

Upgrading Cloudera Data Engineering

Date published: 2020-07-30

Date modified: 2025-07-30



Legal Notice

© Cloudera Inc. 2025. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Cloudera Data Engineering upgrade version compatibility.....	4
Security hardened Spark image migration guide.....	5
Libraries not included in the security hardened image.....	5
Cloudera Data Engineering upgrade scenarios.....	15
Upgrade considerations if Cloudera Data Engineering services include Apache Airflow workloads.....	15
Upgrading Cloudera Data Engineering.....	16
Upgrading to Cloudera Data Lake 7.3.1 with Cloudera Data Engineering.....	18
In-place upgrade with Airflow Operators and Libraries.....	20
Airflow constraints file.....	23
Troubleshooting Python or Airflow compatibility issues.....	57
Debugging Airflow DAGs locally.....	59
Using the backup-restore-based upgrade script.....	61
Upgrading Airflow if DAGs and packages are compatible with new Airflow version.....	64
Handling upgrade failures for Cloudera Data Engineering.....	64
Handling ineligible upgrades for Cloudera Data Engineering.....	67

Cloudera Data Engineering upgrade version compatibility

Cloudera Data Engineering provides the in-place upgrade option from the eligible versions to the latest version. If you are on a version that is ineligible for in-place upgrade, the in-place upgrade option is disabled and you can perform a backup and restore.

Checking your upgrade options

To check if your Cloudera Data Engineering version is eligible for the in-place upgrade option, see [Support lifecycle policy](#).

In-place upgrades

To perform an in-place upgrade, you must be on a version eligible for the in-place upgrade option. The *Cloudera Data Engineering versions eligible for in-place upgrade* table includes the specific AWS and Azure versions from which you can perform an in-place upgrade to the latest Cloudera Data Engineering version.

Table 1: Cloudera Data Engineering versions eligible for in-place upgrade

Supported upgrade versions	Target version
Azure: 1.23.0 and higher AWS: 1.23.0 and higher	1.24.1



Important: Using AWS, Cloudera Data Engineering supports in-place upgrades from version 1.20.3 to the latest version, which includes several Kubernetes version bumps. For this reason, you can expect the upgrade to take longer than usual, with an upgrade time of approximately 3 hours.



Note: If you use services that run on a private network, your Cloudera Data Engineering service is eligible for an in-place upgrade from Cloudera Data Engineering version 1.22.0. onwards.

Backup and restore

If you are on a version that is not eligible for the in-place upgrade option, to upgrade to the latest version, create a new Cloudera Data Engineering service and transfer your jobs to the new service by following the instructions in *Backing up and restoring Cloudera Data Engineering jobs*.

Runtime component version behavior after upgrading

To maintain backwards compatibility, the in-place upgrade does not modify the image type, and after the upgrade to Cloudera Data Engineering version 1.24.1, the old RedHat (deprecated) images continue to be used. To use the new, security hardened image, after the upgrade, you must create new Virtual Clusters and select the new image. For more information on the supported runtime components, see [Compatibility for Cloudera Data Engineering and Runtime components](#).



Important: If you switch from the Redhat UBI image to the security hardened image, note that the updated component versions and libraries can impact the Cloudera Data Engineering operation. Cloudera strongly recommends testing the new implementation in a non-production environment first.

For backward compatibility, you can continue using the Redhat UBI image. The security hardened image is the recommended image, however, note that it requires testing and validation.

For more information, see [Security hardened Spark image migration guide](#).

Related Information

[In-place Upgrade with Airflow Operators and Libraries](#)

[Backing up Cloudera Data Engineering jobs on remote storage](#)

Security hardened Spark image migration guide

From Cloudera Data Engineering version 1.24.1, the Redhat-based images for Apache Spark 3.x versions are deprecated and security hardened images are provided for runtime usage.

To support successful migrations during in-place upgrades and backup-and-restore operations for clusters from previous versions, Redhat images are continued to be used. When manually creating clusters through the Cloudera Data Engineering UI, the API, or the CDP CLI, the default option is the security hardened image.

The security hardened image differs from the Redhat-based images in the following aspects:

- Different Java and Python versions are used. For more information, see [Compatibility for Cloudera Data Engineering and Runtime components](#).
- The security hardened image follows a distroless approach, which excludes packages that are not required for running Spark. Several Linux packages which are present in the Redhat-based images are not included in the security hardened image. As a consequence, any Python or Java module that has a dependency on these excluded packages causes issues in the security hardened image. For more information, see [Libraries not included in the security hardened image](#).



Note: Cloudera strongly recommends testing the applications in the security hardened Virtual Cluster, in a non-production environment first, before using them in production.

For any packages excluded from the security hardened image, missing for Python or Java modules, you can build a custom image with the required libraries from the security hardened base Spark runtime image. For more information, see [Using Custom Spark Runtime Docker Images via API/CLI](#).

Libraries not included in the security hardened image

The security hardened image follows a distroless approach, which excludes packages that are not required for running Spark. The excluded packages are outlined in the following list.

Package	Description
abattis-cantarell-fonts	Default GNOME Cantarell font family for UI text.
acl	System package acl.
adwaita-cursor-theme	GNOME Adwaita cursor theme.
adwaita-icon-theme	GNOME Adwaita icon theme.
apr	Apache Portable Runtime core libraries.
apr-devel	Development headers and static libraries for apr.
apr-util	Utility libraries for the Apache Portable Runtime (APR).
apr-util-bdb	System package apr util bdb.
apr-util-devel	Development headers and static libraries for apr util.
apr-util-ldap	System package apr util ldap.
apr-util-openssl	System package apr util openssl.
at-spi2-atk	System package at spi2 atk.
at-spi2-core	System package at spi2 core.
atk	System package atk.
atlas	System package atlas.
atlas-devel	Development headers and static libraries for atlas.

Package	Description
audit-libs	Shared libraries for audit.
avahi-libs	Shared libraries for avahi.
basesystem	Basic system identification package.
bind-libs-lite	System package bind libs lite.
bind-license	System package bind license.
bind-utils	System package bind utils.
brotli	System package brotli.
bzip2-libs	Shared libraries for bzip2.
cairo	2D vector graphics library.
cairo-gobject	System package cairo gobject.
chkconfig	System package chkconfig.
cmake	System package cmake.
cmake-data	Data files for cmake.
cmake-filesystem	Filesystem layout for cmake.
cmake-rpm-macros	System package cmake rpm macros.
colord-libs	Color management libraries.
copy-jdk-configs	Helpers to manage system JDK configuration.
coreutils-single	Core GNU utilities in a single binary.
cracklib	System package cracklib.
cracklib-dicts	System package cracklib dicts.
crypto-policies	System package crypto policies.
crypto-policies-scripts	System package crypto policies scripts.
cryptsetup-libs	Shared libraries for cryptsetup.
cups-libs	CUPS printing system libraries.
cyrus-sasl-devel	Development headers for Cyrus SASL.
cyrus-sasl-lib	SASL authentication library.
dbus	System message bus (D-Bus).
dbus-common	Common files for dbus.
dbus-daemon	D-Bus message bus daemon.
dbus-glib	GLib bindings for D-Bus.
dbus-libs	Libraries for D-Bus.
dbus-tools	Command-line tools for D-Bus.
dconf	System package dconf.
dejavu-fonts-common	Shared files for the DejaVu font families.
dejavu-sans-mono-fonts	DejaVu Sans Mono monospaced font family.
device-mapper	System package device mapper.
device-mapper-libs	Shared libraries for device mapper.
dmidecode	System package dmidecode.
dwz	DWARF debug information compressor.

Package	Description
efi-srpm-macros	System package efi srpm macros.
elfutils-default-yama-scope	System package elfutils default yama scope.
elfutils-libelf	System package elfutils libelf.
elfutils-libs	Shared libraries for elfutils.
emacs-filesystem	Filesystem layout for emacs.
enchant	System package enchant.
expat	System package expat.
expat-devel	Development headers and static libraries for expat.
file	File type identification utility.
file-libs	Libraries for file type identification.
filesystem	Base filesystem layout and directories.
findutils	File-finding utilities (find, xargs).
fontconfig	System package fontconfig.
fontpackages-filesystem	Filesystem layout for fontpackages.
fribidi	Bidirectional text processing library.
gawk	GNU awk text processing language.
gdb-gdbserver	Remote debugging server for GDB.
gdbm-libs	Shared libraries for gdbm.
gdk-pixbuf2	Image loading and pixel buffer library for GNOME.
gdk-pixbuf2-modules	Image loader modules for GDK-Pixbuf.
geolite2-city	GeoLite2 city geolocation database.
geolite2-country	GeoLite2 country geolocation database.
ghc-srpm-macros	RPM macros for building Haskell packages.
glib-networking	System package glib networking.
glib2	Core application utility library for GNOME/GTK (GLib).
glibc-common	Common files for the GNU C Library.
glibc-gconv-extra	Extra character set conversions for glibc.
glibc-langpack-en	English language support for glibc.
glibc-minimal-langpack	Minimal language data for glibc.
gmp	System package gmp.
gnupg2	System package gnupg2.
gnutls	TLS/SSL and crypto library (GnuTLS).
go-srpm-macros	RPM macros for building Go packages.
gobject-introspection	System package gobject introspection.
gpgme	System package gpgme.
graphite2	Smart font rendering engine.
grep	Text search utility.
groff-base	Typesetting tools (man page formatter).
gsettings-desktop-schemas	GNOME desktop settings schemas.

Package	Description
gtk-update-icon-cache	System package gtk update icon cache.
gtk3	GTK 3 graphical toolkit.
gzip	GNU gzip compression utility.
harfbuzz	OpenType text shaping engine.
hicolor-icon-theme	Fallback icon theme shared by desktops.
hostname	Utility to set or show the system hostname.
httpd	Apache HTTP Server.
httpd-devel	Development headers for building against Apache HTTP Server.
httpd-filesystem	Filesystem layout and directories for Apache HTTP Server.
httpd-tools	Command-line utilities for Apache HTTP Server.
hunspell	System package hunspell.
hunspell-en	System package hunspell en.
hunspell-en-gb	System package hunspell en gb.
hunspell-en-us	System package hunspell en us.
ima-evm-utils	System package ima evm utils.
info	GNU info documentation viewer.
jasper-libs	JPEG-2000 image codec library.
java-11-openjdk	OpenJDK 11 JRE.
java-11-openjdk-devel	OpenJDK 11 development tools and headers.
java-11-openjdk-headless	OpenJDK 11 headless runtime.
javapackages-filestystem	Filesystem layout for Java packaging.
jbigkit-libs	Shared libraries for jbigkit.
json-glib	JSON parser and generator for GLib.
kmod-libs	Shared libraries for kmod.
krb5-workstation	Kerberos 5 client utilities (kinit, klist, etc.).
langpacks-en	System package langpacks en.
libacl	Shared library libacl.
libarchive	Shared library libarchive.
libassuan	Shared library libassuan.
libattr	Shared library libattr.
libblkid	Shared library libblkid.
libcap	Shared library libcap.
libcap-ng	Shared library libcap ng.
libcomps	Shared library libcomps.
libcurl	cURL client library for HTTP/FTP and more.
libdatrie	Shared library libdatrie.
libdb	Shared library libdb.
libdb-devel	Development headers and static libraries for libdb.
libdb-utils	Shared library libdb utils.

Package	Description
libdnf	Shared library libdnf.
libepoxy	OpenGL function pointer management library.
libfdisk	Shared library libfdisk.
libffi-devel	Development headers and static libraries for libffi.
libfontenc	Shared library libfontenc.
libgcrypt	General-purpose cryptographic library.
libgpg-error	GnuPG error description library.
libgusb	Shared library libgusb.
libkadm5	Kerberos administration library.
libksba	Shared library libksba.
libmaxminddb	MaxMind DB reader library (GeoIP2).
libmodman	Shared library libmodman.
libmodulemd	Shared library libmodulemd.
libmount	Shared library libmount.
libnghttp2	Shared library libnghttp2.
libnl3	Shared library libnl3.
libns12	Shared library libns12.
libpath_utils	Shared library libpath utils.
libpkgconf	Shared library libpkgconf.
libproxy	Proxy configuration and discovery library.
libpwquality	Shared library libpwquality.
libquadmath	Shared library libquadmath.
librepo	Shared library librepo.
libreport-filesystem	Filesystem layout for problem reporting (ABRT).
librhsm	Shared library librhsm.
libseccomp	Secure computing (seccomp) filtering library.
libsemanage	SELinux policy management library.
libsigsegv	Segmentation fault handling library.
libsmartcols	Column formatting library from util-linux.
libsolv	Dependency solver library for package managers.
libsoup	Shared library libsoup.
libss	Shared library libss.
libssh	SSH client library.
libssh-config	Configuration files for libssh.
libtalloc	Hierarchical memory allocator (Samba/tdb).
libthai	Thai language support library.
libtiff	TIFF image library.
libtirpc	Shared library libtirpc.
libtool-ltdl	Shared library libtool ltdl.

Package	Description
libusb	Userspace USB programming library.
libuser	Shared library libuser.
libutempter	Shared library libutempter.
libuuid	UUID generation library.
libwayland-client	Wayland client library.
libwayland-cursor	Wayland cursor helper library.
libwayland-egl	Wayland EGL platform library.
libx11-common	Common data files for the X11 client libraries.
libxcomposite	X11 Composite extension client library.
libxcursor	X11 cursor management client library.
libxdamage	X11 Damage extension client library.
libxfixes	X11 Fixes extension client library.
libxft	X FreeType interface library for X11.
libxinerama	X11 multi-head display extension library.
libxkbcommon	Keyboard handling library for Wayland/X11.
libxrandr	X11 RandR extension client library.
libyaml	YAML parser and emitter library.
libzstd	Zstandard compression library.
lksctp-tools	Command-line tools for lksctp.
lua	System package lua.
lua-libs	Shared libraries for lua.
lz4-libs	LZ4 compression library.
mailcap	MIME type handling configuration.
make	Build automation tool.
mod_auth_gssapi	System package mod auth gssapi.
mod_http2	HTTP/2 module for Apache HTTP Server.
mod_ldap	LDAP authentication/authorization module for Apache HTTP Server.
mod_session	Session management module for Apache HTTP Server.
mod_ssl	SSL/TLS module for Apache HTTP Server.
mpfr	System package mpfr.
ncurses-base	System package ncurses base.
ncurses-libs	Shared libraries for ncurses.
nettle	System package nettle.
npth	New GNU portable threads library.
nspr	Netscape Portable Runtime library.
nss	Network Security Services cryptography libraries.
nss-softhkn	NSS software cryptographic module.
nss-softhkn-freebl	NSS low-level crypto library (freebl).
nss-sysinit	NSS system initialization module.

Package	Description
nss-util	Utility libraries for NSS.
nss_wrapper	Name Service Switch wrapper libraries.
nss_wrapper-libs	Shared libraries for nss wrapper.
ocaml-srpm-macros	RPM macros for building OCaml packages.
openblas-srpm-macros	RPM macros for packaging OpenBLAS.
openblas-threads	Optimized BLAS library (OpenBLAS).
openldap	OpenLDAP client utilities and libraries.
openldap-devel	Development headers for OpenLDAP.
openssl-libs	OpenSSL cryptographic and TLS libraries.
pam	System package pam.
pango	Text layout and rendering library.
passwd	Password change utility.
pcre	System package pcre.
pcre2	System package pcre2.
perl-carp	Perl module carp.
perl-constant	Perl module constant.
perl-data-dumper	Perl module data dumper.
perl-digest	Perl module digest.
perl-digest-md5	Perl module digest md5.
perl-encode	Perl module encode.
perl-encoding	Perl module encoding.
perl-errno	Perl module errno.
perl-exporter	Perl module exporter.
perl-file-path	Perl module file path.
perl-file-temp	Perl module file temp.
perl-filter	Perl module filter.
perl-getopt-long	Perl module getopt long.
perl-http-tiny	Perl module http tiny.
perl-interpreter	Perl module interpreter.
perl-io	Perl module io.
perl-io-socket-ip	Perl module io socket ip.
perl-io-socket-ssl	Perl module io socket ssl.
perl-libnet	Perl module libnet.
perl-libs	Shared libraries for perl.
perl-macros	Perl module macros.
perl-mime-base64	Perl module mime base64.
perl-mozilla-ca	Perl module mozilla ca.
perl-net-ssleay	Perl module net ssleay.
perl-open	Perl module open.

Package	Description
perl-parent	Perl module parent.
perl-pathtools	Perl module pathtools.
perl-pod-escapes	Perl module pod escapes.
perl-pod-perldoc	Perl module pod perldoc.
perl-pod-simple	Perl module pod simple.
perl-pod-usage	Perl module pod usage.
perl-podlators	Perl module podlators.
perl-scalar-list-utils	Perl module scalar list utils.
perl-socket	Perl module socket.
perl-srpm-macros	Perl module srpm macros.
perl-storable	Perl module storable.
perl-term-ansicolor	Perl module term ansicolor.
perl-term-cap	Perl module term cap.
perl-text-parsewords	Perl module text parsewords.
perl-text-tabs+wrap	Perl module text tabs and wrap.
perl-threads	Perl module threads.
perl-threads-shared	Perl module threads shared.
perl-time-local	Perl module time local.
perl-unicode-normalize	Perl module unicode normalize.
perl-uri	Perl module uri.
pixman	Low-level pixel manipulation library.
pkgconf	Tool to query and manage library build flags (pkg-config).
pkgconf-m4	Autoconf macros for pkgconf.
pkgconf-pkg-config	pkg-config compatibility symlink provided by pkgconf.
platform-python	System Python 3 interpreter used by the OS.
platform-python-setuptools	Setuptools for the system Python.
popt	Command-line option parsing library.
procps-ng	System package procps ng.
publicsuffix-list-dafsa	System package publicsuffix list dafsa.
python-rpm-macros	System package python rpm macros.
python-srpm-macros	System package python srpm macros.
python2	Python 2 interpreter (legacy).
python2-devel	Development headers for Python 2.
python2-libs	Shared libraries for Python 2.
python2-nose	Unit testing framework for Python 2.
python2-numpy	NumPy for Python 2.
python2-pip	pip package manager for Python 2.
python2-pip-wheel	pip as a wheel for Python 2.
python2-rpm-macros	RPM macros for Python 2 packaging.

Package	Description
python2-setuptools	Python 2 packaging utilities (setuptools).
python2-setuptools-wheel	setuptools as a wheel for Python 2.
python2-virtualenv	Virtual environments for Python 2.
python2-wheel-wheel	wheel package as a wheel for Python 2.
python3-bind	BIND DNS Python 3 tools.
python3-chardet	Character encoding detection for Python 3.
python3-cloud-what	Cloud environment detection helpers for Python 3.
python3-dateutil	Python 3 package dateutil.
python3-dbus	Python 3 package dbus.
python3-decorator	Python 3 package decorator.
python3-dnf	DNF Python 3 bindings.
python3-dnf-plugins-core	Core DNF plugin collection for Python 3.
python3-ethtool	ethtool bindings for Python 3.
python3-gobject-base	GObject introspection bindings for Python 3.
python3-gpg	GnuPG bindings for Python 3.
python3-hawkey	libhawkey bindings for Python 3 (DNF).
python3-idna	IDNA (Internationalized Domain Names) support for Python 3.
python3-iniparse	INI file parser for Python 3.
python3-inotify	inotify bindings for Python 3.
python3-libcomps	Python 3 package libcomps.
python3-libdnf	libdnf bindings for Python 3.
python3-librepo	librepo bindings for Python 3.
python3-libs	Shared libraries for python3.
python3-pip-wheel	Python 3 package pip wheel.
python3-ply	Lex/Yacc parsing tools for Python 3.
python3-pysocks	SOCKS proxy client for Python 3.
python3-requests	HTTP library for Python 3.
python3-rpm	RPM Python 3 bindings.
python3-rpm-generators	RPM dependency generators for Python 3.
python3-rpm-macros	RPM macros for Python 3 packaging.
python3-setuptools-wheel	Setuptools as a wheel for Python 3.
python3-six	Python 3 package six.
python3-subscription-manager-rhsm	Python 3 package subscription manager rhsm.
python3-syspurpose	Python 3 package syspurpose.
python3-systemd	systemd bindings for Python 3.
python3-urllib3	HTTP client library for Python 3.
python38	Python 3.8 interpreter.
python38-devel	Development headers for Python 3.8.
python38-libs	Shared libraries for Python 3.8.

Package	Description
python38-numpy	NumPy for Python 3.8.
python38-pip	pip package manager for Python 3.8.
python38-pip-wheel	pip as a wheel for Python 3.8.
python38-setuptools	setuptools for Python 3.8.
python38-setuptools-wheel	setuptools as a wheel for Python 3.8.
python38-wheel-wheel	wheel package as a wheel for Python 3.8.
qt5-srpm-macros	RPM macros for building Qt 5 packages.
rest	System package rest.
rootfiles	Basic user-level system files.
rpm	RPM Package Manager command-line tools.
rpm-build-libs	Libraries for building RPM packages.
rpm-libs	Libraries for the RPM Package Manager.
rust-srpm-macros	RPM macros for building Rust packages.
sed	Stream editor for filtering and transforming text.
setup	Base system setup files and directories.
shadow-utils	User and group account management utilities.
shared-mime-info	System package shared mime info.
sscg	Simple Signed Certificate Generator.
systemd	System and service manager.
systemd-libs	Systemd shared libraries.
systemd-pam	PAM module for systemd login integration.
tar	Archiving utility for TAR files.
tpm2-tss	TPM 2.0 Software Stack library.
ttmkfdir	Tool to generate fonts.dir for TrueType fonts.
tzdata	Time zone data.
tzdata-java	Time zone data for Java runtime.
unixodbc-devel	Development headers for unixODBC (ODBC driver manager).
usermode	System package usermode.
util-linux	Collection of essential Linux system utilities.
vim-filesystem	Filesystem layout for vim.
vim-minimal	System package vim minimal.
virt-what	Detect if running inside a virtual machine.
which	Locate a command in PATH.
xkeyboard-config	Configuration files for xkeyboard.
xorg-x11-font-utils	Utilities for handling legacy X.Org fonts.
xorg-x11-fonts-type1	Legacy Type1 fonts for X.Org.
xz-libs	Shared libraries for xz.

Cloudera Data Engineering upgrade scenarios

When deciding which upgrade path to take, consider the eligibility of the Cloudera Data Engineering version, the Cloudera Data Engineering services, as well as the recommendations if your Cloudera Data Engineering services include Apache Airflow workloads.

**Note:**

- The Cloudera Data Engineering UI triggers a backup before every upgrade.
- Consider the downtime caused by the upgrade. When the system takes a backup, the Cloudera Data Engineering service enters maintenance mode, pauses jobs and may kill running jobs during the upgrade.
- On AWS, an in-place upgrade from Cloudera Data Engineering version 1.20.3 or higher may result in a downtime of up to 3 hours, as the upgrade includes multiple Kubernetes version updates.

Eligibility for in-place upgrade

1. Check if your Cloudera Data Engineering version is eligible for the in-place upgrade. For more information, see [Cloudera Data Engineering upgrade version compatibility](#).
2. Check if your Cloudera Data Engineering services are eligible for in-place upgrade. For more information, see [Handling ineligible upgrades for Cloudera Data Engineering](#).

For Eligible for in-place upgrade

If both your Cloudera Data Engineering version and services are eligible, take the [in-place upgrade](#), but if you have Apache Airflow workloads in your Cloudera Data Engineering services, before the upgrade, consider the aspects listed in [Upgrade considerations if Cloudera Data Engineering services include Apache Airflow workloads](#).

**Note:** Pre-upgrade Airflow tasks:

- Remove any Airflow libraries or operators, or cancel their activation before upgrading.
- Set the catchup option to False (catchup=False) for all Directed Acyclic Graphs (DAGs). This setting prevents DAGs from rerunning the history if recovery is needed.
- Back up any variables and connections, as they are automatically restored only if the in-place upgrade is successful. If the upgrade fails and backup and restore is performed, variables and connections are not supported.

If the in-place upgrade fails, follow [Handling upgrade failures for Cloudera Data Engineering](#).

For In-eligible for in-place upgrade

If your Cloudera Data Engineering version is not eligible for the in-place upgrade, or if your Cloudera Data Engineering services are in one of the listed in-eligible upgrade cases, take the [manual backup and restore](#).

Upgrading Cloudera Data Lake to version 7.3.1

In Cloudera Data Engineering, Cloudera Data Lake version 7.3.1 only supports Spark 3.5, or higher versions. If you plan to upgrade to Cloudera Data Lake version 7.3.1, first, make sure to migrate your Spark workloads to Spark 3.5. For more information, see [Upgrading to Cloudera Data Lake 7.3.1 with Cloudera Data Engineering](#).

Upgrade considerations if Cloudera Data Engineering services include Apache Airflow workloads


If your Cloudera Data Engineering version and services are eligible for the in-place upgrade, consider these Apache Airflow-related instructions before performing the in-place upgrade.

- **Airflow jobs included in the Cloudera Data Engineering service**
 - Set the catchup option of every Airflow job to false before starting the in-place upgrade. If you do not set the catchup option to false, and if the in-place upgrade fails, you might want to do a manual Cloudera Data Engineering service recovery through the backup. After the backup is restored, any Directed Acyclic Graph (DAG) whose catchup is not set to false might replay its entire run history from the defined DAG start date, which is an undesired behavior in most cases.
 - Set the catchup option to false for every DAG as described in the [DAG runs](#) official Airflow documentation.
- **Airflow Libraries and Operators included in the Cloudera Data Engineering service**

If you use Airflow Libraries and Operators, before starting the in-place upgrade, delete any Airflow Libraries and Operators in Cloudera Data Engineering.

 - If the Airflow Libraries and Operators are in the first step (Configure Repositories) of the configuration process, cancel them.
 - If the Airflow Libraries and Operators have already been built and are only waiting for activation, finish the activation process and delete them.
- **Airflow Variables and Connections included in the Cloudera Data Engineering service**

If you use Airflow Variables and Connections in the Cloudera Data Engineering service, the default backup taken before the upgrade does not include Airflow Variables and Connections.

 -  **Important:** If your in-place upgrade fails and you perform the backup and restore, your Airflow Variables and Connections are lost. As a safety measure, before the upgrade, Cloudera recommends [using the backup-restore-based upgrade script](#) to take a backup.

Alternatively, you can follow the in-place upgrade workflow, and manually back up and restore your Airflow Variables and Connections if the in-place upgrade fails.
 - If your in-place upgrade is successful, your Variables and Connections are kept.
- **Airflow and Airflow-Python version change after in-place upgrade**

For information on the runtime components versions, see [Compatibility for Cloudera Data Engineering and Runtime components](#).

 - If there are Airflow version or Airflow-Python version changes after the in-place upgrade, there is a higher risk of incompatibility or failure. Cloudera recommends to test this with a more rigorous process. For more information, see [In-place upgrade with Airflow Operators and Libraries](#).
 - If there is no Airflow version and Airflow-Python version change, and you are confident in the compatibility of DAGs and libraries, follow: [Upgrading Airflow if DAGs and packages are compatible with new Airflow version](#).

Upgrading Cloudera Data Engineering

Cloudera Data Engineering supports in-place upgrades on both AWS and Azure. Cloudera Data Engineering provides an in-place upgrade option from the eligible versions to the latest version. If you are on a Cloudera Data Engineering version that is not eligible for the in-place upgrade, the in-place upgrade option is disabled and you can perform a backup and restore. The upgrades can be triggered by an Admin from the Cloudera Data Engineering user interface.

About this task

To check whether your Cloudera Data Engineering version is eligible for the in-place upgrade option, see [Support lifecycle policy](#).

When you upgrade to the latest version in Cloudera Data Engineering, the upgrade process prepares your service for the upgrade with a click of a button. Your information is also backed up in this process. Learn how to complete the upgrade process [here](#).

**Important:**

- Virtual clusters for this service are unavailable during the upgrade process and the upgrade puts your service into maintenance. During this time, scheduled jobs are paused and any active jobs are killed. The upgrade can take some time so be advised and alert your team of the downtime.
- If you use Airflow Libraries and Operators, before starting the in-place upgrade, delete any Airflow Libraries and Operators in Cloudera Data Engineering.
 - If the Airflow Libraries and Operators are in the first step (Configure Repositories) of the configuration process, cancel them.
 - If the Airflow Libraries and Operators have already been built and are only waiting for activation, finish the activation process and delete them.

For more information, see [In-place upgrade with Airflow Operators and Libraries](#) and [Upgrading Airflow if DAGs and packages are compatible with new Airflow version](#).

The following are not included in the backup:

- Python-venv resources
- Airflow custom operators and libraries
- Airflow connections
- Airflow variables
- Logs
 - Job Run logs (Driver, Executor, API)
 - Service logs
 - Virtual Cluster event logs
 - Spark Session logs (Including statement history)
 - Airflow DAG logs
- Job Runs
 - Job Run history
- Virtual Cluster end-points
- Resource: Docker runtimes

Before you begin

Ensure that the catchup option is not enabled for any user's Airflow jobs. Before the backup starts, if the Airflow DAG catchup options are enabled, disable them manually.

Procedure

1. In the Cloudera console, click the Data Engineering tile.
The Cloudera Data Engineering **Home** page displays.
2. In the left navigation menu, click Administration, select a Service, and click Upgrade.
3. Click the Maintenance tab and click Upgrade Service.
4. Click Start Preparation.

Cloudera Data Engineering prepares the service for you. If the service cannot be prepared, click Retry or Cancel Upgrade. Canceling the upgrade returns your service to the original state. Optionally, you can click Download Logs to view the Diagnostics page where you can generate a diagnostic bundle. If the preparation is successful, proceed to the next step. If the preparation and cancellation fails, you must contact Cloudera support.

5. Click Start Backup.

Cloudera Data Engineering automatically creates a backup for you. The backup is used for restoring if an irrecoverable failure happens during the upgrade.



Note: If the backup fails, you have the option to click Retry Upgrade or Contact Support. If the upgrade fails a second time, you have the option to contact support or perform a manual upgrade with jobs backup and restore. If the backup is successful, you can proceed with the upgrade.

6. When the backup is complete, to view the backup that was created, select Backup Library.

Status	Backup Id	Name	Description	Created On ↓	Created By
✓	11		23- upgrade		

Items per page: 10 1 - 1 of 1

7. Click Start Upgrade.

8. After the upgrade is complete, to enable all paused jobs and sessions, click Resume Service.

The service restarts.

Results

After the upgrade is complete and you resumed the service, the service operates on the upgraded version of Cloudera Data Engineering. If you cancel the upgrade, your service resumes the previous state, which enables all paused jobs and resumes all sessions.

Related Information

[Handling upgrade failures for Cloudera Data Engineering](#)

[Cloudera Data Engineering upgrade version compatibility](#)

Upgrading to Cloudera Data Lake 7.3.1 with Cloudera Data Engineering

Cloudera Data Engineering 1.23.1 and higher versions support Cloudera Data Lake 7.3.1.

About this task

Upgrading to Cloudera Data Lake 7.3.1 is only supported for specific Cloudera Data Lake versions. For more information on the supported Cloudera Data Lake versions, see [Upgrading to Cloudera Runtime 7.3.1](#).



Note: In Cloudera Data Engineering, Cloudera Data Lake 7.3.1 only supports Apache Spark version 3.5.

Before you begin

If you plan to upgrade to Cloudera Data Lake 7.3.1 with Cloudera Data Engineering 1.23.1 or higher versions, first, make sure to refactor and migrate all your Spark Jobs to Spark 3.5. For more information, see [Migrating Spark applications](#).

Steps

For In Cloudera Data Engineering 1.23 and lower

1. Upgrade to, or create Cloudera Data Engineering Service of versions 1.22 or 1.23, with Data Lake version 7.2.18.

Data Lake 7.2.18 in Cloudera Data Engineering 1.22 and 1.23 supports multiple Spark versions:

- Spark 2.4.8
- Spark 3.2
- Spark 3.3
- Spark 3.5

For more information, see [Compatibility for Cloudera Data Engineering and Runtime components](#).

2. In the Cloudera Data Engineering Service, perform the following.

- a. Create new Virtual Clusters (VCs) with Spark 3.5.
- b. Refactor your current Spark jobs to Spark 3.5.

For information on the Spark application migration instructions, see [Migrating Spark applications](#).

3. Test the Spark 3.5 Jobs by running them.
4. After ensuring that everything works as expected, delete the old VCs with lower Spark versions and keep the new VCs.
5. Upgrade your Cloudera Data Lake version to 7.3.1 in Cloudera Management Console. For more information, see [Upgrade Data Lake Cloudera Runtime version to 7.3.1. and OS to RHEL 8.10](#)
6. Perform in-place upgrade on your Cloudera Data Engineering Service to 1.23.1 or higher versions.

For In Cloudera Data Engineering 1.23.1 and higher

1. Upgrade to, or create Cloudera Data Engineering Service of versions 1.23.1 or higher, with Data Lake version 7.2.18.

Data Lake 7.2.18 in Cloudera Data Engineering 1.23.1 supports multiple Spark versions:

- Spark 2.4.8
- Spark 3.2
- Spark 3.3
- Spark 3.5

For more information, see [Compatibility for Cloudera Data Engineering and Runtime components](#).

2. In the Cloudera Data Engineering Service, perform the following.

- a. Create new Virtual Clusters (VCs) with Spark 3.5.
- b. Refactor your current Spark jobs to Spark 3.5.

For information on the Spark application migration instructions, see [Migrating Spark applications](#).

3. Test the Spark 3.5 Jobs by running them.
4. After ensuring that everything works as expected, delete the old VCs with lower Spark versions and keep the new VCs.
5. Upgrade your Cloudera Data Lake version to 7.3.1 in Cloudera Management Console. For more information, see [Upgrade Data Lake Cloudera Runtime version to 7.3.1. and OS to RHEL 8.10](#)

In-place upgrade with Airflow Operators and Libraries

This section details Airflow considerations to avoid issues during and after the upgrade. Before performing an in-place upgrade of the Airflow operators and libraries, check whether the environment is compatible. This allows you to determine if the DAGs or the Python package dependencies require changes.

About this task



Important: Using AWS, Cloudera Data Engineering supports in-place upgrades from version 1.20.3 to the latest version, which includes several Kubernetes version bumps. For this reason, you can expect the upgrade to take longer than usual, with an upgrade time of approximately 3 hours.



Important:

- If you use Airflow Variables and Connections in the Cloudera Data Engineering service, the default backup taken before the upgrade does not include Airflow Variables and Connections. If your in-place upgrade is successful, your Variables and Connections are kept.
- If your in-place upgrade fails, your Airflow Variables and Connections are lost if you perform the backup and restore. As a safety measure, before the upgrade, Cloudera recommends [using the backup-restore-based upgrade script](#) to take a backup.

Alternatively, you can follow the in-place upgrade workflow, and manually back up and restore your Airflow Variables and Connections if the in-place upgrade fails.



Note:

- Check the compatibility for all used Airflow operators and libraries. If you do not check the compatibility, the recreation of the Airflow operators and libraries, or the DAG execution can fail.

The following is the recommended procedure. However, if you are confident that the DAGs and the packages in the Airflow Custom Libraries and Operators are compatible with the new Airflow version, instead of following this procedure, you can perform the upgrade as described in [Upgrading Airflow if DAGs and packages are compatible with new Airflow version](#).

- After a Cloudera Data Engineering Upgrade, latest Airflow or Python versions such as Airflow 2.9 or Python 3.11 might cause backward compatibility issues with DAGs. These backward compatibility issues might cause downtime for the affected DAGs until they are fixed. To minimize the impact, the upgrade must be planned. The existing DAGs must be verified if they are compliant with the latest Cloudera Data Engineering version before the upgrade to minimize the number of DAGs that have to be fixed post-upgrade (if any).
- On Microsoft Azure, Cloudera Data Engineering version 1.23.0 does not support in-place upgrade from a version that uses the MySQL Single server to a version with MySQL Flexible server. In such a case, perform the upgrade using the procedure described in [Handling upgrade failures for Cloudera Data Engineering](#).

Before you begin

Ensure that the catchup option is not enabled for any user's Airflow jobs. Before the backup starts, if the Airflow DAG catchup options are enabled, disable them manually.

Verify Airflow operators and libraries and DAGs in the newer Cloudera Data Engineering service:

Procedure

1. Create a new test Cloudera Data Engineering service with the latest Cloudera Data Engineering version.

2. Copy the DAGs from the current Cloudera Data Engineering for local checking:
 - a) Update the DAG code for the latest Python version by using the tools mentioned in the [Tools for porting Python from version 3.8 to version 3.11](#) on page 57 section.
 - b) Update and test the DAG code for the latest Airflow version as described in the [Debugging Airflow DAGs locally](#) on page 59 section.
3. Obtain the requirements.txt for the existing Airflow environment. The requirements.txt is required after the upgrade to restore the environments.

To obtain the Airflow environment requirements.txt for a virtual cluster, perform the following steps.

For In Cloudera Data Engineering 1.21 and higher

- a. Obtain the environment information JSON from the Airflow environment API endpoint.
 1. Copy the JOBS API URL. For more information, check the Determine the virtual cluster endpoint URL step in *Configuring the CLI client*.
 2. Use the JOBS API URL as a base and add /admin/airflow/env.

JOBS API URL example with /admin/airflow/env:

```
https://t5kh5fjp.cde-hfrsp8ww.dex-priv.xcu2-8y8x.dev.cldr.work/dex/  
api/v1/admin/airflow/env
```

Example for the JSON:

```
cat pyenv.json  
{  
  "status": "activated",  
  "packages": [  
    {  
      "type": "python-module",  
      "name": "mypy",  
      "version": "1.2.0",  
      "created": "2024-04-03T16:58:07Z"  
    },  
    {  
      "type": "python-module",  
      "name": "mypy-extensions",  
      "version": "1.0.0",  
      "created": "2024-04-03T16:58:07Z"  
    },  
    {  
      "type": "python-module",  
      "name": "tomli",  
      "version": "2.0.1",  
      "created": "2024-04-03T16:58:07Z"  
    }  
  ],  
  "requirements": "mypy\n"
```

- b. From the returned `pyenv.json`, generate a `requirements.txt` file with the `jq` tool:

```
cat pyenv.json | jq -r '.requirements'> requirements.txt
```

Example for the generated `requirements.txt`:

```
mypy
```

For In Cloudera Data Engineering 1.20

Use the CDE CLI to find and download the `requirements.txt` file which belongs to the Airflow environment in a virtual cluster:

- a. Find the resource name that belongs to the Python environment. Optionally use `jq` to get the name:

```
cde resource list --filter "type[eq]airflow-python-env" --filter "status[eq]active" --show-hidden=true | jq '[0].name'
cde-airflow-pyenv-1712224823
```

- b. Download the `requirements.txt` file:

```
cde resource download --name cde-airflow-pyenv-1712224823 --resource-path requirements.txt
```

4. Verify the `requirements.txt` file against an [Airflow constraints file](#).

- Get the Python version used for Airflow from *Compatibility for Cloudera Data Engineering and Runtime components*.
- Check `requirements.txt` against the Cloudera Data Engineering constraint file by creating a temporary Python virtual environment and use the dry-run feature of `pip`. If `pip` does not support dry-run, perform the check by installing the packages without the `--dry-run` flag:

```
# create a temporary local python virtual environment using the same python version used in Airflow in Cloudera Data Engineering.
python -m venv
# source the environment
source venv/bin/activate
# check if there are any conflicts between the packages in the requirements.txt # file and the packages needed by airflow
# At least pip 22.2 is needed for the dry-run feature.
(venv) python -m pip install -r requirements.txt -c "https://raw.githubusercontent.com/apache/airflow/constraints-2.9.3/constraints-3.11.txt" --dry-run
```

5. Create a compatible Python environment in a test virtual cluster in the test Service.

Python packages might need to be updated to work with the Airflow service in the new Cloudera Data Engineering version.

- Backup and restore your Airflow jobs or create all the necessary Airflow jobs in the test Cloudera Data Engineering Service that you created in step 1 on page 20.
- If the DAG parsing fails, go to the Airflow UI to identify the impacted DAGs.
- Fix the DAGs or update the dependencies in the `requirements.txt` to fix the issues.
- Conduct a test run of these DAGs if they call custom libraries at runtime to ensure that the code defined within the tasks is also valid.
- If the DAG execution fails, go to the Airflow job logs to identify the impacted DAGs.
- Fix the DAGs or update the dependencies in the `requirements.txt` to fix the issues.

Backport all DAG changes to the Cloudera Data Engineering service to be upgraded, so that DAGs are compatible with the previous and the new version of Airflow. This may require a retest on an older Cloudera

Data Engineering service test virtual cluster, if the production environment cannot be updated easily. If changes are needed for the older Cloudera Data Engineering service, test the changes again on the newer Cloudera Data Engineering service. Verify if both services work as expected. To fix the compatibility issues regarding Python version change, see [Python upgrade from version 3.8 to version 3.11](#) on page 57 and to fix issues regarding Airflow, see [Airflow upgrade from version 2.7.3 to version 2.9.3](#) on page 58.

12. Delete the test service.
13. Delete all Airflow operators and libraries in all virtual clusters in the Cloudera Data Engineering service before executing the upgrade.
14. Continue with the upgrade.
15. After a successful upgrade, build and activate the Airflow operators and libraries in the virtual clusters from the updated and fixed requirements.txt.
16. Verify that the Airflow jobs are properly running in the Airflow UI and Cloudera Data Engineering Jobs UI.
17. Remove and retest backward compatible changes made in the DAGs for the previous Cloudera Data Engineering version. Keep the DAGs updated to the latest Airflow version in Cloudera Data Engineering.

Related Information

[Configuring the CLI client](#)

[Compatibility for Cloudera Data Engineering and Runtime components](#)

[Upgrading Airflow if DAGs and packages are compatible with new Airflow version](#)

[Airflow constraints file](#)

Airflow constraints file

Before performing an in-place upgrade of the Airflow operators and libraries, check whether the environment is compatible. Use the Airflow constraints file to verify its content against the requirements.txt file that belongs to the Airflow environment in a virtual cluster.

For more information about checking the Airflow constraints file against the requirements.txt file, see step 4 on page 22 in *In-place upgrade with Airflow Operators and Libraries*.

For Upgrade to 1.24.1

Airflow version 2.10.4 constraints file:

```
apache-airflow==2.10.4
waitress==2.1.1
argparse==1.4.0
Authlib==1.3.2
ConfigUpdater==3.2
Deprecated==1.2.15
Events==0.5
Flask-AppBuilder==4.5.2
Flask-Babel==2.0.0
Flask-Bcrypt==1.0.1
Flask-Caching==2.3.0
Flask-JWT-Extended==4.7.1
Flask-Limiter==3.9.2
Flask-Login==0.6.3
Flask-SQLAlchemy==2.5.1
Flask-Session==0.5.0
Flask-WTF==1.2.2
Flask==2.2.5
GitPython==3.1.43
JayDeBeApi==1.2.3
Jinja2==3.1.4
Js2Py==0.74
```

```
Mako==1.3.8
Markdown==3.7
MarkupSafe==3.0.2
PyAthena==3.11.0
PyGithub==2.5.0
PyHive==0.7.0
PyJWT==2.10.1
PyNaCl==1.5.0
PyYAML==6.0.2
Pygments==2.18.0
SQLAlchemy-JSONField==1.0.2
SQLAlchemy-Utils==0.41.2
SQLAlchemy==1.4.54
SecretStorage==3.3.3
Sphinx==5.3.0
WTForms==3.2.1
Werkzeug==2.2.3
adal==1.2.7
adlfs==2024.7.0
aiobotocore==2.15.2
aiofiles==23.2.1
aiohappyeyeballs==2.4.4
aiohttp==3.10.11
aioitertools==0.12.0
aioresponses==0.7.7
aiosignal==1.3.1
airbyte-api==0.52.1
alabaster==0.7.16
alembic==1.14.0
alibabacloud-adb20211201==2.0.0
alibabacloud-tea==0.4.0
alibabacloud_credentials==0.3.6
alibabacloud_endpoint_util==0.0.3
alibabacloud_gateway_spi==0.0.2
alibabacloud_openapi_util==0.2.2
alibabacloud_tea_openapi==0.3.12
alibabacloud_tea_util==0.3.13
alibabacloud_tea_xml==0.0.2
aliyun-python-sdk-core==2.16.0
aliyun-python-sdk-kms==2.16.5
amqp==5.3.1
analytics-python==1.2.9
annotated-types==0.7.0
ansicolors==1.1.8
anyascii==0.3.2
anyio==4.7.0
apache-airflow-providers-airbyte==4.0.0
apache-airflow-providers-alibaba==2.9.1
apache-airflow-providers-amazon==9.1.0
apache-airflow-providers-apache-beam==5.9.1
apache-airflow-providers-apache-cassandra==3.6.0
apache-airflow-providers-apache-drill==2.8.1
apache-airflow-providers-apache-druid==3.12.1
apache-airflow-providers-apache-flink==1.5.1
apache-airflow-providers-apache-hdfs==4.6.0
apache-airflow-providers-apache-hive==8.2.1
apache-airflow-providers-apache-iceberg==1.1.0
apache-airflow-providers-apache-impala==1.5.2
apache-airflow-providers-apache-kafka==1.6.1
apache-airflow-providers-apache-kylin==3.7.0
apache-airflow-providers-apache-livy==3.9.2
apache-airflow-providers-apache-pig==4.5.0
apache-airflow-providers-apache-pinot==4.5.1
apache-airflow-providers-apache-spark==4.11.3
```



```
apache-airflow-providers-apprise==1.4.1
apache-airflow-providers-arangodb==2.6.0
apache-airflow-providers-asana==2.6.0
apache-airflow-providers-atlassian-jira==2.7.1
apache-airflow-providers-celery==3.8.5
apache-airflow-providers-cloudant==4.0.3
apache-airflow-providers-cncf-kubernetes==10.0.1
apache-airflow-providers-cohere==1.3.0
apache-airflow-providers-common-compat==1.2.2
apache-airflow-providers-common-io==1.4.2
apache-airflow-providers-common-sql==1.20.0
apache-airflow-providers-databricks==6.13.0
apache-airflow-providers-datadog==3.7.1
apache-airflow-providers-dbt-cloud==3.11.2
apache-airflow-providers-dingding==3.6.0
apache-airflow-providers-discord==3.8.0
apache-airflow-providers-docker==3.14.1
apache-airflow-providers-elasticsearch==5.5.3
apache-airflow-providers-exasol==4.6.1
apache-airflow-providers-fab==1.5.1
apache-airflow-providers-facebook==3.6.0
apache-airflow-providers-ftp==3.11.1
apache-airflow-providers-github==2.7.0
apache-airflow-providers-google==11.0.0
apache-airflow-providers-grpc==3.6.0
apache-airflow-providers-hashicorp==3.8.0
apache-airflow-providers-http==4.13.3
apache-airflow-providers-imap==3.7.0
apache-airflow-providers-influxdb==2.7.1
apache-airflow-providers-jdbc==4.5.3
apache-airflow-providers-jenkins==3.7.2
apache-airflow-providers-microsoft-azure==11.1.0
apache-airflow-providers-microsoft-mssql==3.9.2
apache-airflow-providers-microsoft-psrp==2.8.0
apache-airflow-providers-microsoft-winrm==3.6.1
apache-airflow-providers-mongo==4.2.2
apache-airflow-providers-mysql==5.7.4
apache-airflow-providers-neo4j==3.7.0
apache-airflow-providers-odbc==4.8.1
apache-airflow-providers-openai==1.4.0
apache-airflow-providers-openfaas==3.6.0
apache-airflow-providers-openlineage==1.14.0
apache-airflow-providers-opensearch==1.5.0
apache-airflow-providers-opsgenie==5.7.0
apache-airflow-providers-oracle==3.12.1
apache-airflow-providers-pagerduty==3.8.1
apache-airflow-providers-papermill==3.8.2
apache-airflow-providers-pgvector==1.3.0
apache-airflow-providers-pinecone==2.1.1
apache-airflow-providers-postgres==5.14.0
apache-airflow-providers-presto==5.7.0
apache-airflow-providers-qdrant==1.2.0
apache-airflow-providers-redis==3.8.0
apache-airflow-providers-salesforce==5.8.0
apache-airflow-providers-samba==4.8.0
apache-airflow-providers-segment==3.6.0
apache-airflow-providers-sendgrid==3.6.0
apache-airflow-providers-sftp==4.11.1
apache-airflow-providers-singularity==3.6.0
apache-airflow-providers-slack==8.9.2
apache-airflow-providers-smtp==1.8.1
apache-airflow-providers-snowflake==5.8.1
apache-airflow-providers-sqlite==3.9.1
apache-airflow-providers-ssh==3.14.0
```

```
apache-airflow-providers-tableau==4.6.1
apache-airflow-providers-tabular==1.6.1
apache-airflow-providers-telegram==4.6.0
apache-airflow-providers-teradata==2.6.1
apache-airflow-providers-trino==5.9.0
apache-airflow-providers-vertica==3.9.1
apache-airflow-providers-weaviate==2.1.0
apache-airflow-providers-yandex==3.12.0
apache-airflow-providers-ydb==2.0.0
apache-airflow-providers-zendesk==4.8.0
apache-beam==2.59.0
apispec==6.8.0
apprise==1.9.0
argcomplete==3.5.2
asana==3.2.3
asgiref==3.8.1
asn1crypto==1.5.1
astroid==2.15.8
asttokens==3.0.0
asyncssh==2.18.0
atlasclient==1.0.0
atlassian-python-api==3.41.16
attrs==24.2.0
aws-sam-translator==1.94.0
aws-xray-sdk==2.14.0
azure-batch==14.2.0
azure-common==1.1.28
azure-core==1.32.0
azure-cosmos==4.9.0
azure-datalake-store==0.0.53
azure-identity==1.19.0
azure-keyvault-secrets==4.9.0
azure-kusto-data==4.6.1
azure-mgmt-containerinstance==10.1.0
azure-mgmt-containerregistry==10.3.0
azure-mgmt-core==1.5.0
azure-mgmt-cosmosdb==9.7.0
azure-mgmt-datafactory==9.0.0
azure-mgmt-datalake-nspkg==3.0.1
azure-mgmt-datalake-store==0.5.0
azure-mgmt-nspkg==3.0.2
azure-mgmt-resource==23.2.0
azure-mgmt-storage==21.2.1
azure-nspkg==3.0.2
azure-servicebus==7.13.0
azure-storage-blob==12.24.0
azure-storage-file-datalake==12.18.0
azure-storage-file-share==12.20.0
azure-synapse-artifacts==0.19.0
azure-synapse-spark==0.7.0
babel==2.16.0
backoff==2.2.1
backports.tarfile==1.2.0
bcrypt==4.2.1
beautifulsoup4==4.12.3
billiard==4.2.1
bitarray==3.0.0
black==24.10.0
blinker==1.9.0
boto3==1.35.36
botocore==1.35.36
cachelib==0.9.0
cachetools==5.5.0
cassandra-driver==3.29.2
```

```
cattrs==24.1.2
celery==5.4.0
certifi==2024.8.30
cffi==1.17.1
cfgv==3.4.0
cfn-lint==1.22.0
cgrouper==0.2.3
chardet==5.2.0
charset-normalizer==3.4.0
checksumdir==1.2.0
ciso8601==2.3.2
click-didyoumean==0.3.1
click-plugins==1.1.1
click-repl==0.3.0
click==8.1.7
clickclick==20.10.2
cloudant==2.15.0
cloudpickle==2.2.1
cohere==4.57
colorama==0.4.6
colorlog==6.9.0
comm==0.2.2
confluent-kafka==2.6.1
connexion==2.14.2
coverage==7.6.9
crcmod==1.7
cron-descriptor==1.4.5
croniter==5.0.1
cryptography==42.0.8
curlify==2.2.1
databricks-sql-connector==2.9.6
dataclasses-json==0.6.7
datadog==0.50.2
db-dtypes==1.3.1
debugpy==1.8.10
decorator==5.1.1
defusedxml==0.7.1
deltalake==0.22.3
diagrams==0.24.1
dill==0.3.1.1
distlib==0.3.9
distro==1.9.0
dnspython==2.7.0
docker==7.1.0
docopt==0.6.2
docstring_parser==0.16
docutils==0.16
duckdb==1.1.3
elastic-transport==8.15.1
elasticsearch==8.16.0
email_validator==2.2.0
entrypoints==0.4
eralchemy2==1.4.1
et_xmlfile==2.0.0
eventlet==0.38.1
execnet==2.1.1
executing==2.1.0
facebook_business==21.0.5
fastavro==1.9.7
fasteners==0.19
fastjsonschema==2.21.1
filelock==3.16.1
flower==2.0.1
frozenlist==1.5.0
```

```
fsspec==2024.10.0
future==1.0.0
gcloud-aio-auth==5.3.2
gcloud-aio-bigquery==7.1.0
gcloud-aio-storage==9.3.0
gcsfs==2024.10.0
geomet==0.2.1.post1
gevent==24.11.1
gitdb==4.0.11
google-ads==25.1.0
google-analytics-admin==0.23.2
google-api-core==2.24.0
google-api-python-client==2.154.0
google-auth-httpplib2==0.2.0
google-auth-oauthlib==1.2.1
google-auth==2.36.0
google-cloud-aiplatform==1.74.0
google-cloud-appengine-logging==1.5.0
google-cloud-audit-log==0.3.0
google-cloud-automl==2.14.1
google-cloud-batch==0.17.31
google-cloud-bigquery-datatransfer==3.17.1
google-cloud-bigquery==3.20.1
google-cloud-bigtable==2.27.0
google-cloud-build==3.27.1
google-cloud-compute==1.22.0
google-cloud-container==2.54.0
google-cloud-core==2.4.1
google-cloud-datacatalog==3.23.0
google-cloud-dataflow-client==0.8.14
google-cloud-dataform==0.5.13
google-cloud-dataplex==2.4.0
google-cloud-dataproc-metastore==1.16.1
google-cloud-dataproc==5.15.1
google-cloud-dlp==3.25.1
google-cloud-kms==3.1.1
google-cloud-language==2.15.1
google-cloud-logging==3.11.3
google-cloud-memcache==1.10.1
google-cloud-monitoring==2.23.1
google-cloud-orchestration-airflow==1.15.1
google-cloud-os-login==2.15.1
google-cloud-pubsub==2.27.1
google-cloud-redis==2.16.1
google-cloud-resource-manager==1.13.1
google-cloud-run==0.10.13
google-cloud-secret-manager==2.21.1
google-cloud-spanner==3.51.0
google-cloud-speech==2.28.1
google-cloud-storage-transfer==1.13.1
google-cloud-storage==2.19.0
google-cloud-tasks==2.17.1
google-cloud-texttospeech==2.21.1
google-cloud-translate==3.18.0
google-cloud-videointelligence==2.14.1
google-cloud-vision==3.8.1
google-cloud-workflows==1.15.1
google-crc32c==1.6.0
google-re2==1.1.20240702
google-resumable-media==2.7.2
googleapis-common-protos==1.66.0
graphql-core==3.2.5
graphviz==0.20.3
greenlet==3.1.1
```

```
grpc-google-iam-v1==0.13.1
grpc-interceptor==0.15.4
grpcio-gcp==0.2.2
grpcio-health-checking==1.62.3
grpcio-status==1.62.3
grpcio-tools==1.62.3
grpcio==1.68.1
gssapi==1.9.0
gunicorn==23.0.0
h11==0.14.0
h2==4.1.0
hatch==1.13.0
hatchling==1.26.3
hdfs==2.7.3
hmsclient==0.1.1
hpack==4.0.0
httpcore==1.0.7
httplib2==0.22.0
httpx==0.27.0
humanize==4.11.0
hvac==2.3.0
hyperframe==6.0.1
hyperlink==21.0.0
ibm-cloud-sdk-core==3.20.3
ibmcloudant==0.9.1
icdiff==2.0.7
identify==2.6.3
idna==3.10
ijson==3.3.0
imagesize==1.4.1
immutabledict==4.2.1
importlib-metadata==6.11.0
impyla==0.19.0
incremental==24.7.2
inflection==0.5.1
influxdb-client==1.48.0
iniconfig==2.0.0
ipdb==0.13.13
ipykernel==6.29.5
ipython==8.30.0
isodate==0.7.2
itsdangerous==2.2.0
jaraco.classes==3.4.0
jaraco.context==6.0.1
jaraco.functools==4.1.0
jedi==0.19.2
jeepney==0.8.0
jiter==0.8.2
jmespath==0.10.0
joserfc==1.0.1
jpypel==1.5.1
json-merge-patch==0.2
jsdiff==2.2.1
jsonpatch==1.33
jsonpath-ng==1.7.0
jsonpath-python==1.0.6
jsonpickle==3.4.2
jsonpointer==3.0.0
jsonschema-path==0.3.3
jsonschema-specifications==2023.12.1
jsonschema==4.23.0
jupyter_client==8.6.3
jupyter_core==5.7.2
keyring==25.5.0
```

```
kombu==5.4.2
krb5==0.7.0
kubernetes==30.1.0
kubernetes_asyncio==30.1.0
kylinpy==2.8.4
lazy-object-proxy==1.10.0
ldap3==2.9.1
limits==3.14.1
linkify-it-py==2.0.3
lockfile==0.12.2
looker-sdk==24.20.0
lxml==5.3.0
lz4==4.3.3
markdown-it-py==3.0.0
marshmallow-oneofschema==3.1.1
marshmallow-sqlalchemy==0.28.2
marshmallow==3.23.1
matplotlib-inline==0.1.7
mdit-py-plugins==0.4.2
mdurl==0.1.2
mergedeep==1.3.4
methodtools==0.4.7
microsoft-kiota-abstractions==1.3.3
microsoft-kiota-authentication-azure==1.1.0
microsoft-kiota-http==1.3.3
microsoft-kiota-serialization-json==1.0.0
microsoft-kiota-serialization-text==1.0.0
mmh3==5.0.1
mongomock==4.3.0
more-itertools==10.5.0
moto==5.0.11
mpmath==1.3.0
msal-extensions==1.2.0
msal==1.31.1
msgraph-core==1.1.7
msrest==0.7.1
msrestazure==0.6.4.post1
multi_key_dict==2.0.3
multidict==6.1.0
mypy-boto3-appflow==1.35.39
mypy-boto3-rds==1.35.72
mypy-boto3-redshift-data==1.35.51
mypy-boto3-s3==1.35.76
mypy-extensions==1.0.0
mypy==1.9.0
mysql-connector-python==9.1.0
mysqlclient==2.2.6
nbclient==0.10.1
nbformat==5.10.4
neo4j==5.27.0
nest-asyncio==1.6.0
networkx==3.4.2
nh3==0.2.19
nodeenv==1.9.1
numpy==1.26.4
oauthlib==3.2.2
objsize==0.7.0
openai==1.57.2
openapi-schema-validator==0.6.2
openapi-spec-validator==0.7.1
openlineage-integration-common==1.25.0
openlineage-python==1.25.0
openlineage_sql==1.25.0
openpyxl==3.1.5
```

```
opensearch-py==2.8.0
opentelemetry-api==1.27.0
opentelemetry-exporter-otlp-proto-common==1.27.0
opentelemetry-exporter-otlp-proto-grpc==1.27.0
opentelemetry-exporter-otlp-proto-http==1.27.0
opentelemetry-exporter-otlp==1.27.0
opentelemetry-exporter-prometheus==0.48b0
opentelemetry-proto==1.27.0
opentelemetry-sdk==1.27.0
opentelemetry-semantic-conventions==0.48b0
opsgenie-sdk==2.1.5
oracledb==2.5.0
ordered-set==4.1.0
orjson==3.10.12
oss2==2.19.1
packaging==24.2
pandas-gbq==0.24.0
pandas-stubs==2.2.3.241126
pandas==2.1.4
papermill==2.6.0
paramiko==3.5.0
parso==0.8.4
pathable==0.4.3
pathspec==0.12.1
pbr==6.1.0
pdpyras==5.3.0
pendulum==3.0.0
pexpect==4.9.0
pgvector==0.3.6
pinecone-client==5.0.1
pinecone-plugin-inference==1.1.0
pinecone-plugin-interface==0.0.7
pinotdb==5.6.0
pipdeptree==2.24.0
pipx==1.7.1
pkginfo==1.12.0
platformdirs==4.3.6
pluggy==1.5.0
ply==3.11
plyvel==1.5.1
portalocker==2.10.1
pprintpp==0.4.0
pre_commit==4.0.1
presto-python-client==0.8.4
prison==0.2.1
prometheus_client==0.21.1
prompt_toolkit==3.0.48
propcache==0.2.1
proto-plus==1.25.0
protobuf==4.25.5
psutil==6.1.0
psycpg2-binary==2.9.10
ptyprocess==0.7.0
pure-sasl==0.6.2
pure_eval==0.2.3
py-partiql-parser==0.5.5
py4j==0.10.9.7
pyOpenSSL==24.3.0
pyarrow-hotfix==0.6
pyarrow==16.1.0
pyasn1==0.6.1
pyasn1_modules==0.4.0
pycountry==24.6.1
pycparser==2.22
```

```
pycryptodome==3.21.0
pydantic==2.10.3
pydantic_core==2.27.1
pydata-google-auth==1.9.0
pydot==1.4.2
pydruidd==0.6.9
pyenchanted==3.2.2
pyexasol==0.27.0
pygraphviz==1.14
pyiceberg==0.8.1
pyjssparser==2.7.1
pykerberos==1.2.4
pymongo==4.10.1
pymssql==2.3.2
pyodbc==5.2.0
pyparsing==3.2.0
pypsrp==0.8.1
pyspark==3.5.3
pyspnego==0.11.2
pytest-asyncio==0.24.0
pytest-cov==6.0.0
pytest-custom-exit-code==0.3.0
pytest-icdiff==0.9
pytest-instafail==0.5.0
pytest-mock==3.14.0
pytest-rerunfailures==15.0
pytest-timeouts==1.2.1
pytest-xdist==3.6.1
pytest==8.3.4
python-arango==8.1.4
python-daemon==3.1.2
python-dateutil==2.9.0.post0
python-dotenv==1.0.1
python-http-client==3.3.7
python-jenkins==1.8.2
python-ldap==3.4.4
python-nvd3==0.16.0
python-slugify==8.0.4
python-telegram-bot==21.9
python3-saml==1.16.0
pytz==2024.2
pywinrm==0.5.0
pyzmq==26.2.0
qdrant-client==1.12.1
reactivex==4.0.4
readme-renderer==43.0
redis==5.2.1
redshift-connector==2.1.4
referencing==0.35.1
regex==2024.11.6
requests-file==2.1.0
requests-kerberos==0.15.0
requests-mock==1.12.1
requests-oauthlib==1.3.1
requests-toolbelt==1.0.0
requests==2.32.3
requests_ntlm==1.3.0
responses==0.25.3
restructuredtext_lint==1.4.0
rfc3339-validator==0.1.4
rfc3986==2.0.0
rich-argparse==1.6.0
rich-click==1.8.5
rich==13.9.4
```



```
rpds-py==0.22.3
rsa==4.9
ruff==0.5.5
s3fs==2024.10.0
s3transfer==0.10.4
scramp==1.4.5
scrapbook==0.5.0
semver==3.0.2
sendgrid==6.11.0
sentinels==1.0.0
sentry-sdk==2.19.2
setproctitle==1.3.4
shapely==2.0.6
shellingham==1.5.4
simple-salesforce==1.12.6
six==1.17.0
slack_sdk==3.33.5
smbprotocol==1.15.0
smap==5.0.1
sniffio==1.3.1
snowballstemmer==2.2.0
snowflake-connector-python==3.12.4
snowflake-snowpark-python==1.26.0
snowflake-sqlalchemy==1.7.1
sortedcontainers==2.4.0
soupsieve==2.6
sphinx-airflow-theme==0.0.12
sphinx-argparse==0.5.0
sphinx-autoapi==2.1.1
sphinx-copybutton==0.5.2
sphinx-jinja==2.0.2
sphinx-rtd-theme==2.0.0
sphinx_design==0.6.0
sphinxcontrib-applehelp==2.0.0
sphinxcontrib-devhelp==2.0.0
sphinxcontrib-htmlhelp==2.1.0
sphinxcontrib-httpdomain==1.8.1
sphinxcontrib-jquery==4.1
sphinxcontrib-jsmath==1.0.1
sphinxcontrib-qthelp==2.0.0
sphinxcontrib-redoc==1.6.0
sphinxcontrib-serializinghtml==1.1.5
sphinxcontrib-spelling==8.0.0
spython==0.3.14
sqlalchemy-bigquery==1.12.0
sqlalchemy-redshift==0.8.14
sqlalchemy-spanner==1.8.0
sqlalchemy_drill==1.1.5
sqlparse==0.5.3
sshtunnel==0.4.0
stack-data==0.6.3
starkbank-ecdsa==2.2.0
statsd==4.0.1
std-uritemplate==2.0.1
strictyaml==1.7.3
sympy==1.13.3
tableauserverclient==0.34
tabulate==0.9.0
tenacity==9.0.0
teradatasql==20.0.0.20
teradatasqlalchemy==20.0.0.3
termcolor==2.5.0
text-unidecode==1.3
thrift-sasl==0.4.3
```

```
thrift==0.16.0
time-machine==2.16.0
tomli_w==1.1.0
tomlkit==0.13.2
tornado==6.4.2
towncrier==24.8.0
tqdm==4.67.1
traitlets==5.14.3
trino==0.331.0
trove-classifiers==2024.10.21.16
twine==6.0.1
types-Deprecated==1.2.15.20241117
types-Markdown==3.7.0.20241204
types-PyMySQL==1.1.0.20241103
types-PyYAML==6.0.12.20240917
types-aiofiles==24.1.0.20240626
types-certifi==2021.10.8.3
types-cffi==1.16.0.20240331
types-croniter==5.0.1.20241205
types-docutils==0.21.0.20241128
types-paramiko==3.5.0.20240928
types-protobuf==5.29.1.20241207
types-pyOpenSSL==24.1.0.20240722
types-python-dateutil==2.9.0.20241206
types-python-slugify==8.0.2.20240310
types-pytz==2024.2.0.20241003
types-redis==4.6.0.20241004
types-requests==2.32.0.20241016
types-setuptools==75.6.0.20241126
types-tabulate==0.9.0.20241207
types-termcolor==1.1.6.2
types-toml==0.10.8.20240310
typing-inspect==0.9.0
typing_extensions==4.12.2
tzdata==2024.2
tzlocal==5.2
uc-micro-py==1.0.3
universal_pathlib==0.2.5
uritemplate==4.1.1
urllib3==2.2.3
userpath==1.9.2
uv==0.4.29
validators==0.33.0
vertica-python==1.4.0
vine==5.1.0
virtualenv==20.28.0
watchtower==3.3.1
wcwidth==0.2.13
weaviate-client==4.8.0
websocket-client==1.8.0
wirerope==0.4.8
wrapt==1.17.0
xmlsec==1.3.14
xmldict==0.14.2
yamllint==1.35.1
yandex-query-client==0.1.4
yandexcloud==0.328.0
yarl==1.18.3
ydb-dbapi==0.1.5
ydb==3.18.10
zeep==4.3.1
zenpy==2.0.56
zipp==3.21.0
zope.event==5.0
```

```
zope.interface==7.2
zstandard==0.23.0
```

For Upgrade to 1.23

Airflow version 2.9.3 constraints file:

```
apache-airflow==2.9.3
waitress==2.1.1
argparse==1.4.0
Authlib==1.3.1
Babel==2.15.0
ConfigUpdater==3.2
Deprecated==1.2.14
Events==0.5
Flask-AppBuilder==4.5.0
Flask-Babel==2.0.0
Flask-Bcrypt==1.0.1
Flask-Caching==2.3.0
Flask-JWT-Extended==4.6.0
Flask-Limiter==3.7.0
Flask-Login==0.6.3
Flask-SQLAlchemy==2.5.1
Flask-Session==0.5.0
Flask-WTF==1.2.1
Flask==2.2.5
GitPython==3.1.43
JPype1==1.5.0
JayDeBeApi==1.2.3
Jinja2==3.1.4
Js2Py==0.74
Mako==1.3.5
Markdown==3.6
MarkupSafe==2.1.5
PyAthena==3.8.3
PyGithub==2.3.0
PyHive==0.7.0
PyJWT==2.8.0
PyNaCl==1.5.0
PyYAML==6.0.1
Pygments==2.18.0
SQLAlchemy-JSONField==1.0.2
SQLAlchemy-Utils==0.41.2
SQLAlchemy==1.4.52
SecretStorage==3.3.3
Sphinx==5.3.0
WTForms==3.1.2
Werkzeug==2.2.3
adal==1.2.7
adlfs==2024.4.1
aiobotocore==2.13.1
aiofiles==23.2.1
aiohttp==3.9.5
aioitertools==0.11.0
aioresponses==0.7.6
aiosignal==1.3.1
alabaster==0.7.16
alembic==1.13.2
alibabacloud-adb20211201==2.0.0
alibabacloud-tea==0.3.9
alibabacloud_credentials==0.3.4
alibabacloud_endpoint_util==0.0.3
alibabacloud_gateway_spi==0.0.1
```

```
alibabacloud_openapi_util==0.2.2
alibabacloud_tea_openapi==0.3.10
alibabacloud_tea_util==0.3.12
alibabacloud_tea_xml==0.0.2
aliyun-python-sdk-core==2.15.1
aliyun-python-sdk-kms==2.16.3
amqp==5.2.0
analytics-python==1.2.9
annotated-types==0.7.0
ansicolors==1.1.8
anyascii==0.3.2
anyio==4.4.0
apache-airflow-providers-airbyte==3.8.1
apache-airflow-providers-alibaba==2.8.1
apache-airflow-providers-amazon==8.25.0
apache-airflow-providers-apache-beam==5.7.1
apache-airflow-providers-apache-cassandra==3.5.1
apache-airflow-providers-apache-drill==2.7.2
apache-airflow-providers-apache-druid==3.10.1
apache-airflow-providers-apache-flink==1.4.2
apache-airflow-providers-apache-hdfs==4.4.2
apache-airflow-providers-apache-hive==8.1.2
apache-airflow-providers-apache-iceberg==1.0.0
apache-airflow-providers-apache-impala==1.4.1
apache-airflow-providers-apache-kafka==1.5.0
apache-airflow-providers-apache-kylin==3.6.2
apache-airflow-providers-apache-livy==3.8.1
apache-airflow-providers-apache-pig==4.4.1
apache-airflow-providers-apache-pinot==4.4.1
apache-airflow-providers-apache-spark==4.8.2
apache-airflow-providers-apprise==1.3.1
apache-airflow-providers-arangodb==2.5.1
apache-airflow-providers-asana==2.5.1
apache-airflow-providers-atlassian-jira==2.6.1
apache-airflow-providers-celery==3.7.2
apache-airflow-providers-cloudant==3.5.2
apache-airflow-providers-cncf-kubernetes==8.3.3
apache-airflow-providers-cohere==1.2.1
apache-airflow-providers-common-io==1.3.2
apache-airflow-providers-common-sql==1.14.2
apache-airflow-providers-databricks==6.7.0
apache-airflow-providers-datadog==3.6.1
apache-airflow-providers-dbt-cloud==3.9.0
apache-airflow-providers-dingding==3.5.1
apache-airflow-providers-discord==3.7.1
apache-airflow-providers-docker==3.12.2
apache-airflow-providers-elasticsearch==5.4.1
apache-airflow-providers-exasol==4.5.2
apache-airflow-providers-fab==1.2.1
apache-airflow-providers-facebook==3.5.2
apache-airflow-providers-ftp==3.10.0
apache-airflow-providers-github==2.6.2
apache-airflow-providers-google==10.21.0
apache-airflow-providers-grpc==3.5.2
apache-airflow-providers-hashicorp==3.7.1
apache-airflow-providers-http==4.12.0
apache-airflow-providers-imap==3.6.1
apache-airflow-providers-influxdb==2.6.0
apache-airflow-providers-jdbc==4.3.1
apache-airflow-providers-jenkins==3.6.1
apache-airflow-providers-microsoft-azure==10.3.0
apache-airflow-providers-microsoft-mssql==3.7.2
apache-airflow-providers-microsoft-psrp==2.7.1
apache-airflow-providers-microsoft-winrm==3.5.1
```

```
apache-airflow-providers-mongo==4.1.2
apache-airflow-providers-mysql==5.6.2
apache-airflow-providers-neo4j==3.6.1
apache-airflow-providers-odbc==4.6.2
apache-airflow-providers-openai==1.2.2
apache-airflow-providers-openfaas==3.5.1
apache-airflow-providers-openlineage==1.9.1
apache-airflow-providers-opensearch==1.3.0
apache-airflow-providers-opsgenie==5.6.1
apache-airflow-providers-oracle==3.10.3
apache-airflow-providers-pagerduty==3.7.2
apache-airflow-providers-papermill==3.7.2
apache-airflow-providers-pgvector==1.2.1
apache-airflow-providers-pinecone==2.0.1
apache-airflow-providers-postgres==5.11.2
apache-airflow-providers-presto==5.5.2
apache-airflow-providers-qdrant==1.1.1
apache-airflow-providers-redis==3.7.1
apache-airflow-providers-salesforce==5.7.2
apache-airflow-providers-samba==4.7.1
apache-airflow-providers-segment==3.5.1
apache-airflow-providers-sendgrid==3.5.1
apache-airflow-providers-sftp==4.10.2
apache-airflow-providers-singularity==3.5.1
apache-airflow-providers-slack==8.7.1
apache-airflow-providers-smtp==1.7.1
apache-airflow-providers-snowflake==5.6.0
apache-airflow-providers-sqlite==3.8.1
apache-airflow-providers-ssh==3.11.2
apache-airflow-providers-tableau==4.5.2
apache-airflow-providers-tabular==1.5.1
apache-airflow-providers-telegram==4.5.2
apache-airflow-providers-teradata==2.4.0
apache-airflow-providers-trino==5.7.2
apache-airflow-providers-vertica==3.8.2
apache-airflow-providers-weaviate==1.4.2
apache-airflow-providers-yandex==3.11.2
apache-airflow-providers-zendesk==4.7.1
apache-beam==2.57.0
apispec==6.6.1
apprise==1.8.0
argcomplete==3.4.0
asana==3.2.3
asgiref==3.8.1
asn1crypto==1.5.1
astroid==2.15.8
asttokens==2.4.1
asyncssh==2.15.0
atlasclient==1.0.0
atlassian-python-api==3.41.14
attrs==23.2.0
aws-sam-translator==1.89.0
aws-xray-sdk==2.14.0
azure-batch==14.2.0
azure-common==1.1.28
azure-core==1.30.2
azure-cosmos==4.7.0
azure-datalake-store==0.0.53
azure-identity==1.17.1
azure-keyvault-secrets==4.8.0
azure-kusto-data==4.5.1
azure-mgmt-containerinstance==10.1.0
azure-mgmt-containerregistry==10.3.0
azure-mgmt-core==1.4.0
```

```
azure-mgmt-cosmosdb==9.5.1
azure-mgmt-datafactory==8.0.0
azure-mgmt-datalake-nspkg==3.0.1
azure-mgmt-datalake-store==0.5.0
azure-mgmt-nspkg==3.0.2
azure-mgmt-resource==23.1.1
azure-mgmt-storage==21.2.1
azure-nspkg==3.0.2
azure-servicebus==7.12.2
azure-storage-blob==12.20.0
azure-storage-file-datalake==12.15.0
azure-storage-file-share==12.16.0
azure-synapse-artifacts==0.19.0
azure-synapse-spark==0.7.0
backoff==2.2.1
backports.tarfile==1.2.0
bcrypt==4.1.3
beautifulsoup4==4.12.3
billiard==4.2.0
bitarray==2.9.2
black==24.4.2
blinker==1.8.2
boto3==1.34.131
botocore==1.34.131
cachelib==0.9.0
cachetools==5.3.3
cassandra-driver==3.29.1
cattr==23.2.3
celery==5.4.0
certifi==2024.7.4
cffi==1.16.0
cfgv==3.4.0
cfn-lint==1.6.0
cgroupspy==0.2.3
charset==5.2.0
charset-normalizer==3.3.2
checksumdir==1.2.0
ciso8601==2.3.1
click-didyoumean==0.3.1
click-plugins==1.1.1
click-repl==0.3.0
click==8.1.7
clickclick==20.10.2
cloudant==2.15.0
cloudpickle==2.2.1
cohere==4.57
colorama==0.4.6
colorlog==4.8.0
comm==0.2.2
confluent-kafka==2.5.0
connexion==2.14.2
coverage==7.6.0
crcmod==1.7
cron-descriptor==1.4.3
croniter==2.0.5
cryptography==41.0.7
curlify==2.2.1
databricks-sql-connector==2.9.6
datadog==0.49.1
db-dtypes==1.2.0
debugpy==1.8.2
decorator==5.1.1
defusedxml==0.7.1
deltalake==0.18.2
```

```
diagrams==0.23.4
dill==0.3.1.1
distlib==0.3.8
distro==1.9.0
dnspython==2.6.1
docker==7.1.0
docopt==0.6.2
docstring_parser==0.16
docutils==0.16
duckdb==1.0.0
elastic-transport==8.13.1
elasticsearch==8.14.0
email_validator==2.2.0
entrypoints==0.4
eralchemy2==1.4.1
et-xmlfile==1.1.0
eventlet==0.36.1
execnet==2.1.1
executing==2.0.1
facebook_business==20.0.0
fastavro==1.9.5
fasteners==0.19
fastjsonschema==2.20.0
filelock==3.15.4
flower==2.0.1
frozenlist==1.4.1
fsspec==2023.12.2
future==1.0.0
gcloud-aio-auth==4.2.3
gcloud-aio-bigquery==7.1.0
gcloud-aio-storage==9.2.0
gcsfs==2023.12.2.post1
geomet==0.2.1.post1
gevent==24.2.1
gitdb==4.0.11
google-ads==24.1.0
google-analytics-admin==0.22.8
google-api-core==2.19.1
google-api-python-client==2.137.0
google-auth-httpplib2==0.2.0
google-auth-oauthlib==1.2.1
google-auth==2.32.0
google-cloud-aiplatform==1.59.0
google-cloud-appengine-logging==1.4.4
google-cloud-audit-log==0.2.5
google-cloud-automl==2.13.4
google-cloud-batch==0.17.22
google-cloud-bigquery-datatransfer==3.15.4
google-cloud-bigquery==3.20.1
google-cloud-bigtable==2.24.0
google-cloud-build==3.24.1
google-cloud-compute==1.19.1
google-cloud-container==2.49.0
google-cloud-core==2.4.1
google-cloud-datacatalog==3.19.1
google-cloud-dataflow-client==0.8.11
google-cloud-dataform==0.5.10
google-cloud-dataplex==2.2.1
google-cloud-dataproc-metastore==1.15.4
google-cloud-dataproc==5.10.1
google-cloud-dlp==3.18.1
google-cloud-kms==2.24.1
google-cloud-language==2.13.4
google-cloud-logging==3.10.0
```

```
google-cloud-memcache==1.9.4
google-cloud-monitoring==2.22.1
google-cloud-orchestration-airflow==1.13.0
google-cloud-os-login==2.14.5
google-cloud-pubsub==2.22.0
google-cloud-redis==2.15.4
google-cloud-resource-manager==1.12.4
google-cloud-run==0.10.7
google-cloud-secret-manager==2.20.1
google-cloud-spanner==3.47.0
google-cloud-speech==2.26.1
google-cloud-storage-transfer==1.11.4
google-cloud-storage==2.17.0
google-cloud-tasks==2.16.4
google-cloud-texttospeech==2.16.4
google-cloud-translate==3.15.4
google-cloud-videointelligence==2.13.4
google-cloud-vision==3.7.3
google-cloud-workflows==1.14.4
google-crc32c==1.5.0
google-re2==1.1.20240702
google-resumable-media==2.7.1
googleapis-common-protos==1.63.2
graphql-core==3.2.3
graphviz==0.20.3
greenlet==3.0.3
grpc-google-iam-v1==0.13.1
grpc-interceptor==0.15.4
grpcio-gcp==0.2.2
grpcio-health-checking==1.62.2
grpcio-status==1.62.2
grpcio-tools==1.62.2
grpcio==1.64.1
gssapi==1.8.3
gunicorn==22.0.0
h11==0.14.0
h2==4.1.0
hatch==1.12.0
hatchling==1.25.0
hdfs==2.7.3
hmsclient==0.1.1
hpack==4.0.0
httpcore==1.0.5
httplib2==0.22.0
httpx==0.27.0
humanize==4.10.0
hvac==2.3.0
hyperframe==6.0.1
hyperlink==21.0.0
icdiff==2.0.7
identify==2.6.0
idna==3.7
ijson==3.3.0
imagesize==1.4.1
importlib-metadata==6.11.0
importlib_resources==6.4.0
impyla==0.19.0
incremental==22.10.0
inflection==0.5.1
influxdb-client==1.44.0
iniconfig==2.0.0
ipdb==0.13.13
ipykernel==6.29.5
ipython==8.26.0
```



```
isodate==0.6.1
itsdangerous==2.2.0
jaraco.classes==3.4.0
jaraco.context==5.3.0
jaraco.functools==4.0.1
jedi==0.19.1
jeepney==0.8.0
jmespath==0.10.0
joserfc==0.12.0
json-merge-patch==0.2
jsdiff==2.1.1
jsonpatch==1.33
jsonpath-ng==1.6.1
jsonpickle==3.2.2
jsonpointer==3.0.0
jsonschema-path==0.3.3
jsonschema-specifications==2023.12.1
jsonschema==4.23.0
jupyter_client==8.6.2
jupyter_core==5.7.2
keyring==25.2.1
kombu==5.3.7
krb5==0.5.1
kubernetes==29.0.0
kubernetes_asyncio==29.0.0
kylinpy==2.8.4
lazy-object-proxy==1.10.0
ldap3==2.9.1
limits==3.13.0
linkify-it-py==2.0.3
lockfile==0.12.2
loguru==0.7.2
looker-sdk==24.10.0
lxml==5.2.2
lz4==4.3.3
markdown-it-py==3.0.0
marshmallow-oneofschema==3.1.1
marshmallow-sqlalchemy==0.28.2
marshmallow==3.21.3
matplotlib-inline==0.1.7
mdit-py-plugins==0.4.1
mdurl==0.1.2
mergedeep==1.3.4
methodtools==0.4.7
microsoft-kiota-abstractions==1.3.3
microsoft-kiota-authentication-azure==1.0.0
microsoft-kiota-http==1.3.2
mmhash3==3.0.1
mongomock==4.1.2
more-itertools==10.3.0
moto==5.0.11
mpmath==1.3.0
msal-extensions==1.2.0
msal==1.29.0
msgraph-core==1.1.1
msrest==0.7.1
msrestazure==0.6.4.post1
multi_key_dict==2.0.3
multidict==6.0.5
mypy-boto3-appflow==1.34.0
mypy-boto3-rds==1.34.135
mypy-boto3-redshift-data==1.34.0
mypy-boto3-s3==1.34.138
mypy-extensions==1.0.0
```

```
mysql==1.9.0
mysql-connector-python==9.0.0
mysqlclient==2.2.4
nbclient==0.10.0
nbformat==5.10.4
neo4j==5.22.0
nest-asyncio==1.6.0
networkx==3.3
nh3==0.2.18
nodeenv==1.9.1
numpy==1.26.4
oauthlib==3.2.2
objsize==0.7.0
openai==1.35.13
openapi-schema-validator==0.6.2
openapi-spec-validator==0.7.1
openlineage-integration-common==1.18.0
openlineage-python==1.18.0
openlineage_sql==1.18.0
openpyxl==3.1.5
opensearch-py==2.6.0
opentelemetry-api==1.25.0
opentelemetry-exporter-otlp-proto-common==1.25.0
opentelemetry-exporter-otlp-proto-grpc==1.25.0
opentelemetry-exporter-otlp-proto-http==1.25.0
opentelemetry-exporter-otlp==1.25.0
opentelemetry-exporter-prometheus==0.46b0
opentelemetry-proto==1.25.0
opentelemetry-sdk==1.25.0
opentelemetry-semantic-conventions==0.46b0
opsgenie-sdk==2.1.5
oracledb==2.2.1
ordered-set==4.1.0
orjson==3.10.6
oss2==2.18.6
packaging==24.1
pandas-gbq==0.23.1
pandas-stubs==2.2.2.240603
pandas==2.1.4
papermill==2.6.0
paramiko==3.4.0
parso==0.8.4
pathable==0.4.3
pathspec==0.12.1
pbr==6.0.0
pdpyras==5.2.0
pendulum==3.0.0
pexpect==4.9.0
pgvector==0.3.1
pinecone-client==4.1.2
pinecone-plugin-interface==0.0.7
pinotdb==5.3.0
pipdeptree==2.23.1
pipx==1.6.0
pkginfo==1.10.0
platformdirs==4.2.2
pluggy==1.5.0
ply==3.11
plyvel==1.5.1
portalocker==2.10.0
pprintpp==0.4.0
pre-commit==3.7.1
presto-python-client==0.8.4
prison==0.2.1
```

```
prometheus_client==0.20.0
prompt_toolkit==3.0.47
proto-plus==1.24.0
protobuf==4.25.3
psutil==6.0.0
psycopg2-binary==2.9.9
ptyprocess==0.7.0
pure-eval==0.2.2
pure-sasl==0.6.2
py-partiql-parser==0.5.5
py4j==0.10.9.7
pyOpenSSL==24.1.0
pyarrow-hotfix==0.6
pyarrow==16.1.0
pyasn1-modules==0.3.0
pyasn1==0.5.1
pycountry==24.6.1
pycparser==2.22
pycryptodome==3.20.0
pydantic==2.8.2
pydantic_core==2.20.1
pydata-google-auth==1.8.2
pydot==1.4.2
pydruid==0.6.9
pyenchanted==3.2.2
pyexasol==0.26.0
pygraphviz==1.13
pyiceberg==0.6.1
pyjparser==2.7.1
pykerberos==1.2.4
pymongo==4.8.0
pymssql==2.3.0
pyodbc==5.1.0
pyparsing==3.1.2
pypsrp==0.8.1
pyspark==3.5.1
pyspnego==0.11.0
pytest-asyncio==0.23.7
pytest-cov==5.0.0
pytest-custom-exit-code==0.3.0
pytest-icdiff==0.9
pytest-instafail==0.5.0
pytest-mock==3.14.0
pytest-rerunfailures==14.0
pytest-timeouts==1.2.1
pytest-xdist==3.6.1
pytest==7.4.4
python-arango==8.0.0
python-daemon==3.0.1
python-dateutil==2.9.0.post0
python-dotenv==1.0.1
python-http-client==3.3.7
python-jenkins==1.8.2
python-ldap==3.4.4
python-nvd3==0.16.0
python-slugify==8.0.4
python-telegram-bot==21.3
python3-saml==1.16.0
pytz==2024.1
pywinrm==0.4.3
pyzmq==26.0.3
qdrant-client==1.10.1
reactivex==4.0.4
readme_renderer==43.0
```

```
redis==5.0.7
redshift-connector==2.1.2
referencing==0.35.1
regex==2024.5.15
requests-file==2.1.0
requests-kerberos==0.15.0
requests-mock==1.12.1
requests-oauthlib==1.3.1
requests-toolbelt==1.0.0
requests==2.31.0
requests_ntlm==1.3.0
responses==0.25.3
restructuredtext_lint==1.4.0
rfc3339-validator==0.1.4
rfc3986==2.0.0
rich-argparse==1.5.2
rich-click==1.8.3
rich==13.7.1
rpds-py==0.19.0
rsa==4.9
ruff==0.3.3
s3fs==2023.12.2
s3transfer==0.10.2
scramp==1.4.5
scrapbook==0.5.0
semver==3.0.2
sendgrid==6.11.0
sentinels==1.0.0
sentry-sdk==2.9.0
setproctitle==1.3.3
shapely==2.0.4
shellingham==1.5.4
simple-salesforce==1.12.6
six==1.16.0
slack_sdk==3.31.0
smbprotocol==1.13.0
smmmap==5.0.1
sniffio==1.3.1
snowballstemmer==2.2.0
snowflake-connector-python==3.11.0
snowflake-sqlalchemy==1.6.1
sortedcontainers==2.4.0
soupsieve==2.5
sphinx-airflow-theme==0.0.12
sphinx-argparse==0.4.0
sphinx-autoapi==2.1.1
sphinx-copybutton==0.5.2
sphinx-jinja==2.0.2
sphinx-rtd-theme==2.0.0
sphinx_design==0.6.0
sphinxcontrib-applehelp==1.0.8
sphinxcontrib-devhelp==1.0.6
sphinxcontrib-htmlhelp==2.0.5
sphinxcontrib-httpdomain==1.8.1
sphinxcontrib-jquery==4.1
sphinxcontrib-jsmath==1.0.1
sphinxcontrib-qthelp==1.0.7
sphinxcontrib-redoc==1.6.0
sphinxcontrib-serializinghtml==1.1.5
sphinxcontrib-spelling==8.0.0
spython==0.3.13
sqlalchemy-bigquery==1.11.0
sqlalchemy-redshift==0.8.14
sqlalchemy-spanner==1.7.0
```

```
sqlalchemy_drill==1.1.5
sqlparse==0.5.0
sshtunnel==0.4.0
stack-data==0.6.3
starkbank-ecdsa==2.2.0
statsd==4.0.1
std-uritemplate==1.0.3
strictyaml==1.7.3
sympy==1.13.0
tableauserverclient==0.31
tabulate==0.9.0
tenacity==8.5.0
teradatasql==20.0.0.13
teradatasqlalchemy==20.0.0.1
termcolor==2.4.0
text-unidecode==1.3
thrift-sasl==0.4.3
thrift==0.16.0
time-machine==2.14.2
tomli_w==1.0.0
tomlkit==0.13.0
tornado==6.4.1
towncrier==23.11.0
tqdm==4.66.4
traitlets==5.14.3
trino==0.329.0
trove-classifiers==2024.7.2
twine==5.1.1
typed-ast==1.5.5
types-Deprecated==1.2.9.20240311
types-Markdown==3.6.0.20240316
types-PyMySQL==1.1.0.20240524
types-PyYAML==6.0.12.20240311
types-aiofiles==24.1.0.20240626
types-certifi==2021.10.8.3
types-cffi==1.16.0.20240331
types-croniter==2.0.0.20240423
types-docutils==0.21.0.20240711
types-paramiko==3.4.0.20240423
types-protobuf==5.27.0.20240626
types-pyOpenSSL==24.1.0.20240425
types-python-dateutil==2.9.0.20240316
types-python-slugify==8.0.2.20240310
types-pytz==2024.1.0.20240417
types-redis==4.6.0.20240425
types-requests==2.32.0.20240712
types-setuptools==70.3.0.20240710
types-tabulate==0.9.0.20240106
types-termcolor==1.1.6.2
types-toml==0.10.8.20240310
typing_extensions==4.12.2
tzdata==2024.1
tzlocal==5.2
uc-micro-py==1.0.3
unicodedsv==0.14.1
universal_pathlib==0.2.2
uritemplate==4.1.1
urllib3==2.0.7
userpath==1.9.2
uv==0.2.22
validators==0.28.3
vertica-python==1.3.8
vine==5.1.0
virtualenv==20.26.3
```

```
watchtower==3.2.0
wcwidth==0.2.13
weaviate-client==4.6.5
websocket-client==1.8.0
wirerope==0.4.7
wrapt==1.16.0
xmlsec==1.3.13
xmlltodict==0.13.0
yamllint==1.35.1
yandex-query-client==0.1.4
yandexcloud==0.291.0
yarl==1.9.4
zeep==4.2.1
zenpy==2.0.49
zipp==3.19.2
zope.event==5.0
zope.interface==6.4.post2
zstandard==0.22.0
```

For Upgrade to 1.22

Airflow version 2.7.3 constraints file:

```
apache-airflow==2.7.3
waitress==2.1.1
argparse==1.4.0
Authlib==1.2.1
Babel==2.13.1
ConfigUpdater==3.1.1
Deprecated==1.2.14
Flask-AppBuilder==4.3.6
Flask-Babel==2.0.0
Flask-Bcrypt==1.0.1
Flask-Caching==2.1.0
Flask-JWT-Extended==4.5.3
Flask-Limiter==3.5.0
Flask-Login==0.6.3
Flask-SQLAlchemy==2.5.1
Flask-Session==0.5.0
Flask-WTF==1.2.1
Flask==2.2.5
GitPython==3.1.40
JPype1==1.4.1
JayDeBeApi==1.2.3
Jinja2==3.1.2
Js2Py==0.74
Mako==1.2.4
Markdown==3.5.1
MarkupSafe==2.1.3
PyGithub==2.1.1
PyHive==0.7.0
PyJWT==2.8.0
PyNaCl==1.5.0
PyYAML==6.0.1
Pygments==2.16.1
SQLAlchemy-JSONField==1.0.1.post0
SQLAlchemy-Utils==0.41.1
SQLAlchemy==1.4.50
SecretStorage==3.3.3
Sphinx==5.3.0
WTForms==3.0.1
Werkzeug==2.2.3
adal==1.2.7
```

```
adlfs==2023.10.0
aiobotocore==2.7.0
aiofiles==23.2.1
aiohttp==3.8.6
aioitertools==0.11.0
aioresponses==0.7.4
aiosignal==1.3.1
alabaster==0.7.13
alembic==1.12.1
alibabacloud-adb20211201==1.0.0
alibabacloud-credentials==0.3.2
alibabacloud-endpoint-util==0.0.3
alibabacloud-gateway-spi==0.0.1
alibabacloud-openapi-util==0.2.2
alibabacloud-tea-openapi==0.3.7
alibabacloud-tea-util==0.3.11
alibabacloud-tea-xml==0.0.2
alibabacloud-tea==0.3.3
aliyun-python-sdk-core==2.14.0
aliyun-python-sdk-kms==2.16.2
amqp==5.1.1
analytics-python==1.4.post1
annotated-types==0.6.0
ansiwrap==0.8.4
anyascii==0.3.2
anyio==4.0.0
apache-airflow-providers-airbyte==3.4.0
apache-airflow-providers-alibaba==2.6.0
apache-airflow-providers-amazon==8.10.0
apache-airflow-providers-apache-beam==5.3.0
apache-airflow-providers-apache-cassandra==3.3.0
apache-airflow-providers-apache-drill==2.5.0
apache-airflow-providers-apache-druid==3.6.0
apache-airflow-providers-apache-flink==1.2.0
apache-airflow-providers-apache-hdfs==4.2.0
apache-airflow-providers-apache-hive==6.2.0
apache-airflow-providers-apache-impala==1.2.0
apache-airflow-providers-apache-kafka==1.2.0
apache-airflow-providers-apache-kylin==3.3.0
apache-airflow-providers-apache-livy==3.6.0
apache-airflow-providers-apache-pig==4.2.0
apache-airflow-providers-apache-pinot==4.2.0
apache-airflow-providers-apache-spark==4.3.0
apache-airflow-providers-apache-sqoop==4.1.0
apache-airflow-providers-apprise==1.1.0
apache-airflow-providers-arangodb==2.3.0
apache-airflow-providers-asana==2.3.0
apache-airflow-providers-atlassian-jira==2.2.0
apache-airflow-providers-celery==3.4.1
apache-airflow-providers-cloudant==3.3.0
apache-airflow-providers-cncf-kubernetes==7.8.0
apache-airflow-providers-common-sql==1.8.0
apache-airflow-providers-daskexecutor==1.1.0
apache-airflow-providers-databricks==4.7.0
apache-airflow-providers-datadog==3.4.0
apache-airflow-providers-dbt-cloud==3.4.0
apache-airflow-providers-dingding==3.3.0
apache-airflow-providers-discord==3.4.0
apache-airflow-providers-docker==3.8.0
apache-airflow-providers-elasticsearch==5.1.0
apache-airflow-providers-exasol==4.3.0
apache-airflow-providers-facebook==3.3.0
apache-airflow-providers-ftp==3.6.0
apache-airflow-providers-github==2.4.0
```

```
apache-airflow-providers-google==10.11.0
apache-airflow-providers-grpc==3.3.0
apache-airflow-providers-hashicorp==3.5.0
apache-airflow-providers-http==4.6.0
apache-airflow-providers-imap==3.4.0
apache-airflow-providers-influxdb==2.3.0
apache-airflow-providers-jdbc==4.1.0
apache-airflow-providers-jenkins==3.4.0
apache-airflow-providers-microsoft-azure==8.1.0
apache-airflow-providers-microsoft-mssql==3.5.0
apache-airflow-providers-microsoft-psrp==2.4.0
apache-airflow-providers-microsoft-winrm==3.3.0
apache-airflow-providers-mongo==3.4.0
apache-airflow-providers-mysql==5.4.0
apache-airflow-providers-neo4j==3.4.0
apache-airflow-providers-odbc==4.1.0
apache-airflow-providers-openfaas==3.3.0
apache-airflow-providers-openlineage==1.2.0
apache-airflow-providers-opensearch==1.0.0
apache-airflow-providers-opsgenie==5.2.0
apache-airflow-providers-oracle==3.8.0
apache-airflow-providers-pagerduty==3.4.0
apache-airflow-providers-papermill==3.4.0
apache-airflow-providers-plexus==3.3.0
apache-airflow-providers-postgres==5.7.1
apache-airflow-providers-presto==5.2.1
apache-airflow-providers-redis==3.4.0
apache-airflow-providers-salesforce==5.5.0
apache-airflow-providers-samba==4.3.0
apache-airflow-providers-segment==3.3.0
apache-airflow-providers-sendgrid==3.3.0
apache-airflow-providers-sftp==4.7.0
apache-airflow-providers-singularity==3.3.0
apache-airflow-providers-slack==8.3.0
apache-airflow-providers-smtp==1.4.1
apache-airflow-providers-snowflake==5.1.0
apache-airflow-providers-sqlite==3.5.0
apache-airflow-providers-ssh==3.8.1
apache-airflow-providers-tableau==4.3.0
apache-airflow-providers-tabular==1.3.0
apache-airflow-providers-telegram==4.2.0
apache-airflow-providers-trino==5.4.0
apache-airflow-providers-vertica==3.6.0
apache-airflow-providers-zendesk==4.4.0
apache-beam==2.51.0
apispec==6.3.0
apprise==1.6.0
argcomplete==3.1.3
arrow==1.3.0
asana==3.2.2
asgiref==3.7.2
asn1crypto==1.5.1
astroid==2.15.8
asttokens==2.4.1
async-timeout==4.0.3
atlasclient==1.0.0
atlassian-python-api==3.41.3
attrs==23.1.0
aws-sam-translator==1.79.0
aws-xray-sdk==2.12.1
azure-batch==14.0.0
azure-common==1.1.28
azure-core==1.29.5
azure-cosmos==4.5.1
```



```
azure-datalake-store==0.0.53
azure-identity==1.15.0
azure-keyvault-secrets==4.7.0
azure-kusto-data==4.2.0
azure-mgmt-containerinstance==10.1.0
azure-mgmt-containerregistry==10.2.0
azure-mgmt-core==1.4.0
azure-mgmt-cosmosdb==9.3.0
azure-mgmt-datafactory==3.1.0
azure-mgmt-datalake-nspkg==3.0.1
azure-mgmt-datalake-store==0.5.0
azure-mgmt-nspkg==3.0.2
azure-mgmt-resource==23.0.1
azure-mgmt-storage==21.1.0
azure-nspkg==3.0.2
azure-servicebus==7.11.3
azure-storage-blob==12.18.3
azure-storage-common==2.1.0
azure-storage-file-datalake==12.13.2
azure-storage-file-share==12.14.2
azure-storage-file==2.1.0
azure-synapse-spark==0.7.0
backcall==0.2.0
backoff==1.10.0
backports.zoneinfo==0.2.1
bcrypt==4.0.1
beautifulsoup4==4.12.2
billiard==4.1.0
bitarray==2.8.2
black==23.10.1
blinker==1.6.3
boto3==1.28.64
botocore==1.31.64
cachelib==0.9.0
cachetools==5.3.2
cassandra-driver==3.28.0
cattr==23.1.2
celery==5.3.4
certifi==2023.7.22
cffi==1.16.0
cfgv==3.4.0
cfn-lint==0.83.1
cgroupspy==0.2.2
chardet==5.2.0
charset-normalizer==3.3.2
checksumdir==1.2.0
ciso8601==2.3.1
click-didyoumean==0.3.0
click-plugins==1.1.1
click-repl==0.3.0
click==8.1.7
clickclick==20.10.2
cloudant==2.15.0
cloudpickle==2.2.1
colorama==0.4.6
colorlog==4.8.0
confluent-kafka==2.3.0
connexion==2.14.2
coverage==7.3.2
crcmod==1.7
cron-descriptor==1.4.0
croniter==2.0.1
cryptography==41.0.5
curlify==2.2.1
```

```
dask==2023.4.1
databricks-sql-connector==2.9.3
datadog==0.47.0
db-dtypes==1.1.1
decorator==5.1.1
defusedxml==0.7.1
deprecation==2.1.0
dill==0.3.1.1
distlib==0.3.7
distributed==2023.4.1
dnspython==2.4.2
docker==6.1.3
docopt==0.6.2
docutils==0.20.1
duckdb==0.9.1
ecdsa==0.18.0
elastic-transport==8.10.0
elasticsearch==8.10.1
email-validator==1.3.1
entrypoints==0.4
eralchemy2==1.3.8
et-xmlfile==1.1.0
eventlet==0.33.3
exceptiongroup==1.1.3
execnet==2.0.2
executing==2.0.1
facebook-business==18.0.3
fastavro==1.9.0
fasteners==0.19
fastjsonschema==2.18.1
filelock==3.13.1
flower==2.0.1
frozenlist==1.4.0
fsspec==2023.10.0
future==0.18.3
gcloud-aio-auth==4.2.3
gcloud-aio-bigquery==7.0.0
gcloud-aio-storage==9.0.0
gcsfs==2023.10.0
geomet==0.2.1.post1
gevent==23.9.1
gitdb==4.0.11
google-ads==22.1.0
google-api-core==2.12.0
google-api-python-client==2.106.0
google-auth-httpplib2==0.1.1
google-auth-oauthlib==1.1.0
google-auth==2.23.4
google-cloud-aiplatform==1.36.0
google-cloud-appengine-logging==1.3.2
google-cloud-audit-log==0.2.5
google-cloud-automl==2.11.3
google-cloud-batch==0.17.2
google-cloud-bigquery-datatransfer==3.12.1
google-cloud-bigquery-storage==2.22.0
google-cloud-bigquery==3.13.0
google-cloud-bigtable==2.21.0
google-cloud-build==3.20.1
google-cloud-compute==1.14.1
google-cloud-container==2.32.0
google-cloud-core==2.3.3
google-cloud-datacatalog==3.16.0
google-cloud-dataflow-client==0.8.5
google-cloud-dataform==0.5.3
```

```
google-cloud-dataplex==1.7.0
google-cloud-dataproc-metastore==1.13.0
google-cloud-dataproc==5.6.0
google-cloud-dlp==3.12.3
google-cloud-kms==2.19.2
google-cloud-language==2.11.1
google-cloud-logging==3.8.0
google-cloud-memcache==1.7.3
google-cloud-monitoring==2.16.0
google-cloud-orchestration-airflow==1.9.2
google-cloud-os-login==2.11.0
google-cloud-pubsub==2.18.4
google-cloud-redis==2.13.2
google-cloud-resource-manager==1.10.4
google-cloud-run==0.10.0
google-cloud-secret-manager==2.16.4
google-cloud-spanner==3.40.1
google-cloud-speech==2.21.1
google-cloud-storage-transfer==1.9.2
google-cloud-storage==2.13.0
google-cloud-tasks==2.14.2
google-cloud-texttospeech==2.14.2
google-cloud-translate==3.12.1
google-cloud-videointelligence==2.11.4
google-cloud-vision==3.4.5
google-cloud-workflows==1.12.1
google-crc32c==1.5.0
google-re2==1.1
google-resumable-media==2.6.0
googleapis-common-protos==1.61.0
graphql-core==3.2.3
graphviz==0.20.1
greenlet==3.0.1
grpc-google-iam-v1==0.12.6
grpcio-gcp==0.2.2
grpcio-status==1.59.2
grpcio==1.59.2
gssapi==1.8.3
unicorn==21.2.0
h11==0.14.0
hdfs==2.7.3
hmsclient==0.1.1
httpcore==0.16.3
httplib2==0.22.0
httpx==0.23.3
humanize==4.8.0
hvac==2.0.0
identify==2.5.31
idna==3.4
ijson==3.2.3
imagesize==1.4.1
importlib-metadata==6.8.0
importlib-resources==6.1.0
impyla==0.18.0
incremental==22.10.0
inflection==0.5.1
influxdb-client==1.38.0
iniconfig==2.0.0
ipdb==0.13.13
ipython==8.12.3
isodate==0.6.1
itsdangerous==2.1.2
jaraco.classes==3.3.0
jedi==0.19.1
```

```
jeepney==0.8.0
jmespath==0.10.0
jschema-to-python==1.2.3
json-merge-patch==0.2
jsongdiff==2.0.0
jsonpatch==1.33
jsonpath-ng==1.6.0
jsonpickle==3.0.2
jsonpointer==2.4
jsonschema-path==0.3.1
jsonschema-specifications==2023.7.1
jsonschema==4.19.2
junit-xml==1.9
jupyter_client==8.5.0
jupyter_core==5.5.0
keyring==24.2.0
kombu==5.3.2
krb5==0.5.1
kubernetes-asyncio==24.2.3
kubernetes==23.6.0
kylinpy==2.8.4
lazy-object-proxy==1.9.0
ldap3==2.9.1
limits==3.6.0
linkify-it-py==2.0.2
locket==1.0.0
lockfile==0.12.2
looker-sdk==23.16.0
lxml==4.9.3
lz4==4.3.2
markdown-it-py==3.0.0
marshmallow-oneofschema==3.0.1
marshmallow-sqlalchemy==0.26.1
marshmallow==3.20.1
matplotlib-inline==0.1.6
mdit-py-plugins==0.4.0
mdurl==0.1.2
mongomock==4.1.2
monotonic==1.6
more-itertools==10.1.0
moto==4.2.7
mpmath==1.3.0
msal-extensions==1.0.0
msal==1.24.1
msgpack==1.0.7
msrest==0.7.1
msrestazure==0.6.4
multi-key-dict==2.0.3
multidict==6.0.4
mypy-boto3-appflow==1.28.42
mypy-boto3-rds==1.28.76
mypy-boto3-redshift-data==1.28.36
mypy-boto3-s3==1.28.55
mypy-extensions==1.0.0
mypy==1.2.0
mysql-connector-python==8.0.29
mysqlclient==2.2.0
nbclient==0.8.0
nbformat==5.9.2
neo4j==5.14.0
networkx==3.1
nh3==0.2.14
nodeenv==1.8.0
numpy==1.24.4
```

```
oauthlib==3.2.2
objsize==0.6.1
openapi-schema-validator==0.6.2
openapi-spec-validator==0.7.1
openlineage-integration-common==1.4.1
openlineage-python==1.4.1
openlineage_sql==1.4.1
openpyxl==3.1.2
opensearch-py==2.3.2
opentelemetry-api==1.20.0
opentelemetry-exporter-otlp-proto-common==1.20.0
opentelemetry-exporter-otlp-proto-grpc==1.20.0
opentelemetry-exporter-otlp-proto-http==1.20.0
opentelemetry-exporter-otlp==1.20.0
opentelemetry-exporter-prometheus==1.12.0rc1
opentelemetry-proto==1.20.0
opentelemetry-sdk==1.20.0
opentelemetry-semantic-conventions==0.41b0
opsgenie-sdk==2.1.5
oracledb==1.4.2
ordered-set==4.1.0
orjson==3.9.10
oscrypto==1.3.0
oss2==2.18.3
packaging==23.2
pandas-gbq==0.19.2
pandas==2.0.3
papermill==2.4.0
paramiko==3.3.1
parso==0.8.3
partd==1.4.1
pathable==0.4.3
pathspect==0.11.2
pbr==5.11.1
pdpyras==5.1.2
pendulum==2.1.2
pexpect==4.8.0
pickleshare==0.7.5
pinotdb==5.1.0
pipdeptree==2.13.0
pipx==1.2.1
pkginfo==1.9.6
pkgutil_resolve_name==1.3.10
platformdirs==3.11.0
pluggy==1.3.0
ply==3.11
plyvel==1.5.0
portalocker==2.8.2
pre-commit==3.5.0
presto-python-client==0.8.4
prison==0.2.1
prometheus-client==0.18.0
prompt-toolkit==3.0.39
proto-plus==1.22.3
protobuf==4.24.4
psutil==5.9.6
psycpg2-binary==2.9.9
ptyprocess==0.7.0
pure-eval==0.2.2
pure-sasl==0.6.2
py-partiql-parser==0.4.1
py4j==0.10.9.7
pyOpenSSL==23.3.0
pyarrow==11.0.0
```

```
pyasn1-modules==0.3.0
pyasn1==0.5.0
pycountry==22.3.5
pycparser==2.21
pycryptodome==3.19.0
pycryptodomex==3.19.0
pydantic==2.4.2
pydantic_core==2.10.1
pydata-google-auth==1.8.2
pydot==1.4.2
pydruidd==0.6.5
pyenchanted==3.2.2
pyexasol==0.25.2
pygraphviz==1.11
pyjssparser==2.7.1
pykerberos==1.2.4
pymongo==4.5.0
pymssql==2.2.10
pyodbc==5.0.1
pyparsing==3.1.1
pypsrp==0.8.1
pyspark==3.5.0
pyspnego==0.10.2
pytest-asyncio==0.21.1
pytest-cov==4.1.0
pytest-httpx==0.21.3
pytest-instafail==0.5.0
pytest-mock==3.12.0
pytest-rerunfailures==12.0
pytest-timeouts==1.2.1
pytest-xdist==3.3.1
pytest==7.4.3
python-arango==7.7.0
python-daemon==3.0.1
python-dateutil==2.8.2
python-dotenv==1.0.0
python-http-client==3.3.7
python-jenkins==1.7.0
python-jose==3.3.0
python-ldap==3.4.3
python-nvd3==0.15.0
python-slugify==8.0.1
python-telegram-bot==20.2
pytz==2023.3.post1
pytzdata==2020.1
pywinrm==0.4.3
pyzmq==25.1.1
reactivex==4.0.4
readme-renderer==42.0
redis==4.6.0
redshift-connector==2.0.915
referencing==0.30.2
regex==2023.10.3
requests-file==1.5.1
requests-kerberos==0.14.0
requests-mock==1.11.0
requests-ntlm==1.2.0
requests-oauthlib==1.3.1
requests-toolbelt==1.0.0
requests==2.31.0
responses==0.23.3
rfc3339-validator==0.1.4
rfc3986==1.5.0
rich-argparse==1.4.0
```

```
rich-click==1.7.1
rich==13.6.0
rpds-py==0.10.6
rsa==4.9
ruff==0.1.3
s3transfer==0.7.0
sarif-om==1.0.4
scramp==1.4.4
scrapbook==0.5.0
semver==3.0.2
sendgrid==6.10.0
sentinels==1.0.0
sentry-sdk==1.33.1
setproctitle==1.3.3
shapely==2.0.2
simple-salesforce==1.12.5
six==1.16.0
slack-sdk==3.23.0
smbprotocol==1.11.0
smmmap==5.0.1
sniffio==1.3.0
snowballstemmer==2.2.0
snowflake-connector-python==3.3.1
snowflake-sqlalchemy==1.5.0
sortedcontainers==2.4.0
soupsieve==2.5
sphinx-airflow-theme==0.0.12
sphinx-argparse==0.4.0
sphinx-autoapi==2.1.1
sphinx-copybutton==0.5.2
sphinx-jinja==2.0.2
sphinx-rtd-theme==1.3.0
sphinxcontrib-applehelp==1.0.4
sphinxcontrib-devhelp==1.0.2
sphinxcontrib-htmlhelp==2.0.1
sphinxcontrib-httpdomain==1.8.1
sphinxcontrib-jquery==4.1
sphinxcontrib-jsmath==1.0.1
sphinxcontrib-qthelp==1.0.3
sphinxcontrib-redoc==1.6.0
sphinxcontrib-serializinghtml==1.1.5
sphinxcontrib-spelling==8.0.0
spython==0.3.1
sqlalchemy-bigquery==1.8.0
sqlalchemy-drill==1.1.4
sqlalchemy-redshift==0.8.14
sqlalchemy-spanner==1.6.2
sqlparse==0.4.4
sshpubkeys==3.3.1
sshtunnel==0.4.0
stack-data==0.6.3
starkbank-ecdsa==2.2.0
statsd==4.0.1
sympy==1.12
tableauserverclient==0.25
tabulate==0.9.0
tblib==3.0.0
tenacity==8.2.3
termcolor==2.3.0
text-unidecode==1.3
textwrap3==0.9.2
thrift-sasl==0.4.3
thrift==0.16.0
time-machine==2.13.0
```

```
tomli==2.0.1
tomlkit==0.12.1
toolz==0.12.0
tornado==6.3.3
towncrier==23.10.0
tqdm==4.66.1
traitlets==5.13.0
trino==0.327.0
twine==4.0.2
types-Deprecated==1.2.9.3
types-Markdown==3.5.0.0
types-PyMySQL==1.1.0.1
types-PyYAML==6.0.12.12
types-certifi==2021.10.8.3
types-croniter==2.0.0.0
types-docutils==0.20.0.3
types-paramiko==3.3.0.0
types-protobuf==4.24.0.4
types-pyOpenSSL==23.3.0.0
types-python-dateutil==2.8.19.14
types-python-slugify==8.0.0.3
types-pytz==2023.3.1.1
types-redis==4.6.0.8
types-requests==2.31.0.6
types-setuptools==68.2.0.0
types-tabulate==0.9.0.3
types-termcolor==1.1.6.2
types-toml==0.10.8.7
types-urllib3==1.26.25.14
typing_extensions==4.8.0
tzdata==2023.3
tzlocal==5.2
uc-micro-py==1.0.2
unicodcsv==0.14.1
uritemplate==4.1.1
urllib3==1.26.18
userpath==1.9.1
vertica-python==1.3.6
vine==5.0.0
virtualenv==20.24.6
watchtower==3.0.1
wcwidth==0.2.9
websocket-client==1.6.4
wrapt==1.15.0
xmldict==0.13.0
yamllint==1.32.0
yarl==1.9.2
zeep==4.2.1
zenpy==2.0.41
zict==3.0.0
zipp==3.17.0
zope.event==5.0
zope.interface==6.1
zstandard==0.22.0
```

Related Information

[In-place Upgrade with Airflow Operators and Libraries](#)

Troubleshooting Python or Airflow compatibility issues

You can learn more about how to identify the compatibility issues related to Python or Airflow versions during Cloudera Data Engineering upgrade and troubleshoot them.

Python upgrade from version 3.8 to version 3.11

- If you are upgrading Python from version 3.8 to version 3.9, for more information, see:
 - [What's New in Python 3.9](#)
 - [Removed in Python 3.9](#)
 - [Porting to Python 3.9](#)
 - [Check for DeprecationWarning](#)
- If you are upgrading Python from version 3.9 to version 3.10, for more information, see:
 - [What's New In Python 3.10](#)
 - [Deprecated in Python 3.10](#)
 - [Removed in Python 3.10](#)
 - [Porting to Python 3.10](#)
 - Some of the important deprecations, removals, and restrictions in Python 3.10:
 - [PEP 644 – Require OpenSSL 1.1.1 or newer](#)
 - [PEP 632 – Deprecate distutils module](#)
 - [PEP 623 – Remove wstr from Unicode](#)
 - [PEP 624 – Remove Py_UNICODE encoder APIs](#)
 - [PEP 597 – Add optional EncodingWarning](#)
- If you are upgrading Python from version 3.10 to version 3.11, for more information, see:
 - [What's New In Python 3.11](#)
 - [Deprecated in Python 3.11](#)
 - [Porting to Python 3.11](#)
 - Some of the important deprecations, removals, and restrictions in Python 3.11:
 - [PEP 594: Many legacy standard library modules have been deprecated](#). These will be removed in Python 3.13.
 - [PEP 624: Py_UNICODE encoder APIs are removed](#)

Tools for porting Python from version 3.8 to version 3.11

This list contains automated tools that allow you to check the Python source code and modernize it to the latest version if necessary:

- Python development mode: It can be enabled using the `-X dev` command line option or by setting the `PYTHONDEVMODE` environment variable to 1.
- DeprecationWarnings: To check for deprecation warnings in your code, see [DeprecationWarnings](#).
- refurb: A tool for modernizing Python codebases. For more information, see [refurb](#). For example,

```
refurb --enable-all --python-version 3.8 dag_file.py
```

- pyupgrade: A tool to automatically upgrade syntax for newer versions of Python. For more information, see [pyupgrade](#). For example,


```
# create a temporary git repository for the DAG files and run pyupgrade on them
# NOTE: pyupgrade will rewrite the files automatically
pyupgrade --py311-plus *.py
```

```
# check the modifications done by pyupgrade before updating the DAG files
in CDE
```

Airflow upgrade from version 2.7.3 to version 2.9.3

- Significant changes in Airflow version 2.8.0:
 - Raw HTML code in DAG docs and DAG params descriptions are disabled by default ([#35460](#))
 - Account for change in UTC offset when calculating next schedule ([#35887](#))
 - For the complete list of changes, see [Airflow 2.8.0 documentation](#).
- Significant changes in Airflow version 2.8.1:
 - Target version for core dependency pendulum package set to 3 ([#36281](#))
 - Airflow packaging specification follows modern Python packaging standards ([#36537](#))
 - Graphviz dependency is now an optional one, not required one ([#36647](#))
 - Do not let EventsTimetable schedule past events if catchup=False ([#36134](#))
 - For the complete list of changes, see [Airflow 2.8.1 documentation](#).
- Significant changes in Airflow version 2.8.2:
 - The allowed_deserialization_classes flag now follows a glob pattern ([#36147](#))
 - The audit_logs permissions have been updated for heightened security ([#37501](#))
 - AirflowTimeoutError is no longer except by default through Exception ([#35653](#))
 - Revoking audit_log permission from all users except admin ([#37501](#))
 - For the complete list of changes, see [Airflow 2.8.2 documentation](#).
- Significant changes in Airflow version 2.8.3:
 - The smtp provider is now pre-installed when you install Airflow ([#37713](#))
 - For the complete list of changes, see [Airflow 2.8.3 documentation](#).
- Significant changes in Airflow version 2.8.4:
 - There are no significant changes in Airflow 2.8.4. For full list of changes, see [Airflow 2.8.4 documentation](#).
- Significant changes in Airflow version 2.9.0:
 - Listener API methods are considered stable and can be used for production system (They were experimental features in older Airflow versions) ([#36376](#))
 - Dataset URIs are now validated on input ([#37005](#))
 - Add REST API actions to Audit Log events ([#37734](#))
 - Prevent large string objects from being stored in the Rendered Template Fields ([#38094](#))
 - The xcom change allows the storage of larger Xcom values in Airflow metadata 2^32 instead of 2^16 bytes. ([#38401](#))
 - Stronger validation for key parameter defaults in taskflow context variables ([#38015](#))
 - Deprecate smtp configs in airflow settings / local_settings ([#37711](#))
 - For the complete list of changes, see [Airflow 2.9.0 documentation](#).
- Significant changes in Airflow version 2.9.1:
 - Stackdriver logging bugfix requires Google provider 10.17.0 or later ([#38071](#))
 - For the complete list of changes, see [Airflow 2.9.1 documentation](#).
- Significant changes in Airflow version 2.9.2:
 - There are no significant changes in Airflow 2.9.2. For full list of changes, see [Airflow 2.9.2 documentation](#).
- Significant changes in Airflow version 2.9.3:
 - Time unit for scheduled_duration and queued_duration changed (metrics) ([#37936](#))
 - For the complete list of changes, see [Airflow 2.9.3 documentation](#).

Airflow upgrade from version 2.9.3 to version 2.10.4

- Significant changes in Airflow 2.10.0:
 - Scarf-based telemetry: Airflow now collects telemetry data ([#39510](#))
-  **Note:** Disabled in Cloudera Data Engineering.
- Datasets no longer trigger inactive DAGs ([#38891](#))
 - try_number is no longer incremented when a task is performed ([#39336](#))
 - /logout endpoint in Flask AppBuilder (FAB) Auth Manager is now Cross-Site Request Forgery (CSRF) protected ([#40145](#))
 - OpenTelemetry Traces for Apache Airflow ([#37948](#))
 - Decorator for Task Flow (@skip_if, @run_if) to make it simple to skip a task. ([#41116](#))
 - Using multiple executors concurrently ([#40701](#))
- Significant changes in Airflow 2.10.3:
 - Enhancing BashOperator to execute templated bash scripts as temporary files ([#44641](#))
- Significant changes in Airflow 2.10.4:
 - TaskInstance priority_weight is capped in 32-bit signed integer ranges ([#43611](#))

Debugging Airflow DAGs locally

You can learn about how to debug Airflow DAGs locally to identify any compatibility issues while upgrading Cloudera Data Engineering.

About this task

The example is based on the [Airflow DAG debug documentation](#).

Procedure

- Add the following two lines at the end of the DAG file that you want to test:

```
if __name__ == "__main__":
    dag.test()
```

For example,

```
from datetime import timezone
from dateutil import parser
from airflow.operators.bash_operator import BashOperator
from airflow.operators.python_operator import PythonOperator
from airflow import DAG
default_args = {
    'owner': 'test',
    'depends_on_past': False,
    'email': ['airflow@example.com'],
    'email_on_failure': False,
    'start_date': parser.isoparse('2020-11-11T20:20:04.268Z').replace(tzinfo=timezone.utc),
    'email_on_retry': False,
    'user': 'notused'
}

dag = DAG(
    'some-dag-3',
    default_args=default_args,
```

```

    catchup=False,
    schedule_interval=None,
    is_paused_upon_creation=False,
)

def print_function():
    print("Task executed.")

print_task = PythonOperator(
    task_id='print_task',
    dag=dag,
    python_callable=print_function
)

echo = BashOperator(
    task_id='run_this_first',
    bash_command='echo 1',
    dag=dag,
)

echo >> print_task

if __name__ == "__main__":
    dag.test()

```

2. Prepare a DAG testing environment by running the following commands:

```

python -m venv venv-airflow
source venv-airflow/bin/activate

(venv-airflow) airflow pip install 'apache-airflow[http,celery,kubernetes]
==2.9.3' --constraint "https://raw.githubusercontent.com/apache/airflow/
constraints-2.9.3/constraints-3.11.txt"

# check the configured DB used for testing the DAGs (a local sqlite db sho
uld be enough)
airflow config get-value database sql_alchemy_conn
sqlite:///Users/$USER/airflow/airflow.db

# if there are DB errors missing tables or other reset the DB
airflow db reset

```

3. Run the DAG and check for the warnings. For example,

```

python -W default example_dag.py
/Users/$USER/workload/airflow/bashop_pythonop_dag.py:4 DeprecationWarning:
The `airflow.operators.bash_operator.BashOperator` class is deprecated.
Please use `airflow.operators.bash.BashOperator`.
/Users/$USER/workload/airflow/bashop_pythonop_dag.py:5 DeprecationWarning:
The `airflow.operators.python_operator.PythonOperator` class is
deprecated. Please use `airflow.operators.python.PythonOperator`.
/Users/$USER/workload/airflow/bashop_pythonop_dag.py:19
RemovedInAirflow3Warning: Param `schedule_interval` is deprecated and
will be removed in a future release. Please use `schedule` instead.
[2024-08-22T12:20:33.673+0200] {dag.py:4199} INFO - dagrun id: some-dag-3
[2024-08-22T12:20:33.678+0200] {dag.py:4215} INFO - created dagrun <DagRu
n some-dag-3@2024-08-22T10:20:33.616669+00:00:manual__2024-08-22T10:20:33
.616669+00:00,state:ru
...

```

where the -W flag checks for DeprecationWarnings in the code. For more information, see [Check for DeprecationWarning in your code](#).

Using the backup-restore-based upgrade script

Learn about how to use the backup-restore-based upgrade script in Cloudera Data Engineering on cloud.

About this task

This procedure applies to Cloudera Data Engineering versions 1.20.3-h2, 1.21.0-h2, 1.22.0, and higher.

If Apache Airflow connections and variables are involved, the original backup-restore-based upgrade is not applicable. This document describes a script-based fallback for performing the backup-restore-based upgrade that you can use when the default upgrade method described in [Handling upgrade failures for Cloudera Data Engineering](#) does not work.

Table 2: Contents included in the backup

Artifacts	Included in the automation script
Service config	Y
Service Logs	N
Virtual Cluster config	Y
Virtual Cluster end-points	N
Virtual Cluster event logs	N
Job: Spark	Y
Job: Airflow	Y
Resource: files	Y
Resource: docker runtimes	Y
Resource: python-venv (for Spark)	Y
Resource: python-venv (for Airflow)	Y (Since 1.21.0)
Git Repository	Y
Credentials	Y
Spark Session	N
Spark Session logs (Including statement history)	N
Job Runs	Y (Since 1.22.0)
Job Run logs (Driver, Executor, API)	N
Airflow DAG logs	N
Airflow connections	Y
Airflow variables	Y

Once the backup is complete, Service and Virtual Cluster metadata are stored in the Object Store. Data related to jobs and resources is stored locally, at the location where you invoked the backup command.

Before you begin

Before running the script, you must install the listed tools:

- CDP CLI

- `cdpcurl`
 - CDE CLI
 - `jq`
 - `kubectl`
1. Install CDP CLI. For more information, see [Installing Cloudera client](#).
 2. Install `cdpcurl`. For more information, see [cdpcurl](#).
 3. Download the CDE CLI from your Virtual Cluster page and add its path to the `PATH` environment variable with the export `PATH="$PATH:[***CLI-PATH***]"` command.
 4. Install the other tools according to your system requirements.

Procedure

1. Configure the CDP CLI.

- a) On the Cloudera Management Console, navigate to your Profile page.
- b) On the Access Keys tab, click Generate Access Key.
- c) Download the credentials file and copy the contents to `~/.cdp/credentials`

The credentials are set as the default profile. You can also rename the profile. If you do so, set the CDP profile to your preference using one of the listed methods:

- The `CDP_PROFILE` environment variable
- The `cde-service-backup-restore-utils.properties` file
- The `--cdp-profile` command line option

2. Configure the CDE CLI.

Configure the CDE CLI following the instructions in [Configuring the CLI client](#). Cloudera recommends that you use the Cloudera credentials, as they are needed for CDP CLI too.

- a) Set the CDE CLI using one of the listed methods:

- The environment variable
- The `cde-service-backup-restore-utils.properties` file
- The `~/.cde/config.yaml` file



Note: Command line options are not available.

- b) Run the following command once to make sure that the configuration works:

```
cde job list --vcluster-endpoint [***JOBS-API-URL***]
```

3. Back up the Cluster.

- a) Download the backup-restore-based upgrade script (cde-service-backup-restore-utils.sh) to your local machine: navigate to <https://github.com/cloudera/cde-public/blob/master/scripts/cde-service-backup-restore-utils.sh> and download the raw file using the download icon on the right.
- b) Authenticate to kubectl to access your Cluster.
- c) To back up the Cluster, run the following command.



Note: Using --pre-check is optional, but recommended on the first attempt. With --pre-check, you can check the output of the tools in advance, to make sure the tools are properly installed and configured.

```
./cde-service-backup-restore-utils.sh backup-service -s <cluster-id> -b <backup-base-directory> --pre-check
```

Some of the contents are stored remotely and some are stored locally in the directory you specify.



Note: -b indicates the backup base directory, which is optional. If not specified, the default \$HOME/.cde/service-backups value is used. Backup files are saved under [***BASE-DIR**]/[***SERVICE-ID**]/, which means that the same directory is shared between multiple backups with the same base directory and service ID. If you want to preserve multiple backups of the same service, specify a different base directory. When you are sure that the backup is no longer needed, delete the backup files.

4. Restore the Cluster.

- a) Run the following command to provision the Cluster and Virtual Clusters. This can take around an hour.



Note: To retain the endpoints of a VC, obtain the original service ID, after that, make sure that you delete the original Cloudera Data Engineering service. To obtain the service ID, on the Cloudera Data Engineering UI, select Administration in the left navigation menu. Select the service and check the ID on the Service Details page.

To retain the endpoints of a VC, include the --service-id [***ORIGINAL-SERVICE-ID***] parameter in the restore-service CLI command.

If you do not specify a cluster ID in the restore-service command, a random ID is generated.

```
./cde-service-backup-restore-utils.sh restore-service -s <original-cluster-id> -t <new-cluster-id> -n <new-cluster-name> -b <backup-base-directory> --service-only
```

- b) Authenticate again to kubectl to access your Cluster.

If the version of the new service is different from the original one, download the CDE CLI tool of the new service and add its path to the *PATH* environment variable.



Note:

Ideally, all scheduled jobs are paused before taking the backup. If the scheduled jobs were not paused, add the --pause-jobs flag to the command to pause them on restore. If you add this flag, the script modifies the Airflow DAGs also to stop catch up.

To restore your contents, run:

```
./cde-service-backup-restore-utils.sh restore-service -s <original-cluster-id> -t <new-cluster-id> -n <new-cluster-name> -b <backup-base-directory> --contents-only
```

Upgrading Airflow if DAGs and packages are compatible with new Airflow version

You can upgrade Airflow after ensuring the DAGs and the packages in the Airflow Custom Libraries and Operators are compatible with the new Airflow version by using the requirements.txt file.

Before you begin

**Note:**

To avoid issues after the upgrade, ensure that the DAGs and the packages in the Airflow Custom Libraries and Operators are compatible with the new Airflow version. Fixing the issues after the upgrade process is complete might cause further downtime in performing the job.

Ensure that the catchup option is not enabled for any user's Airflow jobs. Before the backup starts, if the Airflow DAG catchup options are enabled, disable them manually.

Procedure

1. Get the requirements.txt file for the current Airflow Custom Libraries and Operators.
After the upgrade, you need the requirements.txt file to restore Python environments. For information on how to get the requirements.txt file, see *In-place upgrade with Airflow Operators and Libraries*.
2. Verify the requirements.txt file against an Airflow constraints file.
For more information, see *In-place upgrade with Airflow Operators and Libraries*.
 - a) If you run into errors during the verification, update the dependencies in the requirements.txt file.
3. Disable all Airflow Custom Libraries and Operators in all virtual clusters in the Cloudera Data Engineering service.
4. Perform the upgrade process.
5. Using the saved requirements.txt file from step 2 on page 64, build and activate the Airflow Python environment in the virtual clusters.
If the build fails because of non-python related dependency issues, you may need to make version adjustments in the requirements.txt file.
6. On the Airflow UI and Cloudera Data Engineering jobs UI, verify that the Airflow jobs are running.
In the case of DAG-related errors:
 - If the DAG parsing fails, look into the Airflow UI to identify the impacted DAGs.
 - If the DAG execution fails, look into the Airflow job logs to identify the impacted DAGs.
7. Fix the impacted DAGs or update the dependencies in the requirements.txt file.

Related Information

[In-place Upgrade with Airflow Operators and Libraries](#)

Handling upgrade failures for Cloudera Data Engineering

If your upgrade of Cloudera Data Engineering fails, you have the option to clone the service with the latest version of Cloudera Data Engineering. Learn how to handle an upgrade failure.

About this task

During a Cloudera Data Engineering upgrade, a backup is created as part of the upgrade preparation process. The backup comprises all metadata and associated resource files, which are stored in the Object Store. This procedure uses that backup to be restored in a new cluster.

The list of service backups is available in the Backup Library. To locate the Backup Library, in the left navigation menu of Cloudera Data Engineering select Administration, select Service Details, select the Maintenance tab, and select Backup Library.

To obtain the list of all available backups, in the CDP CLI, run:

```
cdp de list-backups
```

To obtain the list of service backups associated with a specific Cloudera environment, run "cdp de list-backups --filter "environment(eq)[***CDP ENVIRONMENT NAME***]"

The Cloudera Data Engineering backup includes the following:

- Cloudera Data Engineering Service configurations
- Virtual cluster names
- Virtual cluster configurations
- Virtual cluster file-based resources
- Spark job definitions
- Airflow job definitions
- Spark Python-env resources

The following are not yet included in the backup:

- Non file-based resources, for example, Python-venv resources and custom runtimes
- Airflow custom operators & libraries
- Logs
- Job run history
- Endpoints

Before you begin



Important: This procedure changes the virtual cluster endpoints.

1. Ensure that the catchup option is not enabled for any user's Airflow jobs.

Before the backup starts, if the Airflow DAG catchup options are enabled, disable them manually.

2. By default, the restored Cloudera Data Engineering service is assigned the name of the original backed-up service, and a new service ID is generated. To prevent backup failure due to naming conflicts, choose one of the following options:
 - a. Delete the original service that failed to upgrade during the Cloudera Data Engineering upgrade.
 - b. Rename the service using the `--service-name` option.



Note: To retain endpoint stability, you can explicitly assign the restored service the ID of the original backed-up service, using the `--service-id` option of the `cdp de restore-service` command.

Assigning a new ID results in new Fully Qualified Domain Names (FQDNs) for the service and all its virtual clusters, which affects the endpoint stability.

A valid service ID is:

- An 8 character-long alphanumeric string
- Does not contain vowels
- Unique

Procedure

1. Restore the service from the backup.

```
cdp de restore-service --backup-id <backup-id> --environment-crn
<environment-crn>
```

Where:

backup-id

The ID of the backup that you are restoring from.

environment-crn

The Customer Resource Number (CRN) of the Cloudera environment with which a restored Cloudera Data Engineering service is associated. Currently, you can restore the Cloudera Data Engineering service only to the same Cloudera environment to which the backed-up service is associated.

For example:

```
cdp de restore-service --backup-id 2 --environment-crn crn:cdp:environme
nts:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:environment:c67b9089-
2d3b-4579-861d-c0df12a105b1
```

2. To obtain a list of backups, run:

```
cdp de list-backups
```

3. To describe a particular backup, run:

```
cdp de describe-backup --backup-id <backup-id>
```

For example:

```
$ cdp de describe-backup --backup-id 2 --profile priv
{
  "backup": {
    "id": 2,
    "serviceID": "cluster-cf6h74lq",
    "serviceName": "dex-priv-default-azure-env-1689008683873",
    "environmentName": "dex-priv-default-azure-env",
    "environmentCrn": "crn:cdp:environments:us-west-1:9d74eee4-1cad-45
d7-b645-7ccf9edbb73d:environment:c67b9089-2d3b-4579-861d-c0df12a105b1",
    "creator": "crn:altus:iam:us-west-1:9d74eee4-1cad-45d7-b645-7ccf
9edbb73d:user:0f9a97a7-23a7-43bd-bc71-ecdb2aa34ed5",
    "cloudPlatform": "AZURE",
    "status": "completed",
    "created": "2023-07-17T18:02:58.385455Z"
  }
}
```

4. In the case of an Airflow DAG failure, identify the impacted DAG on the Airflow UI and fix it.

For more information, see the DAG-related steps in *In-place upgrade with Airflow Operators and Libraries*.

Related Information

[Upgrading Cloudera Data Engineering](#)

[In-place Upgrade with Airflow Operators and Libraries](#)

Handling ineligible upgrades for Cloudera Data Engineering

In certain scenarios, your Cloudera Data Engineering service is not eligible for an in-place upgrade to the latest Cloudera Data Engineering version. In such a scenario, follow a manual upgrade process where administrators create a new service and import all jobs and resources. Learn how to conduct a workaround for each scenario.

Older Cloudera Data Engineering versions

If the version of your Cloudera Data Engineering service is older than six months, it is outside of the upgrade window. Older Cloudera Data Engineering versions do not support in-place upgrade and you cannot use the process listed in *Handling upgrade failures for Cloudera Data Engineering*. To check whether your Cloudera Data Engineering version is eligible for the in-place upgrade option, see [Support lifecycle policy](#).

Cloudera Data Engineering provides the in-place upgrade option from the eligible versions to the latest version. If you are on a release that is ineligible for the in-place upgrade, the in-place upgrade option is disabled and you can perform a backup and restore. For more information, see *Cloudera Data Engineering upgrade version compatibility*.

Solution: A manual upgrade option is available. To upgrade, create a new Cloudera Data Engineering Service and transfer your jobs to the new service by following the instructions in *Backing up and restoring Cloudera Data Engineering jobs*.

Older Data Lake versions

If your Cloudera Data Engineering service uses an old Data Lake version, and therefore it is not compatible with the latest Cloudera Data Engineering version, your Cloudera Data Engineering service is not eligible to perform either an in-place upgrade or a manual upgrade. For information about Cloudera Data Engineering and its compatible Data Lake version, see *Compatibility for Cloudera Data Engineering and Runtime components*.

Solution: Upgrade your Data Lake to a compatible version associated with the Cloudera Data Engineering version. For more information about upgrading Data Lake, see *Upgrading a Data Lake*.

Hotfix version of Cloudera Data Engineering

There is no upgrade available to Cloudera Data Engineering hotfix versions.

Related Information

[Handling upgrade failures for Cloudera Data Engineering](#)

[Cloudera Data Engineering upgrade version compatibility](#)

[Backing up Cloudera Data Engineering jobs on remote storage](#)

[Compatibility for Cloudera Data Engineering and Runtime components](#)

[Upgrading a Data Lake](#)

[In-place Upgrade with Airflow Operators and Libraries](#)