Open for Innovation
KNIME

Kerberos User Guide

KNIME AG, Zurich, Switzerland
Version 5.1 (last updated on 2020-04-28)

Table of Contents

Introduction 1
Overview 1
Kerberos Configuration 3
Use system defaults (**discouraged**) 3
Use Kerberos client configuration file (`krb5.conf`) 3
Use realm and KDC 4
Log into Kerberos 5
How to log in 5
Status bar 6
Status 8
Debug Logging 10
Glossary 12

Introduction

This guide describes step-by-step how to configure Kerberos on KNIME Analytics Platform.

Kerberos, which dates back to 1993, is a network authentication protocol for distributed applications, and support for it is also integrated into KNIME Big Data Extensions and KNIME Extension for Apache Spark.

You can use Kerberos authentication to connect to a wide-array of Kerberos-secured services. Using KNIME Analytics Platform with the Big Data Extensions you can e.g connect to Kerberos-secured Hive clusters using Hive Connector node, or Impala using Impala Connector node.

Overview

To configure Kerberos in KNIME Analytics Platform, go to *File > Preferences > KNIME > Kerberos* and open the Kerberos preferences page.

As shown in Figure 1, connecting to Kerberos consists of the following two high-level steps:

1. Setup Kerberos configuration.
2. Log into Kerberos (i.e. obtain a Kerberos ticket).

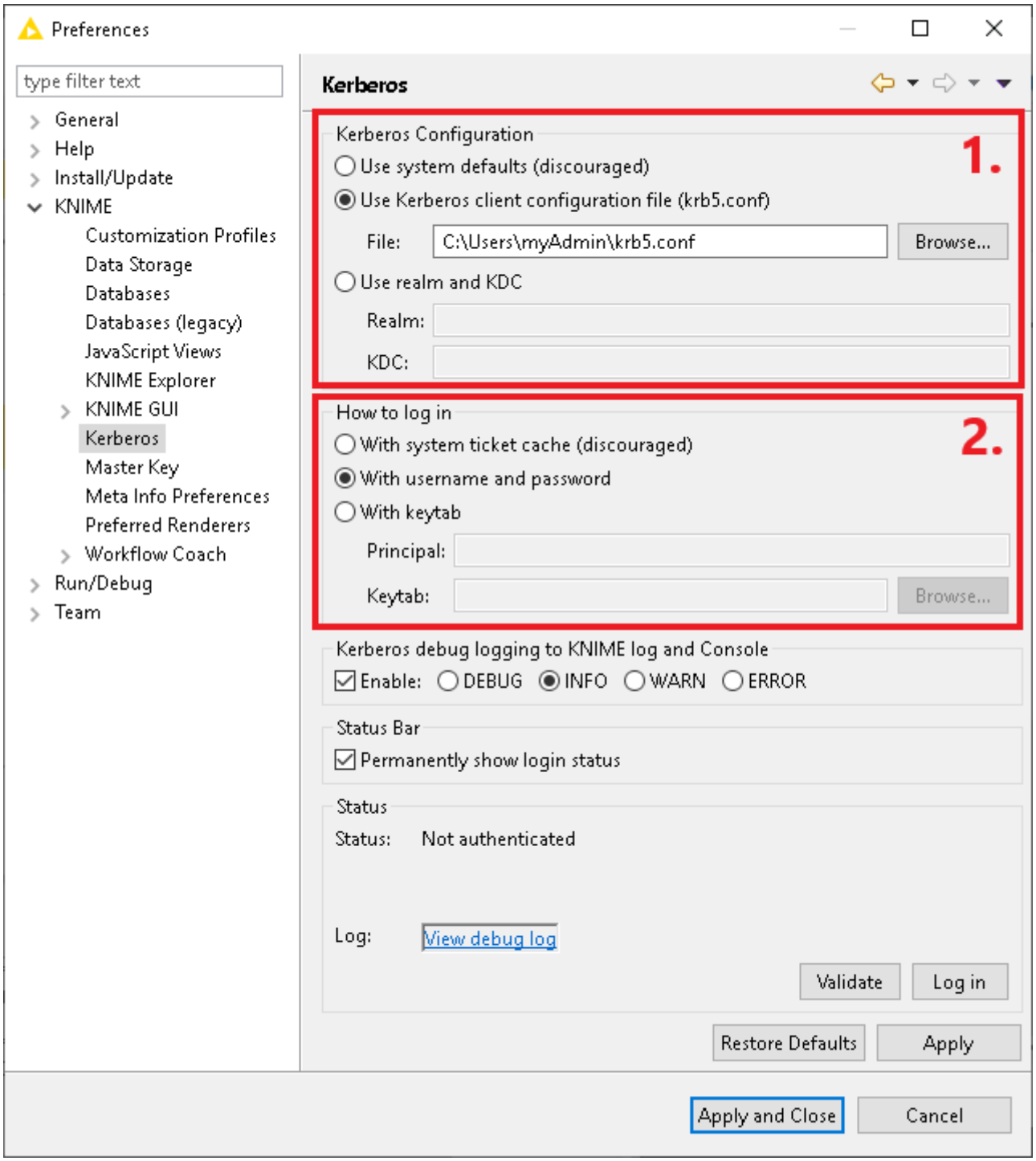

Figure 1. Kerberos preferences page

The following sections will explain these steps in more details.

If any of the Kerberos terminology is unclear, you can refer to the Glossary at the bottom of this guide.

Kerberos Configuration

The first step is to configure Kerberos in KNIME Analytics Platform (see Figure 1). One of the three options provided can be selected, depending on your needs and the environment setup of the system in use.

Use system defaults (**discouraged**)

This option is **discouraged**, because the correct setup is highly dependent on the system environment. In addition, if you move the preferences file or KNIME Analytics Platform to another machine, it might not work anymore. However, select this option if you want to keep existing setup running, or if you want to manage Kerberos configuration outside KNIME Analytics Platform.

When selecting this option KNIME Analytics Platform will look at a set of default locations for the `krb5.conf` (Kerberos client configuration file).

Possible locations for `krb5.conf`

KNIME Analytics Platform will try the following locations in the given order:

1. First, it checks whether the `java.security.krb5.conf` system property is set. If so, it will try to read the file from the location specified in this system property.
2. Otherwise, it will try to read the `krb5.conf` from the Java Runtime Environment of KNIME Analytics Platform:

   ```
   <knime-analytics-platform-
   installation>/plugins/org.knime.binary.jre.<version>/jre/lib/security/krb5.conf
   ```

3. If the above also fails, it will try the following operating system dependent locations:
   - Windows: `C:\Windows\krb5.ini`
   - Linux/MacOS: `/etc/krb5.conf`

For more information, please refer to the Kerberos documentation.

Use Kerberos client configuration file (`krb5.conf`)

This option is the recommended way to configure KNIME Analytics Platform, since it allows

for full configurability of the settings in the Kerberos client configuration file (`krb5.conf`).

First, you need to obtain a valid `krb5.conf` file. If you don't know where the file is located or if you don't have one already, please contact your local administrator. Alternatively, you can write a `krb5.conf` file yourself. A minimal configuration file could look like this:

```
[libdefaults]
default_realm = MYCOMPANY.COM

[realms]
MYCOMPANY.COM = {
  kdc = kdc.mycompany.com
  admin_server = kdc.mycompany.com
}
```

The above example declares that you are in a Kerberos realm called `MYCOMPANY.COM` and that the hostname of the Kerberos KDC is `kdc.mycompany.com`. Adjust these values as appropriate for your setup. Depending on your setup, more configuration settings may be necessary. The `krb5.conf` format is fully described as part of the MIT Kerberos documentation.

Now, move the `krb5.conf` file into a location of your choice, where it can be accessed by KNIME Analytics Platform, and enter the file path in the Kerberos preferences page. It is recommended to store the file outside of the KNIME Analytics Platform installation folder, to avoid accidentally deleting it during upgrades.

Use realm and KDC

The easiest way is to insert, directly in the Kerberos preferences page:

- The name of the realm (the name needs to be in uppercase letters)
- The IP or hostname of the KDC

Based on the input realm and KDC, the `krb5.conf` file will be generated.

Log into Kerberos

How to log in

After Kerberos is configured, the next step is to select one of the following authentication methods to log into Kerberos.

With system ticket cache (**discouraged**)

This option is **discouraged**.

Select this option if you want the ticket-granting ticket (TGT) to be obtained from the system ticket cache. The ticket cache will be searched in the following locations:

- On Solaris and Linux: `/tmp/krb5cc_uid` where `uid` is numeric user identifier.
- On Windows:
  - `C:\Users\<username>\krb5cc_<username>`.
  - Otherwise, if the file does not exist or if it does not contain a valid TGT, the TGT will be obtained from the Local Security Authority (LSA) API.

On recent Windows versions, further changes to the Windows Registry are required for Java processes such as KNIME Analytics Platform to read the TGT from LSA. This type of setup is not recommended as it poses a security risk.

If a valid TGT is present on the system, no further action is required to log into Kerberos.

For more information on this, please check the Oracle documentation.

With username and password

Using username and password is the recommended way to log into Kerberos. Username and password will be prompted at login time.

With keytab

Select this option to use *Principal* and *Keytab*. No further user interaction is then required to

log into Kerberos.

Status bar

Further down in the preferences page, under the *Status Bar* section, the option *Permanently show login status* can be selected, as shown in Figure 2.

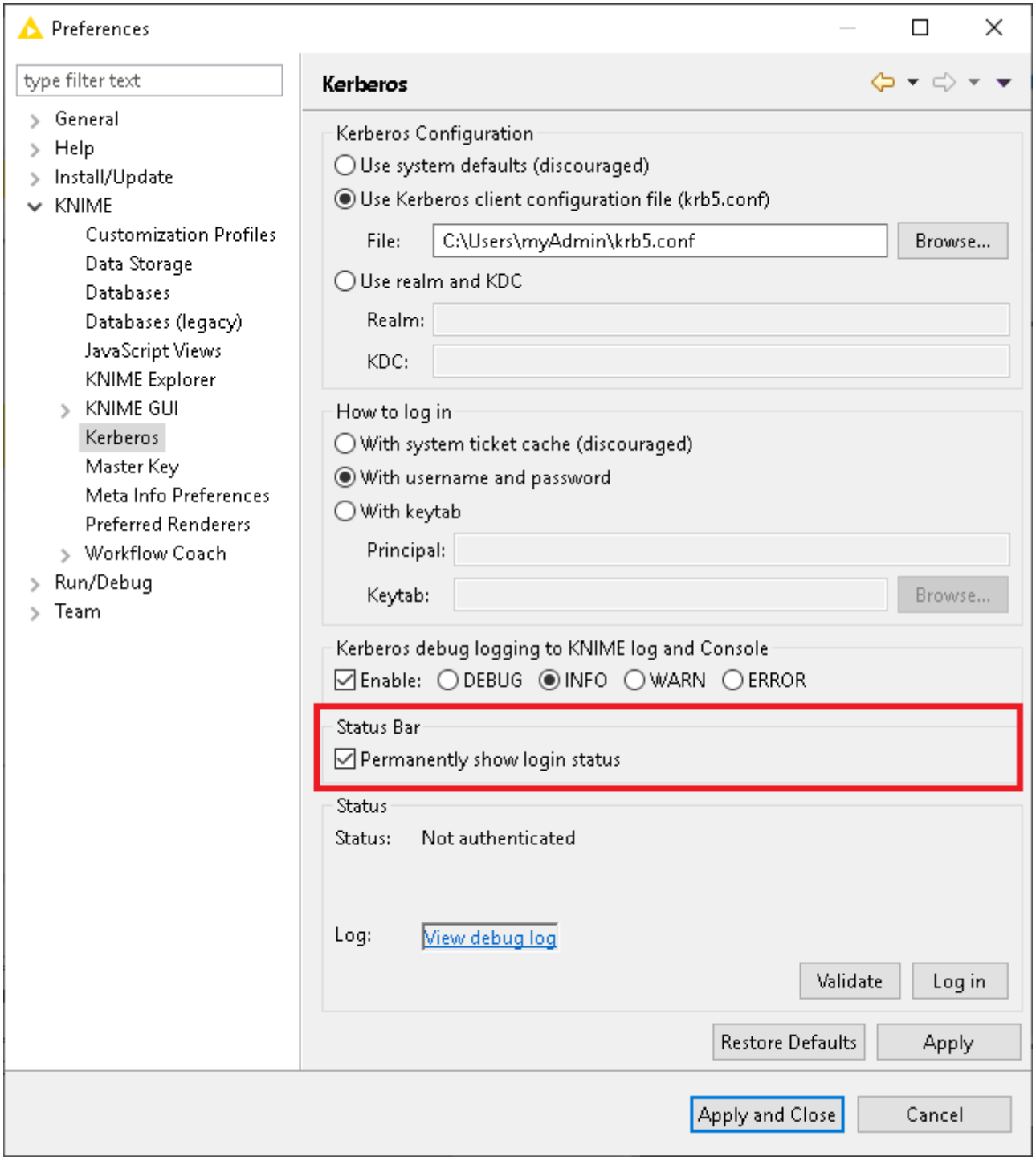

Figure 2. Kerberos status bar

Enabling this option will show the Kerberos login status in the lower bar of KNIME Analytics

Platform (see Figure 3). The advantage of this option is that you can do login/logout and check the Kerberos status anytime without having to open the Kerberos preferences page.

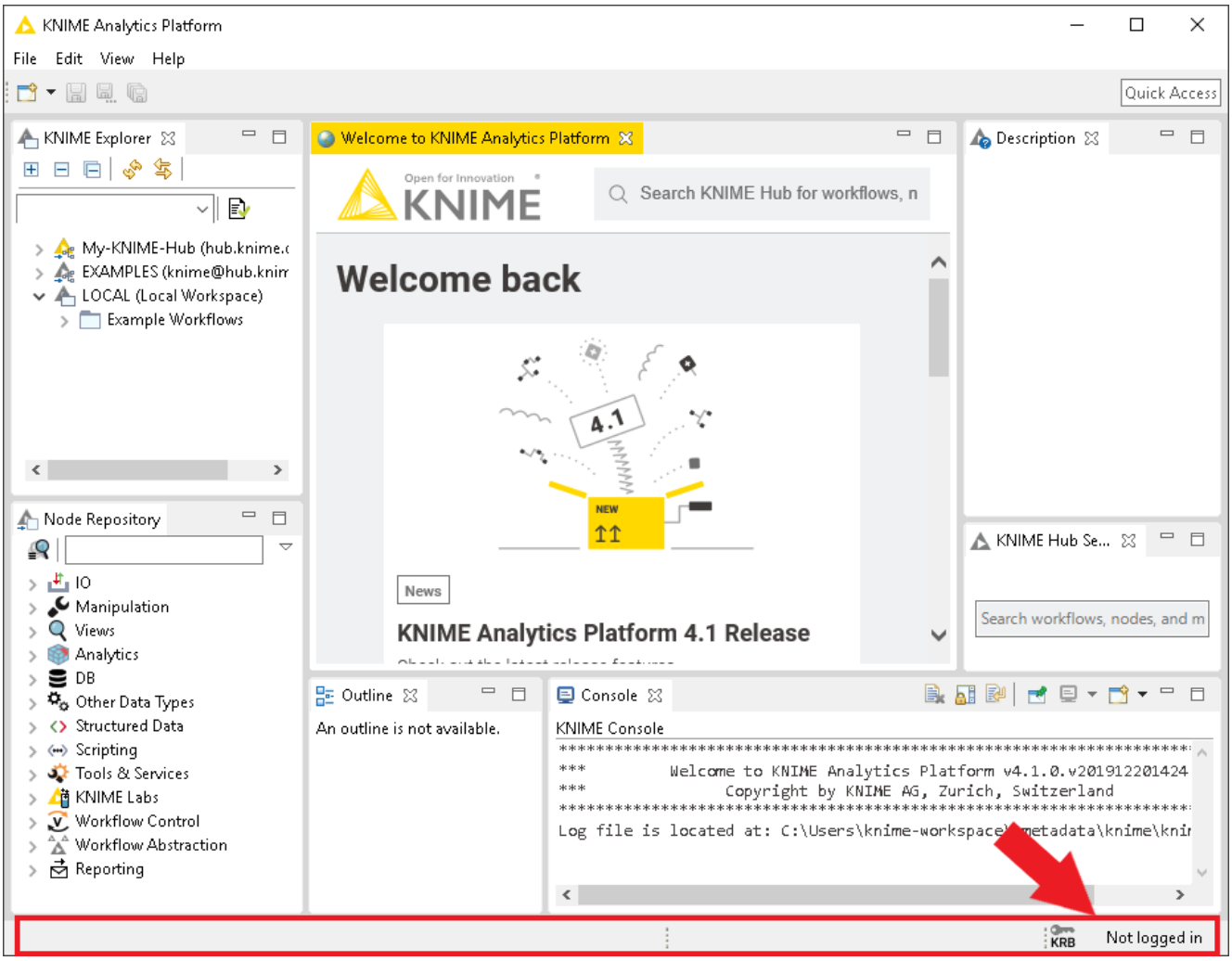

Figure 3. Kerberos status bar in KNIME Analytics Platform

Right-clicking on the status bar will open a menu containing two options:

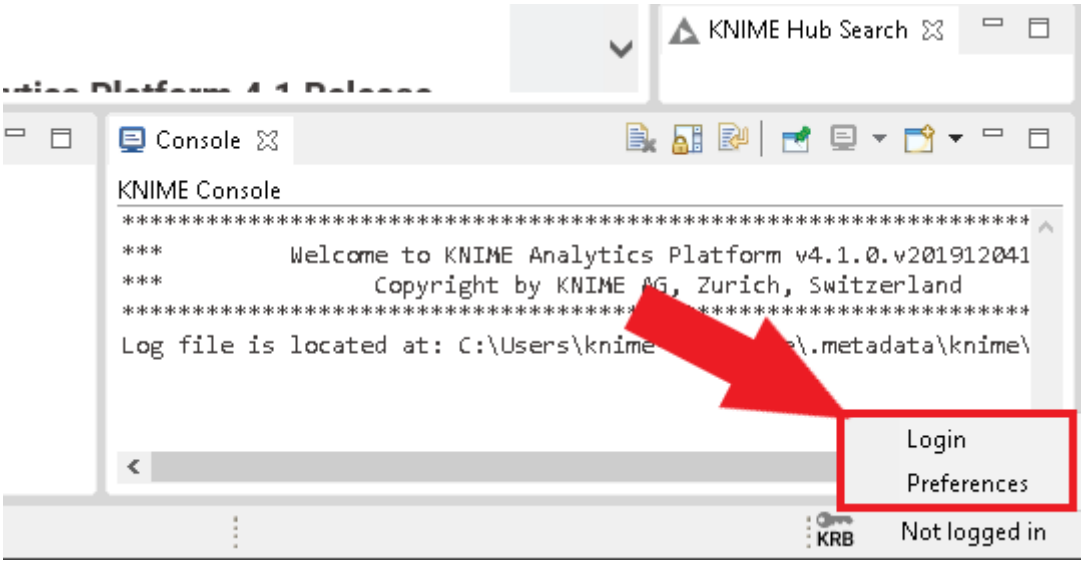

Figure 4. Kerberos status bar menu

- *Login* triggers the login process. Selecting username and password to login, a pop-up window will open. Here, you can enter your Kerberos credentials (see Figure 5 below).

Double-clicking on the status bar will also trigger the login process.

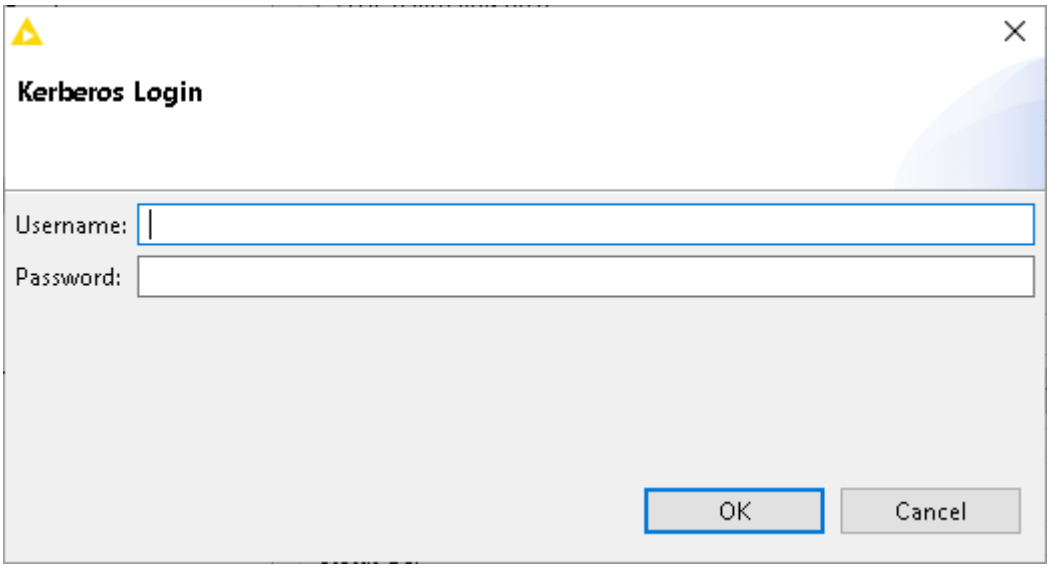

Figure 5. Login prompt for username and password

- *Preferences* opens the Kerberos preferences page. This can also be achieved by going to *File* > *Preferences* > *KNIME* > *Kerberos*.

Status

After configuring and selecting the authentication method, you can check the validity of your settings in the lower part of the preferences page.

Here, several information are shown:

- *Status* shows the Kerberos status in general, e.g it will show red messages if any error occurs.
- Under *Log*, click *View debug log* to view Kerberos debug log messages. A pop-up window will appear showing all log messages related to Kerberos. Please make sure to enable Kerberos debug logging beforehand (please check the Debug Logging section for more information).
- Click the *Validate* button to validate the Kerberos configuration.
- Click the *Log in* button to log into Kerberos.

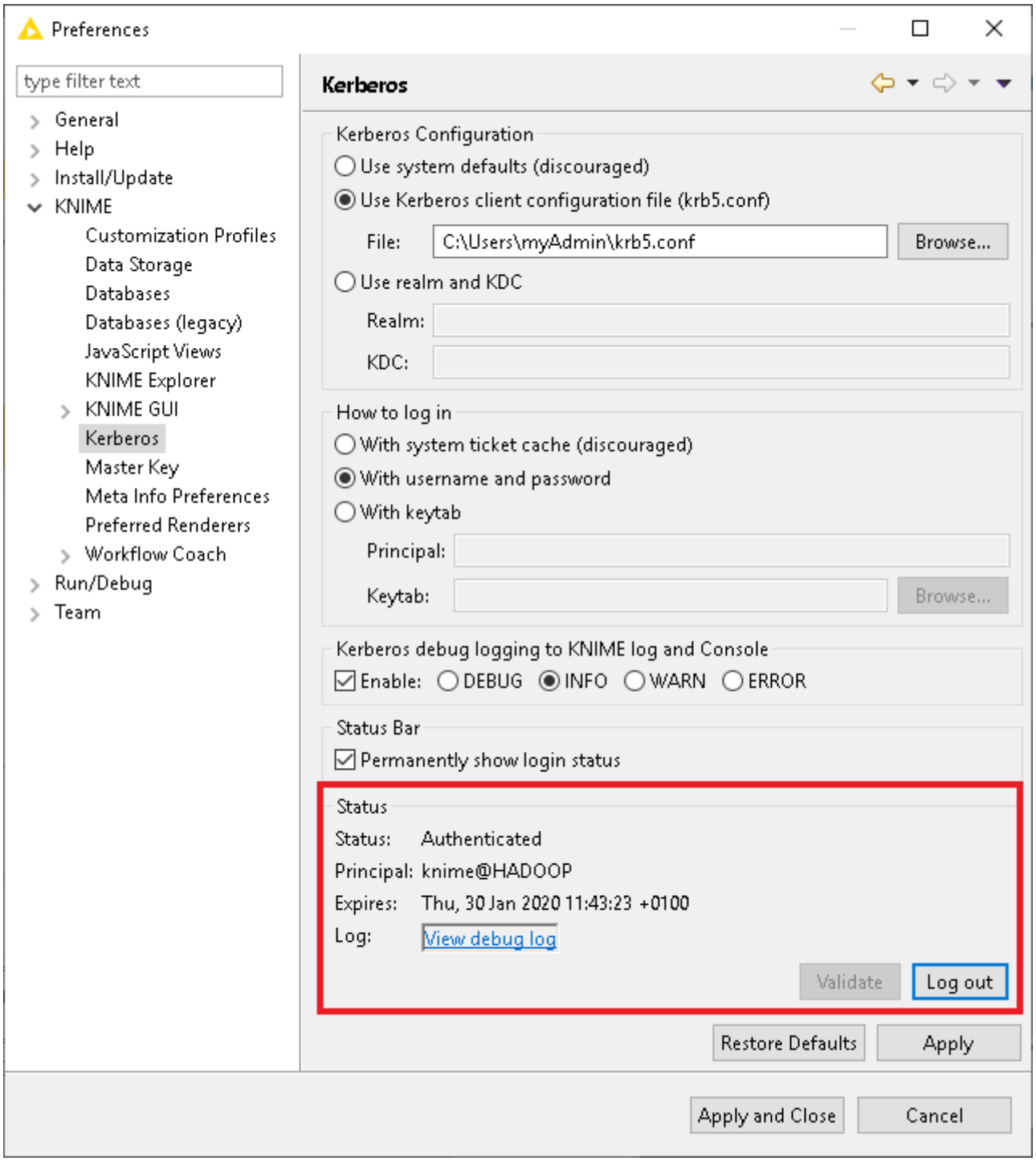

Figure 6. Kerberos status

Debug Logging

If you encounter problems with the Kerberos setup it is helpful to enable Kerberos logging to get more information about the problem. To enable Kerberos logging, simply check the option *Kerberos debug logging to KNIME log and Console* in Kerberos preferences page (go to *File > Preferences > KNIME > Kerberos*).

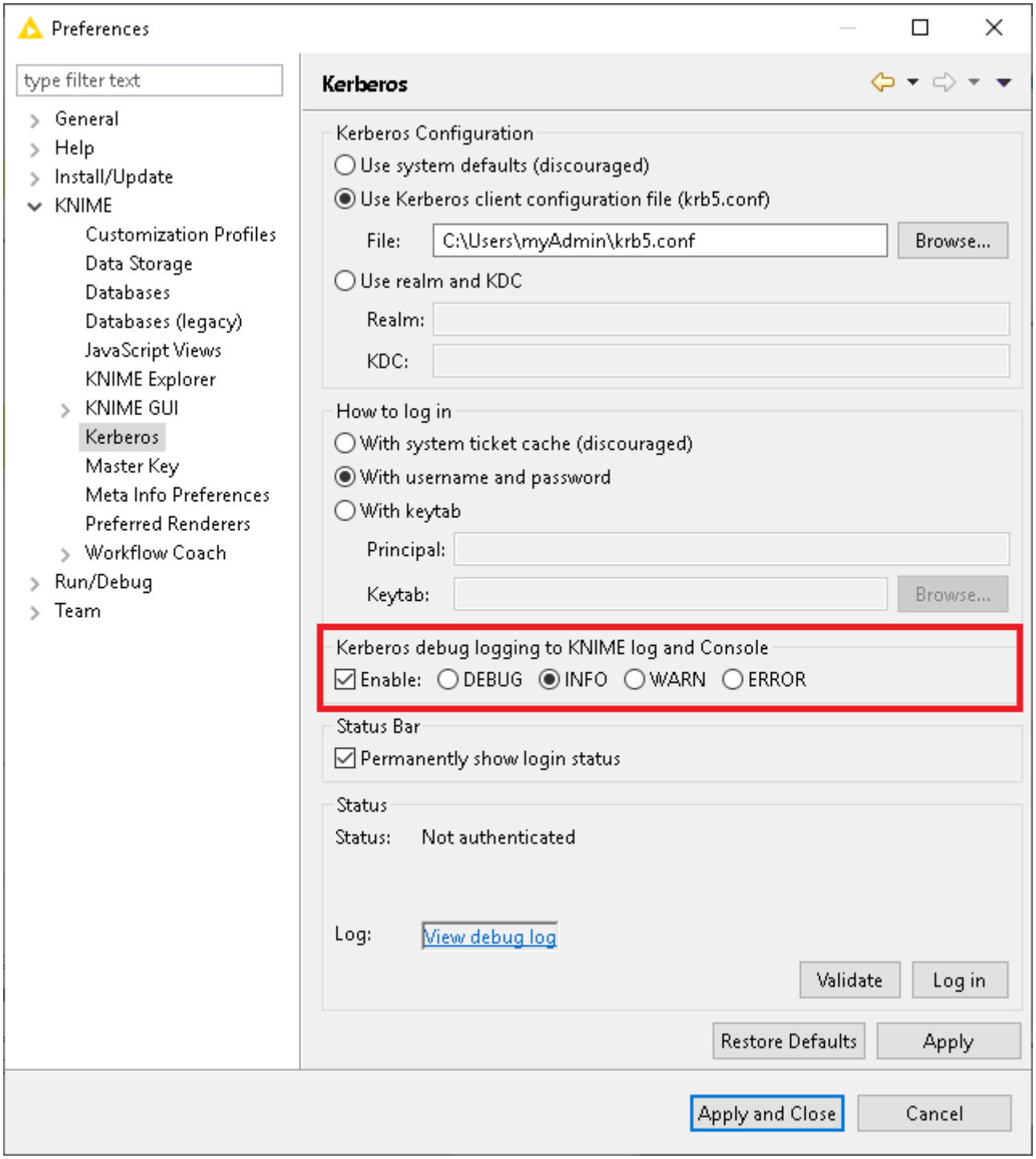

Figure 7. Enable debug log

After a restart of KNIME Analytics Platform, additional Kerberos information will be displayed

in the KNIME Console and KNIME log file.

You should restart KNIME Analytics Platform for the changes to be effective.

Figure 8. View debug log

To see only Kerberos-related log messages, click *View debug log* in the Kerberos preferences page. A new pop-up window will open containing Kerberos debug log messages (see Figure 8). Note that you need to enable the option *Kerberos debug logging to KNIME log and Console* beforehand to be able to see the log messages.

Glossary

- **KDC** Key Distribution Center, a server that handles Kerberos authentication.
- **Principal** The Kerberos-equivalent to a username. In Kerberos, principals identify users or services. Examples:
  - A user principal: `joe@MYCOMPANY.COM`
  - A service principal (in this case for Hive Server 2): `hive/server.mycompany.com@MYCOMPANY.COM`
- **Realm** Indicates an administrative domain. Both users and services are registered as principals with their passwords in a realm. Example: `MYCOMPANY.COM`
- **Ticket** A piece of data that serves as proof that you have authenticated yourself as a principal.
- **Ticket cache** Holds your Kerberos tickets. In order to work with KNIME Analytics Platform, tickets need to be stored in a file-based ticket cache, as specified by the KRB5CCNAME environment variable.

KNIME AG
Talacker 50
8001 Zurich, Switzerland
www.knime.com
info@knime.com