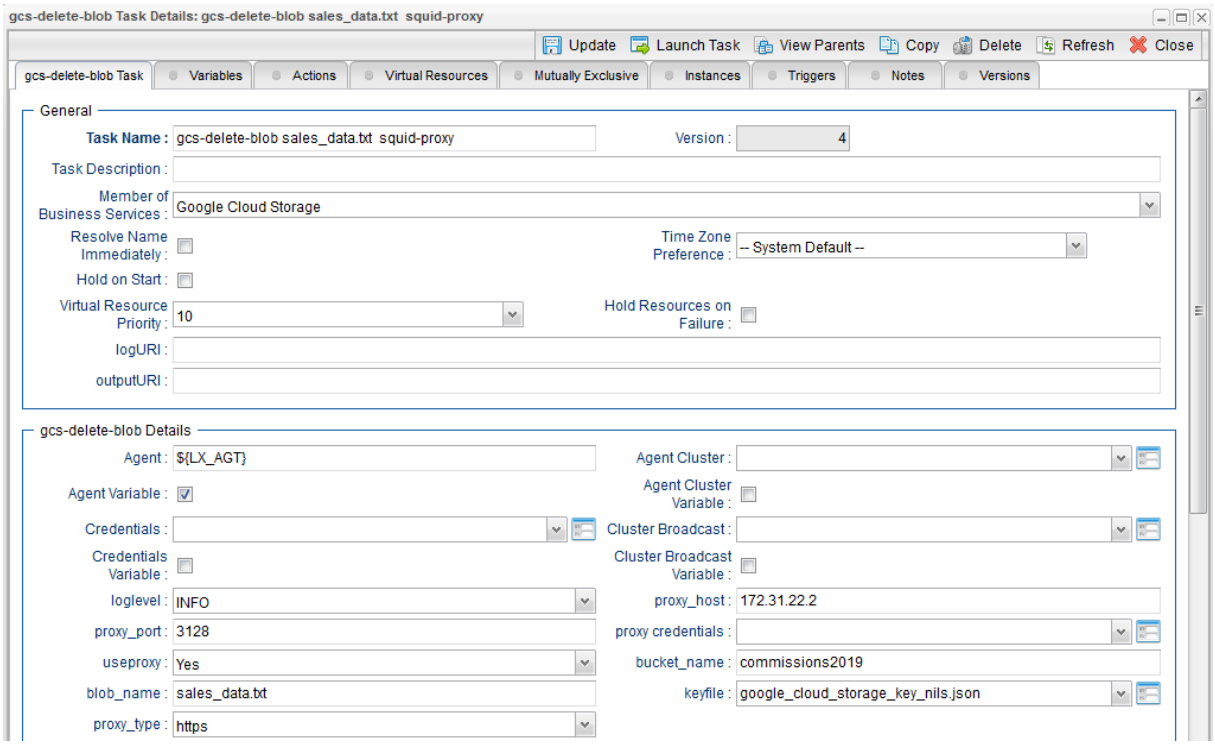
Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server
Use Proxy (yes/no)	Mandatory	Decide If a Proxy Server should be used yes or no (If “no” is selected the fields Proxy,Port and proxycred are ignored)
Proxy_host	Optional	Proxy Server IP or hostname (only used in case Use Proxy = yes)
Proxy credentials	Optional	Proxy Server Credentials (only used in case Use Proxy = “yes” is selected)
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
Destination_file_name	Mandatory	Target file name and directory e.g. */home/stonebranch/demo/from_gcs/sales_data1.txt*
Source_blob_name	Mandatory	Source Blob file to download from given bucket
Bucket	Mandatory	Source bucket
Overwrite	Mandatory	• Timestamp => if file exists add a copy with a timestamp at the end

Field	Required	Description
		• False => do not overwrite • True => overwrite
Operation	Mandatory	• Copy => keep file on AWS after download • Move => delete file on AWS after download
Keyfile	Mandatory	• Link to a script in Universal Controller containing the keyfile

6.6 Delete-blob-from-bucket

Command	UT Name	Description
delete-blob-from-bucket	ut-google-blobstorage-delete-blob-from-bucket-linux	Deletes a blob from a bucket

Task Screenshot:



Field Description:

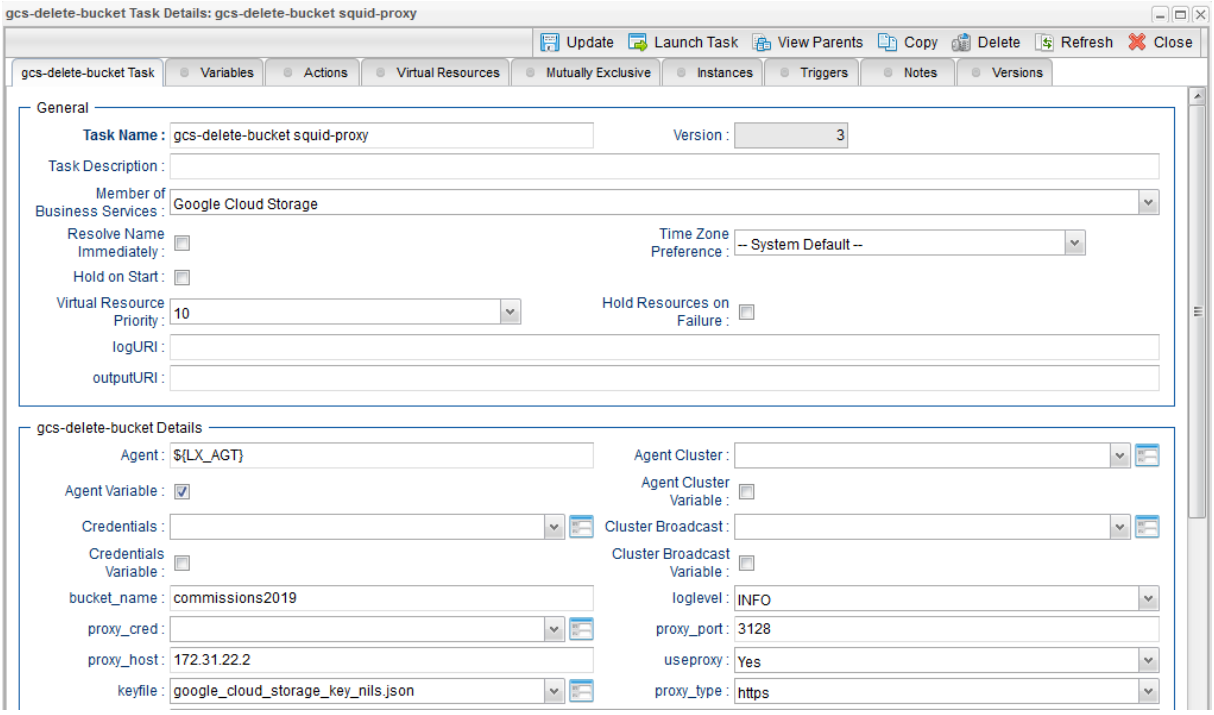
Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server

Field	Required	Description
Use Proxy (yes/no)	Mandatory	Decide If a Proxy Server should be used yes or no (If “no” is selected the fields Proxy,Port and proxycred are ignored)
Proxy	Optional	Proxy Server IP/ hostname (used if Use Proxy = “yes”)
Proxy_credentials	Optional	Proxy Server Credentials (used if Use Proxy = “yes”)
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
Blob_name	Mandatory	Blob to delete from the given bucket
Bucket_name	Mandatory	Bucket, which contains the blob to delete
Keyfile	Mandatory	Link to a script in Universal Controller containing the keyfile

6.7 Delete-empty-bucket

Command	UT Name	Description
delete-empty-bucket	ut-google-blobstorage-delete-empty-bucket-linux	Deletes an empty bucket

Task Screenshot:



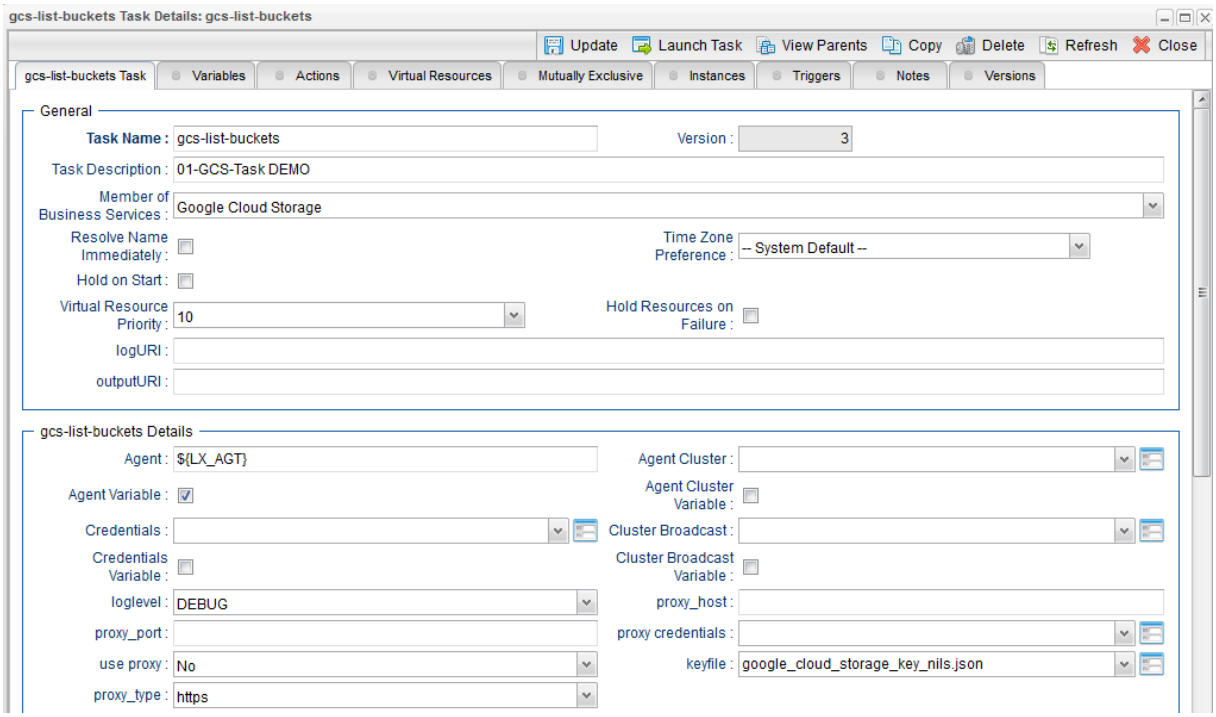
Field Description:

Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server
Use Proxy (yes/no)	Mandatory	Decide If a Proxy Server should be used yes or no (If “no” is selected the fields Proxy,Port and proxycred are ignored)
Proxy_host	Optional	Proxy Server IP/ hostname (used if Use Proxy = “yes”)
Proxy_cred	Optional	Proxy Server Credentials (used if Use Proxy = “yes”)
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
Bucket_name	Mandatory	Bucket to delete
Keyfile	Mandatory	Link to a script in Universal Controller containing the keyfile

6.8 List-buckets

Command	UT Name	Description
list-buckets	ut-google-blobstorage-list-buckets-linux	List all buckets of a Google account

Task Screenshot:



Field Description:

Field	Required	Description
Agent	Mandatory	The Linux Universal Agent, which runs the Python Google-storage-blob module to call the GOOGLE BLOB STORAGE commands
Credentials	Optional	The Credentials used on the Linux Server
Proxy	Optional	Proxy Server IP/ hostname (used if Use Proxy = “yes”)
Proxycred	Optional	Proxy Server Credentials (used if Use Proxy = “yes”)
Proxy_Port	Optional	Proxy Server Port (only used in case Use Proxy = yes)
Loglevel	Mandatory	logging settings DEBUG, INFO, WARNING, ERROR, CRITICAL
Keyfile	Mandatory	Link to a script in Universal Controller containing the keyfile

7 Test Cases

The following basic test cases has been performed:

Case#	Result
Create a bucket in GOOGLE BLOB STORAGE	Correct
Creates a bucket in GOOGLE BLOB STORAGE (bucket already exists)	Correct
Monitors at a given interval (10s) for a blob in a bucket	Correct

Copies a file to a bucket, if the blob exists (flag overwrite is set)	Correct
Copies a file to a bucket, if the blob exists (flag overwrite is not set)	Correct
Copies a file to a bucket (file does not exist in bucket)	Correct
Copies a file to a bucket (bucket does not exist)	Correct
Lists all blobs in a bucket	Correct
Downloads a GOOGLE BLOB STORAGE blob to a local file if file does not yet exist	Correct
Downloads a GOOGLE BLOB STORAGE blob to a local file if already exist (flag overwrite is set)	Correct
Downloads a GOOGLE BLOB STORAGE blob to a local file if already exist (flag overwrite is not set)	Correct
Downloads a GOOGLE BLOB STORAGE blob to a local file if file does not yet exist and overwrite flag is set.	Correct
Deletes a blob from a bucket	Correct
Deletes a blob from a bucket (blob does not exist)	
Deletes an empty bucket	Correct
Deletes a bucket, which does not exist	Correct
List all buckets of a GOOGLE account	Correct
Checks for the existence of a blob in a bucket	Correct
Checks for the existence of a blob in a bucket, if blob does not exist	Correct

8 Document References

There are no document references.