Documentation

- ERP
  - Resources
- Big Data Components
  - Apache Bigtop
    - Resources
  - ODPi
    - Big Data Core Components
      - Apache Hadoop
      - ELK - ElasticSearch, Logstash and Kibana
      - Apache Sqoop
      - Apache Arrow
    - Big Data Operations
      - Apache Ambari
      - Apache Zookeeper
      - Apache Oozie
      - Apache Falcon
      - Ganglia
    - Big Data Streaming Tools
      - Apache Spark
      - Apache Flink
      - Apache Beam
      - Apache Tez
      - Apache Flume
      - Apache Storm
      - Apache Tachyon
      - Apache Kafka
      - Apache NiFi
      - Apache MiNiFi
    - Big Data Data warehousing and Database Tools
      - Apache Hive
      - Apache HBase
      - Apache Cassandra
      - Postgres
      - Memcached
      - MySQL
      - Redis
      - Apache Drill
    - Big Data Data Governance and Security
      - Apache Ranger
      - Apache Knox
      - Apache Atlas
      - Apache Sentry
    - Big Data File Formats
      - Apache Parquet
      - Apache Avro
    - Big Data Datascience Notebooks
      - Apache Jupyter
      - Apache Zeppelin
    - Big Data Analytics
      - Apache Kudu
    - Big Data ML - Machine Learning
      - H2O
    - Big Data component dependencies

Tests
- Smoke Tests
- Integration Tests
- ODPi Spec Tests

Benchmarking
- HiBench
- Build and Port
- Machine Learning
- Misc

Blogs/Presentations
- Big Data Roadmap
- Strategic Engineering
- Big Data Epics
- JIRAInProgress

ERP
Big Data Components

1. Apache BigTop
   - Building Hadoop 2.7.2, Spark 2.0, Hive 2.0.1 using Apache BigTop
   - Building BigTop using Docker container
   - BigTop Sandbox with Hadoop, Spark, Hive, HBase for Aarch64
   - BigTop Smoke Tests
   - Resources
     - BigTop 1.2 Components
     - BigTop Docker images
     - Upstream github
     - CI Loop
     - v1.2 Release
     - BigTop AArch64 patches
     - BigTop Project Ideas and Roadmap

2. ODPi
   - Building ODPi BigTop packages
   - Setup, Configure and Install ODPi Hadoop
   - ODPi Hadoop Cluster Setup Guide
   - ODPi BigTop Hadoop Installation, setup and running, ODPi BigTop Hadoop Installation, setup and running

3. Big Data Core Components
   - Apache Hadoop
     - Setup, Configure and Install ODPi Hadoop
     - Building and Running Apache Hadoop
     - Building Hadoop 2.7.2, Spark 2.0, Hive 2.0.1 using Apache BigTop
     - Apache Hadoop Build, Install, Configure and Run Guide
     - Building, Running, Configuring and Profiling Apache Hadoop
     - Apache Hadoop Tuning Notes
     - Apache Hadoop Map Reduce Notes
     - OpenJDK javac Nullpointerexception building Hadoop
     - Patch 1/3 Introduce the HyperCrc32C Checksum class
     - Patch 2/3 libhadoop: CRC: ARM NEON Support
     - Patch 3/3 ModifyCRC to target NEON routine
     - hadoop-lca13
   - ELK - ElasticSearch, Logstash and Kibana
     - Building ELK (ElasticSearch, LogStash and Kibana) on AArch64
     - ELK Setup and Run on AArch64

4. Big Data Operations
   - Apache Ambari
     - Build Apache Ambari on AArch64
     - Build Apache Ambari on AArch64 (Copy)
     - Install, Setup and Configure Apache Ambari
     - JVM Null pointer error in javac when building Apache Ambari
   - Apache Zookeeper
     - Enablement of Apache Zookeeper on AArch64
   - Apache Oozie
   - Apache Falcon
   - Ganglia

5. Big Data Streaming Tools
   - Apache Spark
     - Building Hadoop 2.7.2, Spark 2.0, Hive 2.0.1 using Apache Bigtop
     - Spark 2.0 - Build, configure and Installation steps using Apache Bigtop
     - Profiling Spark on AArch64
     - Building Apache Spark 1.6 on AArch64
     - Building Apache Spark 2.0
     - Spark Dependency Libraries
   - Apache Flink
   - Apache Beam
   - Apache Tez
   - Apache Flume
   - Apache Storm
   - Apache Tachyon
   - Apache Kafka
6. **Big Data Data warehousing and Database Tools**
   - Apache Hive
     - Building Apache Hive
   - Apache HBase
     - Enablement of Apache HBase on AArch64
   - Apache Cassandra
   - PostgreSQL
   - Memcached
   - MySQL
   - Redis
   - Apache Drill

7. **Big Data Data Governance and Security**
   - Apache Ranger
   - Apache Knox
   - Apache Atlas
   - Apache Sentry

8. **Big Data File Formats**
   - Apache Parquet
   - Apache Avro

9. **Big Data Datascience Notebooks**
   - Apache Jupyter
   - Apache Zeppelin

10. **Big Data Analytics**
    - Apache Kudu

11. **Big Data ML - Machine Learning**
    - H20
      - H20 Install and Run Guide
      - H20 Scaling Study
      - Sparkling Water - Spark + H20

12. **Big Data component dependencies**
    - NETLIB-JAVA AArch64 Natives Support

**Tests**

- **Smoke Tests**
  - BigTop smoke tests
- **Integration Tests**
- **ODPi Spec Tests**

**Benchmarking**

- TeraSort
- Building, Running, Configuring and Profiling Apache Hadoop
- Spark Bench
- TPC-H
- TPCxHS
- Apache Bench
- BigBench
- Building and Running HiBench on AArch64 Platform
- HiveTestBench
- HiBench
  - Spark profiling on Arm64

**Build and Port**

- Build Apache Ambari on AArch64
- HBase Enablement on AArch64
- Apache Flink on AArch64
- Zookeeper Enablement on AArch64
- NETLIB-JAVA AArch64 Natives Support
- Apache Ambari Install, Setup and Configuration
Machine Learning

- Building and Installing Tensorflow on AArch64

Misc

- Onboarding info - Welcome to Team
- Big Data team work summary v1.0
- CRC32 vs Non-CRC32 Study
- Configuring archiva with Tomcat

Bigtop

Blogs/Presentations

State of Big Data on Aarch64 - Apache Bigtop

Big Data benchmarking

Himalayan Odyssey

Big Data Roadmap

Strategic Engineering

Big Data and OpenJDK Strategic Engineering - 2018

Big Data and OpenJDK Strategic Engineering - 2017

Big Data Epics

<table>
<thead>
<tr>
<th>key</th>
<th>summary</th>
<th>type</th>
<th>created</th>
<th>updated</th>
<th>due</th>
<th>assignee</th>
<th>reporter</th>
<th>priority</th>
<th>status</th>
<th>resolution</th>
</tr>
</thead>
</table>

⚠️ Jira project doesn't exist or you don't have permission to view it.

View these issues in Jira

JIRA Inprogress

<table>
<thead>
<tr>
<th>key</th>
<th>summary</th>
<th>type</th>
<th>created</th>
<th>updated</th>
<th>due</th>
<th>assignee</th>
<th>reporter</th>
<th>priority</th>
<th>status</th>
<th>resolution</th>
</tr>
</thead>
</table>
⚠ Jira project doesn't exist or you don't have permission to view it.

View these issues in Jira