

CLUDERA AND OUTERTHOUGHT

The Problem:

Apache Hadoop allows us to store and process huge amounts of data, while Apache HBase provides an interactivity and data management layer on top of Hadoop - based on the Google BigTable model.

However, this current technology lacks in key aspects of usability for the enterprise developer. On the analytics and visualization side, there still is a serious lack of developer-friendly offerings that enable enterprise developers to become fully productive with Apache HBase to attack real-world problems.

We wanted to make Apache HBase easier to use for software engineers working on enterprise projects, providing the benefit of new technology with an application model and APIs they can quickly start working with, and that's how Lily was born: it allows organizations to Smartly manage Big Data, Easily.

The Solution:

Lily is a jumpstart for every content-centric CDH/Hadoop/HBase project: it offers developer-centric access, development, and test tools to increase development momentum. It combines the interactivity and flexibility of Apache HBase with enterprise indexing and search services Apache Solr and Elastic Search. Lily is specifically useful in retail, e-commerce, news and media projects, as a data scaling and aggregation platform.

Its flexible schema supports progressive insights in data models, with Lily being used to store and make accessible large data streams coming from heterogeneous sources.

Lily goes beyond storage, access, index and search: it unobtrusively tracks your users and how they behave around your data, generating real-time insights on behavioral patterns, and automatically creating a data store for rapid recommendations.

All this functionality is specifically made available with the enterprise developer in mind, with unobtrusive low-barrier APIs and no user interface technology leaking its abstractions into their architecture. Lily is designed to be a back-end data platform: something to rely on, something to build upon, but not something imposing a full technology stack. Lily is created with ease of integration in mind, doing the heavy lifting so that enterprise devs can focus on the money-making features.



Outerthought

Industries:

- Online
- Retail
- Media and News

Website: <http://outerthought.org/>

Company Overview

Outerthought is an industrious software engineering company specialized in Java-based open source development with a focus on internet-scale content applications, infrastructure-level services and great developer convenience. The company develops Lily, the first content repository for the Age of Data - Smart Data, at Scale, made Easy, and Daisy, a sophisticated collaborative content management system. Outerthought is based in Zwijnaarde, Belgium.

Product Overview

Lily is smart data at scale, made easy. It combines leading-edge Apache HBase with enterprise index and search technology in a developer-friendly open source package. It also offers subscription-based enterprise deployment support tools and services. Lily is ideally suited as a data management, aggregation and scale-out platform, combining the benefits of large-scale distributed data management with flexible indexing and search based on Apache. Its rich schema capabilities and easy Java and REST APIs help enterprise developers to become rapidly productive with best-of-breed technology for the age of data.

Features

- Deep, robust integration between Apache HBase and Solr with automatic index updating, maintenance, sharing, text extraction and index regeneration (using Mapreduce);
- Consistent secondary indexes for alternative, high-speed access to your data;
- A flexible, rich data model layered on top of the HBase/BigTable model, with a variety of smart data types (e.g. Links and BLOB) to support real-world software projects;
- Support for schema and data versioning;
- Enterprise-scale deployment and management tools and admin applications - tightly integrated with CDH;
- Tools for building Lily applications with the enterprise developer in mind; and
- Java and REST API.

Benefits

- Lily's ease of use actively reduces project lead time;
- Lily's fully distributed architecture allows you to utilize your entire infrastructure to cope with large volumes of data and usage;
- Lily's tight integration between Apache HBase and Solr brings flexible big data index and search in reach of every enterprise developer;
- Lily's built-in usage statistics gathering offers access to real-time insights on how your data is being used; and
- Lily comes with built-in recommendations, offering cross- and up-selling, personalized and targeted content recommendations, and creating tighter bounds between your users and your data.

About Cloudera's Distribution Including Apache Hadoop

Available for free download at www.cloudera.com/downloads, CDH delivers a streamlined path for putting Apache Hadoop to work solving business problems in production. Ideal for enterprises seeking a stable, tested Hadoop solution without proprietary vendor lock-in, CDH is the bridge between the insights of organizations using Hadoop in production and the continuous stream of innovations from the Apache community.

About Cloudera

Cloudera is the leading provider of Apache Hadoop-based software and services and works with customers in financial services, web, telecommunications, government and other industries. The company's products, Cloudera Enterprise and Cloudera's Distribution including Apache Hadoop, help organizations profit from all of their information. Cloudera's Distribution including Apache Hadoop is the most comprehensive Apache Hadoop-based platform in the industry. Cloudera Enterprise is the most cost-effective way to perform large-scale data storage and analysis and includes the tools, platform and support necessary to use Apache Hadoop in a production environment. For more on Cloudera, please visit www.cloudera.com.

About Outerthought

Outerthought is a leading technology supplier of Lily, a content and data platform for large-scale media, news and retail customers. With a long history of active open source development and contributions, we deliver a developer-friendly, easy-to-use application that wraps Apache Hadoop/HBase and enterprise search technology in a well-integrated, productive development environment with enterprise management tools. Located in Belgium with customers worldwide, we work together with implementation partners to help our customers benefit from leading-edge Big Data technology.