****

**Shopzilla Implements a Cloudera Enterprise Data Hub to Enhance its EDW and Capture Unparalleled Retail Insights**

***Processes 15,000 Feeds and 100 Million Products Daily from Retailers in Just Hours or Minutes with Cloudera***

**PALO ALTO, Calif. – August 5, 2014** – [Cloudera](http://www.cloudera.com/content/cloudera/en/home.html), the leader in enterprise analytic data management powered by [Apache Hadoop](http://www.cloudera.com/content/cloudera/en/why-cloudera/hadoop-and-big-data.html)™, today announced that Shopzilla, Inc., a leading source for connecting buyers and sellers online, has deployed a Cloudera enterprise data hub to complement its existing Oracle enterprise data warehouse (EDW). In this hybrid Big Data environment, Shopzilla now has unlimited capacity to process and deliver new insights on millions of pageviews and ten billion ad requests daily, reaching over 100 million unique visitors and gaining valuable insights in hours or minutes instead of days.

Shopzilla recently announced that it combined three marketing-centric business units that will operate under the Connexity brand name. Connexity will uniquely combine consumer insights and media buying within the same programmatic platform, helping marketers to learn more about their customers, discover valuable audiences, and activate new consumers at scale. Data from Shopzilla's global portfolio of retail websites connects more than 40 million shoppers with over 100 million products from tens of thousands of retailers and that crucial data is now an offering via Connexity. It is important that data was processed as rapidly and efficiently as possible in order to keep up with growing customer engagement. With its 500-terabyte EDW growing by five terabytes a day, Shopzilla’s existing legacy data warehouse had outgrown its capacity, impacting the company’s ability to provide business analytics in a timely and effective manner.

“Our legacy system delivers great performance for analytics and reporting, but didn’t have the bandwidth for the intensive data transformations we needed—it would take hours to process 100 million products per day,” said Paramjit Singh, director of data for Shopzilla. “We needed enormous processing capabilities, scalability, full redundancy, and extensive storage—at a cost-effective price. Our Cloudera platform provides all that and more, while complementing our current data warehouse system. We were able to reduce latency from days to hours and soon minutes.”

“Cloudera provides an exploration environment for our data scientists that reveals tremendous insights, which would be virtually impossible to obtain otherwise,” explained Singh. “We’re able to answer complex questions on multi-structured data, such as how a user is behaving on a particular site and what ads would be most effective, as well as execute other sophisticated data mining queries. It improves Shopzilla’s ability to provide relevant results to users—a core tenet of our business. Many of the things we do as a business would not be possible without this platform running alongside our Oracle data warehouse.”

This improved processing performance also benefits Shopzilla’s search engine marketing (SEM) activities, allowing the company to score and bid on ten million keywords each day. Reaching over 100 million users, Shopzilla is able to collect billions of data points to create some of the most targeted and rich shopping-intent data available.

“By 2017, US online retail sales will total $434.2 billion. In a data-driven industry such as online retail, which is experiencing such explosive growth, providing profound and timely insights to both shoppers and retailers is key in boosting marketing ROI,” said Alan Saldich, vice president of marketing at Cloudera. “Connecting social and transactional data provides businesses with a 360-degree view of customer behaviors, interactions, interests, and activities in a way that was just not possible before.”

Shopzilla augmented its Oracle EDW with a multi-tenant Cloudera Enterprise system to create a hybrid environment. While Hadoop is the primary engine for data processing and analytics, aggregated data is stored in the EDW using Apache Sqoop for reporting on the back end. Users can access Cloudera Enterprise directly using Apache Pig and Apache Hive, and Shopzilla plans to upgrade to Cloudera Impala and Apache Spark in the near future.

A Cloudera-powered enterprise data hub delivers the most secure, managed, governed, and open data management platform to give customers a choice over legacy data management for storing, accessing, and analyzing any amount and any kind of data in one centralized repository. Cloudera has all of the key attributes necessary for customers to make data the true focal point of any business.

To learn more, view the Shopzilla success story: [http://www.cloudera.com](http://www.cloudera.com/" \t "_blank)/content/cloudera/en/our-customers/Shopzilla.html

**About Shopzilla**

Shopzilla, Inc. is a leading source for connecting buyers and sellers online. Reaching a global audience of over 50 million shoppers each month through both its destination websites and affiliate network, Shopzilla connects shoppers with over 100 million products from tens of thousands of retailers a month. Shopzilla, Inc. manages a premier portfolio of online shopping brands in the US and Europe, consisting of [Bizrate](http://www.bizrate.com), [Beso](http://www.beso.com), [Shopzilla](http://www.shopzilla.com), [Retrevo](http://www.retrevo.com), [TaDa](http://www.tada.com/), [PrixMoinsCher](http://prixmoinscher.com), and [SparDeinGeld](http://www.spardeingeld.com), as well as B2B businesses including [Connexity](http://www.connexity.com), a consumer insights and audience activation platform, and the Shopzilla Publisher Program.

**About Cloudera**

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for Big Data, an enterprise data hub built on Apache Hadoop™. Cloudera offers enterprises one place to store, process and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data. Only Cloudera offers everything needed on a journey to an enterprise data hub, including software for business critical data challenges such as storage, access, management, analysis, security and search. As the leading educator of Hadoop professionals, Cloudera has trained over 22,000 individuals worldwide. Over 1,000 partners and a seasoned professional services team help deliver greater time to value. Finally, only Cloudera provides proactive and predictive support to run an enterprise data hub with confidence. Leading organizations in every industry plus top public sector organizations globally run Cloudera in production.

[www.cloudera.com](http://www.cloudera.com)

**Connect with Cloudera**

Read our blog:<http://www.cloudera.com/blog/>

Follow us on Twitter:<http://twitter.com/cloudera>

Visit us on Facebook:<http://www.facebook.com/cloudera>

*Cloudera, Cloudera Platform for Big Data, Cloudera Enterprise Basic Edition, Cloudera Enterprise Flex Edition, Cloudera Enterprise Data Hub Edition* and *CDH* are trademarks or registered trademarks of Cloudera in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

[###](http://www.facebook.com/cloudera)#

**Press Contacts**

**North America**

Deborah Mullan

Bhava Communications for Cloudera

cloudera@bhavacom.com

+1-510-984-6122

**Europe**

Richard Botley

Ketchum for Cloudera

LON-Cloudera@ketchum.com

+44 (0) 20 7611 3788

***Source: (1)*** *Retail Ecommerce Set to Keep a Strong Pace Through 2017: eMarketer 2013.*

*http://www.emarketer.com/Article/Retail-Ecommerce-Set-Keep-Strong-Pace-Through-2017/1009836*