This document lists the steps to be followed to build Hadoop 3 with native libraries and Intel-ISAL library. Both are optional features in Hadoop by default.

Native Hadoop

Hadoop has native implementations of certain components for performance reasons and for non-availability of Java implementations. These components are available in a single, dynamically-linked native library called the native hadoop library. The library is named libhadoop.so in GNU/Linus platforms. The native hadoop library is written in ANSI C and is built using the GNU autotools-chain (autoconf, autoheader, automake, autoscan, libtool).

1. Install native libraries required for Hadoop on all instances (both master and slave nodes):

```
sudo apt -y install build-essential autoconf automake libtool
cmake zlib1g-dev pkg-config libssl-dev libsasl2-dev
```

2. Compile Hadoop source with the native flag on

```
sudo mvn package -Pdist,native -DskipTests -Dtar
```

3. After successful build, the newly built native libraries of the project are located at ~/hadoop-dist/target/Hadoop-3.0.0/lib/native

Intel Intelligent Storage Acceleration library (ISAL)

1. ISAL needs to be configured on all nodes (both master and slaves)
2. Download ISAL from the GitHub page https://github.com/intel/isa-l
3. Follow instructions as given on the page. I got initial error that ISAL requires YASM 1.2.0 or higher.
4. From the website:

http://manpages.ubuntu.com/manpages/trusty/man1/yasm.1.html
install latest version from this page (downloaded deb and ran dpkg).

5. While running ./autogen.sh for configuring ISAL,
   error: possibly undefined macro: AC_PROG_LD
   This is because some essential build tools were missing.
6. Fixed with the following:

```
sudo apt install autogen autoconf libtool
```

7. 'make' for building ISAL is not available by default, install it using

```
sudo apt install make
```

8. Now Hadoop source can be built with ISAL support using the following command:

```
sudo mvn package -Pnative -Pdist -Dtar -DskipTests -Drequire.isal
-Disal.lib=/usr/lib -Dbundle.isal
```

* Use -Drequire.isal to fail the build if libisal.so is not found.
* Use -Disal.lib to specify a nonstandard location for the libisal library files (mine got installed in the location /usr/lib).
* Use -Dbundle.isal to copy the contents of the isal.lib directory into the final tar file. This option requires that -Disal.lib is also given, and it ignores the -Disal.prefix option. If -Disal.lib isn't given, the bundling and building will fail.

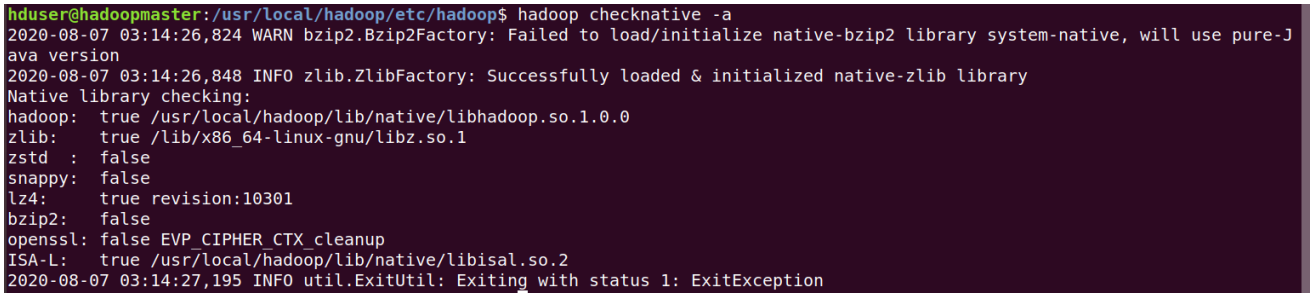
Important step

After enabling native library and ISA-L please add the following code to the project configuration file at $HADOOP_HOME/etc/hadoop/hadoop-env.sh:

```
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_COMMON_LIB_NA-
TIVE_DIR"
```

This is for the name node to recognize the libraries.
We can check the status of native libraries by issuing the hdfs command
`hadoop checknative`

Below is the screenshot:



It shows that native libraries and ISA-L are enabled.

References:

https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-common/NativeLibraries.html
https://hadoop.apache.org/docs/r3.0.0/hadoop-project-dist/hadoop-hdfs/HDFSErasureCoding.html
https://github.com/intel/isa-l