Powered By Bigtop

This page aims to listing all the organizations and individuals who are using Apache Bigtop either as a basis for their own distributions or as-is. Please add yourself and tell us a bit about your environment, use cases and your cluster size. If your system is powered-by Apache Bigtop you can use the official logo attached to this page.

In alphabetical order:

- **Cloudera's Distribution Including Apache Hadoop (CDH) 4** CDH 4 is Cloudera's 100% open source Hadoop distribution based on Apache Bigtop. Comprised of Apache Hadoop and a number of other leading open source projects, CDH delivers the core elements of Hadoop – scalable storage and distributed computing – as well as all of the necessary enterprise capabilities such as security, high availability, and integration with a broad range of hardware and software solutions. In addition, Cloudera offers its enterprise customers a family of product and services that complement the open-source Apache Hadoop platform. These include comprehensive training sessions, architectural services and technical support for Hadoop clusters in development or in production. We serve a wide range of customers including retail, government, financial service, healthcare, life sciences, digital media, advertising, networking and telephony enterprises.

- At **EMC/Greenplum**, we have been using BigTop extensively as a build framework for our 1000-node Analytics Workbench Cluster.

- The **Juju Charms for Hadoop, HBase, Hive and Zookeeper** and the associated packages for Ubuntu are a derivation of Apache Bigtop. Juju is a service orchestration tool which allows you to deploy and scale these Bigtop components (and a large number of other services) on Ubuntu Server 12.04 on public and private clouds and on bare metal. Juju also has a local development mode to allow Charm developers to quickly and easily develop Juju Charms.

- **Magna Tempus Group** provides ready to use, well integrated open-source stack for intensive and high-performance in-memory data analysis based on such widely accepted technologies as BigTop, Hadoop, HBase, Hive and many others. This is The standard distribution of Hadoop and analytic components that is the same as the official Apache releases. We’ll keep up with the stream of updates going into open-source projects, so your software will always be tested and up-to-date.

- **Trend Micro** uses Bigtop as the basis for our internal custom distribution of Hadoop, which starts with Bigtop but then pulls features from different upstream versions and includes ASF licensed non-core contributions as our platform needs dictate. We use Bigtop's packaging infrastructure for build automation and Bigtop's integration tests as part of our QA processes. We plan to contribute as much as possible back to the community.

- At **Uniting Data** we deliver on a promise of fast, green Big data analytics leveraging the potential of the multicore clusters and providing tight integration with HPC solutions. Uniting Data Platform is a 100% open source Hadoop distribution based on the Apache Bigtop. We also offer our enterprise customers a family of services including technical training, architectural services and technical support.

- At **Red Hat**, our work on the **GlusterFS Hadoop** project leverages Apache BigTop’s easily consumed big data distribution and smoke tests. Different than other hadoop distributions which may bias towards a particular architecture and implementation: the transparent materials used to build the BigTop product allow us to validate, install, deploy, and test cutting edge hadoop related technology, and modify it in ways that satisfy our inner curiosities - in a manner that is synergistic with community oriented innovation.

- **WANdisco** is a committed member & sponsor of the Apache Software community and has active committers on several project including Apache Hadoop. WANdisco provides:
  - **WANdisco Distro (WDD)**: a fully tested, 100% open source production-ready platform powered by Apache Hadoop 2.
  - **Apache Hadoop Support**: support for Apache Hadoop for both end users and technology vendors.
  - **WANdisco Non-Stop NameNode**: WANdisco's patented replication technology turns the NameNode into an active-active shared-nothing cluster that delivers optimum performance, scalability and availability on a 24-by-7 basis without any downtime or data loss.
  - **WANdisco Hadoop Console**: a comprehensive wizard-driven management dashboard that enables any Linux administrator to fully deploy, manage and monitor a Hadoop big data implementation without HBase or HDFS expertise.
  - **WANdisco S3-HDFS**: Supports migration from Amazon to private in-house clouds.