Support   Developers

Contact Us   Downloads

Search 🔍

COMMUNITY | DOCUMENTATION | DOWNLOADS | TRAINING | BLOGS

Sneak Preview: HBaseCon 2015 Development & Internals Track

by **Justin Kestelyn (@kestelyn)**   April 17, 2015   **no comments**

This year's HBaseCon Development & Internals track covers new features in HBase 1.0, what's to come in 2.0, best practices for tuning, and more.

In this post, I'll give you a window into the **HBaseCon 2015's** (May 7 in San Francisco) Development & Internals track.

Thanks, **Program Committee**!

- **"Meet HBase 1.0"**
  Enis Söztutar (Hortonworks) & Solomon Duskis (Google)

  HBase 1.0 is the new stable major release, and the start of "semantic versioned" releases. We will cover new features, changes in behavior and requirements, source/binary and wire compatibility details, and upgrading. We'll also dive deep into the new standardized client API in 1.0, which establishes a separation of concerns, encapsulates what is needed from how it's delivered, and guarantees future compatibility while freeing the implementation to evolve.

- **"HBase 2.0 and Beyond: Panel"**
  Matteo Bertozzi (Cloudera), Sean Busbey (Cloudera), Jingcheng Du (Intel), Lars Hofhansl (Salesforce.com), Jon Hsieh (Cloudera), Enis Söztutar (Hortonworks) & Jimmy Xiang (Cloudera) (Pinterest)

  Now that you've seen Base 1.0, what's ahead in HBase 2.0, and beyond—and why? Find out from this panel of people who have designed and/or are working on 2.0 features.

- **"HBase Performance Tuning"**
  Lars Hofhansl (Salesforce.com)

  At Salesforce, we have deployed many thousands of HBase/HDFS servers, and learned a lot about tuning during this process. This talk will walk you through the many relevant HBase, HDFS, Apache ZooKeeper, Java/GC, and Operating System configuration options and provides guidelines about which options to use in what situation, and how they relate to each other.

- **"Solving HBase Performance Problems with Apache HTrace"**
  Abraham Elmahrek & Colin McCabe (Cloudera)

  HTrace is a new Apache incubator project which makes it much easier to diagnose and detect performance problems in HBase. It provides a unified view of the performance of requests, following them from their origin in the HBase client, through the HBase region servers, and finally into HDFS. System administrators can use a central web interface to query and view aggregate performance information for the whole cluster. This talk will cover the motivations for creating HTrace, its design, and some examples of how HTrace can help diagnose real-world HBase problems.

- **"HBase and Spark"**
  Ted Malaska (Cloudera)

  In this session, learn how to build an Apache Spark or Spark Streaming application that can interact with HBase. In addition, you'll walk through how to implement common, real-world batch design patterns to optimize for performance and scale.

- **"Events @ Box: Using HBase as a Message Queue"**
  David Mackenzie (Box)

  Box's /events API powers our desktop sync experience and provides users with a realtime, guaranteed-delivery event stream. To do that, we use HBase to store and serve a separate message queue for each of 30+ million users. Learn how we implemented queue semantics, were able to replicate our queues between clusters to enable transparent client failover, and why we chose to build a queueing system on top of HBase.

- **"State of HBase Docs and How to Contribute"**
  Misty Stanley-Jones (Cloudera)

  In this session, learn about the move to Asciidoc in HBase docs, some of the other notable changes lately, and things we've done to make it easier for you to contribute to the docs.

- **"Reusable Data Access Patterns with CDAP Datasets"**
  Gary Helmling (Cask Data)

  In this talk, you'll learn about Datasets, part of the open source Cask Data Application Platform (CDAP), which provide reusable implementations of common data access patterns. We will also look at how Datasets provide a set of common services that extend the capabilities of HBase: global transactions for multi-row or multi-table updates, read-less increments for write-optimized counters, and support for combined batch and real-time access.

Getting interested? Wait until you see the Ecosystem sneak preview next week!

Archives by Month

HBASECON

Thank you to **our sponsors** – **Bloomberg, Cask Data, Hortonworks, MapR, Pepperdata, Salesforce.com (Gold); Facebook, Google (Silver); Apache Software Foundation, The Hive Think Tank (Community); O'Reilly Media (Media)** — without which HBaseCon would be impossible!

Filed under:

Community

Events

HBase

No Responses

Leave a comment

Name REQUIRED

Email REQUIRED
(WILL NOT BE PUBLISHED)

Website

Comment

Leave Comment

Prove you're human! *

$5 \times two =$ _____

Products

Cloudera Enterprise
Cloudera Express
Cloudera Manager
CDH
All Downloads
Professional Services
Training

Solutions

Enterprise Solutions
Partner Solutions
Industry Solutions

Partners

Resource Library
Support

About

Hadoop & Big Data
Management Team
Board
Events
Press Center
Careers
Contact Us
Subscription Center

English

Follow us:

Share:

Cloudera, Inc.
1001 Page Mill Road Bldg 2
Palo Alto, CA 94304

www.cloudera.com
US: 1-888-789-1488
Intl: 1-650-362-0488