Presentations

- Hive Meetups
  - January 2016 Hive User Group Meetup
  - November 2015 Hive Contributor Meetup
  - April 2015 Hive Contributor Meetup Presentations
  - February 2015 Hive User Meetup Presentation
  - November 2013 Hive Contributors Meetup Presentations
  - June 2013 Hadoop Summit Hive Meetup Presentations
  - February 2013 Hive User Group Meetup
  - June 2012 Hadoop Summit Hive Meetup Presentations
  - November 2011 NYC Hive Meetup Presentations
- Older Hive Presentations
- Related Work

Hive Meetups

January 2016 Hive User Group Meetup

- Monitoring Hive - Szehon Ho
- Data Organization for Performance - Gopal V
- What's new in Hive 2.0 - Sergey Shelukhin
- Hive on Spark: now and future - Xuefu Zhang

November 2015 Hive Contributor Meetup

- Metrics in Hive - Mohit Sabharwal
- Hive at Uber - Mohammad Islam
- Capybara Hive Integration Testing - Alan Gates
- Random Query Gen- Hive Meetup - Szehon Ho
- Lets talk sessionstate - Anthony Hsu
- Lessons learned from enterprise users - Xuefu Zhang
- Hardening Hive - John P
- LLAP: Sub-Second Execution in Hive - Gopal V

April 2015 Hive Contributor Meetup Presentations

- SQLLine - The tale of the twins that everyone loved too much – Julian Hyde
- ResultSet Compression in Apache Hive – Rohil Dholakia, George Chow
- High availability.Rolling upgrade for Hiveserver2 - Vaibhav Gumashita
- The birds and bees talk (Replication) - Sushanth Sowmyan

February 2015 Hive User Meetup Presentation

- Hive Join Optimizations: MR and Spark (Szehon Ho)
- Cascading and Hive (Ryan Desmond)

November 2013 Hive Contributors Meetup Presentations

- Using Dynamic Compilation with Hive (Edward Capriolo)
- Let There Be Tez: Current Status and Demo (Gunther Hagleitner)
- Hive Authorization Improvement Proposal - HIVE-5837 (Thejas Nair)
- Insert, update, and delete in Hive with full ACID support - HIVE-5317 (Owen O'Malley)
- Apache Sentry (Brock Noland)
- Cost Based Optimizer Design - HIVE-5775 (John Pullokaran)
- Join syntax improvement - HIVE-5555 (Ashutosh Chauhan)

June 2013 Hadoop Summit Hive Meetup Presentations

- Hive Correlation Optimizer (Yin Hua)

February 2013 Hive User Group Meetup

- Hive at KIXEYE and in the Gaming Industry (Aaron Sun)
June 2012 Hadoop Summit Hive Meetup Presentations

- What’s News in Apache Hive 0.9.0 (Ashutosh Chauhan, Hortonworks)
- HiveServer2 Project (Carl Steinbach, Cloudera)
- Column Statistics in Hive (Shreepadma Venugopalan, Cloudera)
- HCatalog: Extending Hive’s Metadata to Pig and MR (Alan Gates, Hortonworks)
- Continuous Aggregations with Hive (Viral Bajaria, Hulu)

November 2011 NYC Hive Meetup Presentations

- Breaking First Normal Form (Edward Capriolo, Media 6 Degrees)
- HAWK: Performance Monitoring for Hive (JunHo Cho, NexR)
- RHive: Integrating R and Hive (JunHo Cho, NexR)
- Replacing the Hive Execution Layer with Pervasive Turbo Rush (Jim Falgout, Pervasive Software)
- Using JDBC with Hive (Bennie Schut, eBuddy)

Older Hive Presentations

- High Volume Updates in Hive from 2012 Hadoop Summit (Owen O’Malley)
- Join Strategies in Hive from the 2011 Hadoop Summit (Liyn Tang, Namit Jain)
- Replacing an Oracle DB/DW with Hadoop/Hive from the 2011 Hadoop Summit Hive Contributor Meeting (JunHo Cho, NexR)
- RCFile Paper at ICDE 2011, Hannover, Germany (Yongqiang He (Facebook), Rubao Lee (OSU), Yin Huai (OSU), Zheng Shao (Facebook), Namit Jain (Facebook), Xiaodong Zhang (OSU), Zhwei Xu (ICT, CAS))
- Join Optimization in Hive (Liyn Tang, Facebook)
- Hive Presentation at ApacheCon NA 2010 (John Sichi, Facebook)
- Using Hadoop and Hive to Optimize Travel Search (Jonathan Seidman and Ramesh Venkataramaiah, Orbitz)
- HBase Meetup HUG10 - Hive/HBase Integration (John Sichi, Facebook)
- Hive User Group Presentation - Hive Quick Start Tutorial (Carl Steinbach, Cloudera)
- Hive User Group Presentation - New Features and APIs from Facebook (Ning Zhang, Yongqiang He, Namit Jain, John Sichi, Paul Yang, Zheng Shao, Facebook)
- Hive User Group Presentation from Netflix (Eva Tse and Jerome Boulon, Netflix)
- Hive Paper at QCon Nov 2009 (Ashish Thusoo and Namit Jain, Facebook)
- Hive Training – Motivations and Real World Use Cases, (Ning Zhang, Facebook, at Cloudera’s training session)
- User Defined Table Generating Functions, (Paul Yang, Facebook)
- Hive Anatomy – System & Pseudo-code level Architecture Review, (Ning Zhang, internal presentations, Facebook)
- Rethinking the Data Warehouse with Hadoop and Hive (Ashish Thusoo, Facebook at Hadoop World NYC 2009)
- Hive User Group Meeting August 2009, (Facebook)
- Hive Object Model, (Zheng Shao, Facebook)
- Hive: VLDB 2009, Lyon, France (Facebook)
- Hive: Hadoop Summit 2009, Santa Clara, CA, USA (Namit Jain, Zheng Shao, Facebook)
- Data Warehousing & Analytics on Hadoop, Percon Conference, Santa Clara, CA, USA (Ashish Thusoo, Prasad Chakka, Facebook)
- Large Scale Data Processing using commodity SW/HW, IIT-Delhi CS Dept., (Joydeep Sen Sarma, Facebook)
- An Introduction to Hive, Jeff Hammerbacher, Facebook
- Hive: Data Warehousing Analytics on Hadoop, UC Berkeley, (Joydeep Sarma, Namit Jain, Zheng Shao, Facebook)
- Hive: Data Warehousing with Hadoop, NYC Hadoop User Meetup (Jeff Hammerbacher, Cloudera)
- Facebook and Open Source, UIUC, (Zheng Shao, Facebook)
- Hive ApacheCon 2008, New Orleans, LA (Ashish Thusoo, Facebook)

Related Work

- Processing Theta-Joins using MapReduce (A. Okcan, M. Riedewald)
- Optimizing Joins in a Map-Reduce Environment (F. Afrati, J. Ullman)
- Efficient Parallel Set-Similarity Joins Using MapReduce (R. Vernica, M. Carey, C. Li)
- HadoopDB: An Architectural Hybrid of MapReduce and DBMS Technologies for Analytical Workloads (A. Abouzeid, K. Bajda-Pawlikowski, D. Abadi, A. Silberschatz, A. Raspin)
- Tenzing: A SQL Implementation On The MapReduce Framework
- Building a High-Level Dataflow System on top of Map-Reduce: The Pig Experience
- YSmart: Yet Another SQL-to-MapReduce Translator (R. Lee, T. Luo, Y. Huai, F. Wang, Y. He, X. Zhang)
- Query Optimization Using Column Statistics in Hive (A. Gruenheid, E. Omiecinski, L. Mark)