

Configuration and Setup for HDF5 HDFS VFD

The HDF5 configure options to enable the HDFS vfd are:

- Autotools configure `--with-libhdfs=$HADOOP_HOME`
- CMake option `HDF5_ENABLE_HDFS`

Both will require setting `HADOOP_HOME`. For autotools configure, these environment variable settings may be necessary:

- add `$HADOOP_HOME/lib/native` and `$JAVA_HOME/jre/lib/amd64/server` to `LD_LIBRARY_PATH`.
- export `CLASSPATH=`$HADOOP_HOME/bin/hadoop classpath --glob``

Installation and Use

The following video contains information on installing and using the HDF5 HDFS VFD.

Click on the arrow and select "YouTube" at the bottom.

An HDFS instance is required to verify the proper functioning of the VFD. Fortunately, HDFS can be deployed in standalone or pseudo-distributed mode in non-clustered environments for testing and development. The HDFS VFD bundle includes a set of tests and the scripts to set up such transient HDFS deployments.