The Apache Software Foundation Announces Apache™ Hadoop™ 2

Wednesday 16 October, 2013

Foundation of next-generation Open Source Big Data Cloud computing platform runs multiple applications simultaneously to enable users to quickly and efficiently leverage data in multiple ways at supercomputing speed.

Forest Hill, MD –16 October 2013– The Apache Software Foundation (ASF), the all-volunteer developers, stewards, and incubators of nearly 150 Open Source projects and initiatives, today announced Apache™ Hadoop™ 2, the latest version of the Open Source software framework for reliable, scalable, distributed computing.

A foundation of Cloud computing and at the epicenter of "big data" solutions, Apache Hadoop enables data-intensive distributed applications to work with thousands of nodes and exabytes of data. Hadoop enables organizations to more efficiently and cost-effectively store, process, manage and analyze the growing volumes of data being created and collected every day. Apache Hadoop connects thousands of servers to process and analyze data at supercomputing speed.

The project's latest release marks a major milestone more than four years in the making, and has achieved the level of stability and enterprise-readiness to earn the General Availability designation.

"With the release of stable Hadoop 2, the community celebrates not only an iteration of the software, but an inflection point in the project's development. We believe this platform is capable of supporting new applications and research in large-scale, commodity computing," said Apache Hadoop Vice President Chris Douglas. "The Apache Software Foundation creates the conditions for innovative, community-driven technology like Hadoop to evolve. When that process converges, the result is inspiring."

"Hadoop 2 marks a major evolution of the open source project that has been built collectively by passionate and dedicated developers and committers in the Apache community who are committed to bringing greater usability and stability to the data platform," said Arun C. Murthy, release manager of Apache Hadoop 2 and Founder of Hortonworks Inc. "It has been an honor and pleasure to work with the community and a personal thrill to see our four years of work on YARN finally coming to fruition in the GA of Hadoop 2. Hadoop is truly becoming a cornerstone of the modern data architecture by enabling organizations to leverage the value of all their data, including capturing net-new data types, to drive innovative new services and applications."

"What started out a few years ago as a scalable batch processing system for Java programmers has now emerged as the kernel of the operating system for big data," said original Hadoop creator and ASF Board member Doug Cutting. "Over a dozen Apache projects integrate with Hadoop, with ten more in the Apache Incubator poised to soon join their ranks."

Dubbed a "Swiss army knife of the 21st century" and named "Innovation of the Year" by the 2011 Media Guardian Innovation Awards, Apache Hadoop is widely deployed at enterprise organizations around the globe, including industry leaders from across the Internet and social networking landscape such as Amazon Web Services, AOL, Apple, eBay, Facebook, foursquare, HP, LinkedIn, Netflix, The New York Times, Rackspace, and Twitter. Other technology leaders such as Microsoft, IBM, Teradata, SAP have integrated Apache Hadoop into their offerings. Yahoo!, an early pioneer, hosts the world’s largest known Hadoop production environment to date, spanning more than 35,000 nodes.

Under the Hood
Apache Hadoop 2 reflects intensive community- development, production experience, extensive testing, and feedback from hundreds of knowledgeable users, data scientists and systems engineers, bringing a highly stable, enterprise-ready release of the fastest-growing big data platform.

New in Hadoop 2 is the addition of YARN that sits on top of HDFS and serves as a large-scale, distributed operating system for big data applications, enabling multiple applications to run simultaneously for more efficient support of data throughout its entire lifecycle. The culmination of so many other releases in the Hadoop 2.x line, the most current release --2.2.0-- is the first stable release in the 2.x line. Features include support support for:

Related Sectors:

Computing & Telecoms ::

Related Keywords:

About The Apache Software Foundation :: Apache :: Software :: AFS ::

Scan Me:

- Apache Hadoop YARN, a cornerstone of next generation Apache Hadoop, for running both data-processing applications (e.g. Apache Hadoop MapReduce, Apache Storm etc.) and services (e.g. Apache HBase)
- High Availability for Apache Hadoop HDFS
- Federation for Apache Hadoop HDFS for significant scale compared to Apache Hadoop 1.x.
- Binary Compatibility for existing Apache Hadoop MapReduce applications built for Apache Hadoop 1.x.
- Support for Microsoft Windows.
- Snapshots for data in Apache Hadoop HDFS.
- NFS-v3 Access for Apache Hadoop HDFS.

"The community has stepped up to the challenge of making Hadoop enterprise-ready, hardening the filesystem, providing high availability, adding critical security capabilities,and delivering integrations to enable consolidation of any kind or amount of enterprise data," said Aaron Myers, member of the Apache Hadoop Project Management Committee and Engineer at Cloudera.

"Today, with the announcement of Hadoop 2 and YARN, we've taken another step. Beyond the basic multitenancy customers have enjoyed for the past year, enabling them to mix batch, interactive and real-time workloads, they now have the ability to do so from within a stable foundational part of the Hadoop ecosystem. It's a testament to the community's work that now every distribution of Apache Hadoop will enjoy these benefits, ensuring that customers can deliver the applications they need, on a single Hadoop platform."

"It has been an honor and pleasure to work with the community and a personal thrill to see our four years of work on YARN finally coming to fruition in the GA of Hadoop 2," added Murthy. "Apache Hadoop is truly becoming a cornerstone of the modern data architecture by enabling organizations to leverage the value of all their data, including capturing net-new data types, to drive innovative new services and applications."

"A large portion of the credit for this success is due to Apache's open-source model, which has permitted a wide range of users and vendors to productively collaborate on a platform shared by all," added Cutting.

Availability and Oversight

As with all Apache products, Apache Hadoop software is released under the Apache License v2.0, and is overseen by a self-selected team of active contributors to the project. A Project Management Committee (PMC) guides the Project's day-to-day operations, including community development and product releases. Apache Hadoop release notes, source code, documentation, and related resources are available at http://hadoop.apache.org/.

About The Apache Software Foundation (ASF)

Established in 1999, the all-volunteer Foundation oversees nearly one hundred fifty leading Open Source projects, including Apache HTTP Server --the world's most popular Web server software. Through the ASF's meritocratic process known as "The Apache Way," more than 400 individual Members and 3,500 Committers successfully collaborate to develop freely available enterprise-grade software, benefiting millions of users worldwide: thousands of software solutions are distributed under the Apache License; and the community actively participates in ASF mailing lists, mentoring initiatives, and ApacheCon, the Foundation's official user conference, trainings, and expo. The ASF is a US 501(c)(3) charitable organization, funded by individual donations and corporate sponsors including AMD, Basis Technology, Budget Direct, Citrix, Cloudera, Comcast, Facebook, Go Daddy, Google, HP, Hortonworks, Huawei, IBM, InMotion Hosting, Matt Mullenweg, Microsoft, PSW
Group, Pivotal, WANdisco, and Yahoo!. For more information, visit http://www.apache.org/ or follow @TheASF on Twitter.

"Apache", "Apache Hadoop", "Hadoop", and "ApacheCon" are trademarks of The Apache Software Foundation. All other brands and trademarks are the property of their respective owners.

# # #

NOTE: you are receiving this message because you are subscribed to the announce@apache.org distribution list. To unsubscribe, send email from the recipient account to announce-unsubscribe@apache.org with the word "Unsubscribe" in the subject line.

Company Contact:

The Apache Software Foundation

T. 16179218656
E. press@apache.org
W. https://www.apache.org

View Online

Newsroom: Visit our Newsroom for all the latest stories:
https://www.apache.pressat.co.uk