

Kafka papers and presentations

Papers

- [Streams and Tables: Two Sides of the Same Coin](#), Matthias J. Sax, Guozhang Wang, Matthias Weidlich, Johann-Christoph Freytag
- [Building a Replicated Logging System with Apache Kafka](#), Guozhang Wang, Joel Koshy, Sriram Subramanian, Kartik Paramasivam, Mammad Zadeh, Neha Narkhede, Jun Rao, Jay Kreps, Joe Stein
- [Liquid: Unifying Nearline and Offline Big Data Integration](#), Raul Castro Fernandez, Peter Pietzuch, Jay Kreps, Neha Narkhede, Jun Rao, Joel Koshy, Dong Lin, Chris Riccomini, Guozhang Wang
- [Kafka, Samza and the Unix Philosophy of Distributed Data](#), Martin Klepmann, Jay Kreps
- [Kafka: A Distributed Messaging System for Log Processing](#), Jay Kreps, Neha Narkhede, Jun Rao from LinkedIn, at NetDB workshop 2011
- [Building LinkedIn's Real-time Activity Data Pipeline](#), Ken Goodhope, Joel Koshy, Jay Kreps, Neha Narkhede, Richard Park, Jun Rao, Victor Yang Ye
- [The Log: What Every Software Engineer Should Know About Real-time Data's Unifying Abstraction](#), Jay Kreps

Presentations

- [Stream Processing made Easy with Apache Kafka](#), Guozhang Wang, Hadoop Summit SF, 2016 ([video](#), [slides](#))
- [Commander: Decoupled, Immutable REST APIs with Kafka Streams](#), Bobby Calderwood of Capital One, Abstractions Con 2016, Aug 2016
- [New consumer internals: rebalancing, rebalancing...](#) -- Jason Gustafson (Confluent) & Onur Karaman (LinkedIn), August 2016
- [Application Development and Data in the Emerging World of Stream Processing](#), Neha Narkhede, Berlin Buzzwords, June 2016 ([video](#), [slides](#))
- [Introducing Kafka Streams, the new stream processing engine of Apache Kafka](#), Michael G. Noll, Berlin Buzzwords, June 2016 ([video](#), [slides](#))
- [Building Stream Infrastructure across Multiple Data Centers with Apache Kafka](#), Guozhang Wang, Strata San Jose 2016 ([slides](#))
- [Tuning Kafka for low latency guaranteed messaging](#), Jiangjie (Becket) Qin (LinkedIn), June 2016
- [Kafka quotas](#), Aditya Auradkar (LinkedIn), November 2015
- [Kafka Security](#), Harsha Chintalapani & Parth Brahmabhatt (Hortonworks), November 2015
- [Building a Replicated Logging System with Apache Kafka](#), Guozhang Wang, VLDB 2015, September ([slides](#))
- [Developing with the Go client for Apache Kafka](#), Joe Stein, January 2015
- [Apache Kafka 0.8 Basic Training](#), Michael G. Noll, July 2014
- [Apache Kafka at LinkedIn](#), Guozhang Wang, BDTc 2016, December ([slides](#))
- [I Heart Logs: Apache Kafka and Real-time Data Integration](#), Jay Kreps, June 2014
- [Developing real-time data pipelines with Apache Kafka](#), Joe Stein, DataDayTexas 01/2014
- [Infrastructure at Scale: Apache Kafka, Apache Storm & Elasticsearch](#), Jim Nisbet, Philip O'Toole, AWS re:invent 2013
- [Real-time streaming and data pipelines with Apache Kafka](#) ([slides](#)), Joe Stein, NYC Storm Meetup 12/2013
- [Building a realtime data pipeline apache Kafka at LinkedIn](#), Joel Koshy, Hadoop Summit 2013
- [Intra-cluster replication in Apache Kafka](#), Jun Rao from LinkedIn, ApacheCon 2013
- [Kafka: A Distributed pub-sub messaging system](#), Neha Narkhede from LinkedIn, ApacheCon 2011
- [Kafka: A Distributed Messaging System for Log Processing](#), Jun Rao from LinkedIn, at NetDB workshop 2011
- [Kafka: LinkedIn's open source distributed pub-sub system](#), Neha Narkhede from LinkedIn, SVForum Software Architecture & Platform, SIG, July 2011
- [Kafka - Bringing Reliable Stream Processing To a Cold, Dark World](#) ([slides](#)), Jakob Homan from LinkedIn, at Yahoo! London HUG
- [Kafka - Bringing Reliable Stream Processing To a Cold, Dark World](#) ([video](#)), Jakob Homan from LinkedIn, at Berlin Buzzwords 2011
- [Kafka: high-throughput, persistent, multi-reader streams](#) Jay Kreps from LinkedIn, at HUG January 2011
- [Event Stream Processing with Kafka](#) Tim Lossen from Wooga, Ruby User Group Berlin, February 2012
- [Kafka: Demo of features using producer and consumer shells](#) Edward Capriolo m6d.com
- [Event Stream Processing with Kafka](#) ([video](#)) Tim Lossen from Wooga, Berlin Buzzwords, June 2012
- [Dashboard Mania](#) Tim Lossen from Wooga, O'Reilly Velocity Conf, June 2012
- [Kafka: The Great Log File In The Sky](#) John Pignata, GroupMe, August 2012
- [Kafka + Storm = Realtime data at GumGum](#) ([slides](#)), Vaibhav Puranik from GumGum, LA Hadoop Users Group, August 2012
- [Building LinkedIn's Real-time Data Pipeline](#) ([slides](#)) Jay Kreps, Airbnb tech talk, San Francisco August, 2012
- [Scaling near-real-time analytics with Kafka and HBase](#) Nate Putnam and Dave Revell, OSCON 2012
- [Data Models and Consumer Idioms Using Apache Kafka](#) Erik Onnen, Surge 2012
- [Kafka: A Little Introduction](#)
- [Introduction to Kafka](#) ([slides](#)) ([video](#)) Chris Curtin, Atlanta Java Users Group, March 2013
- [Introduction and Overview to Kafka](#) David Arthur, TriHUG July 2013

Blog Posts

- [Kafka Streams - Not Looking at Facebook](#), by Timothy Renner, Aug 2016
 - Describes an (imaginary) app that monitors user behavior and notifies users when they are not looking at Facebook
- [Processing Tweets with Kafka Streams](#), August 2016
 - Describes implementing a stream processing application in Scala that processes tweets (windowed wordcount). Uses docker to containerize the app.
- [Applying Kafka Streams for internal message delivery pipeline](#), blog post by [LINE Corp.](#), August 2016
 - Describes the use of Kafka Streams (and Kafka) in production at LINE for processing 1+ million msg/s. As of 2016 LINE has 220+ million active monthly users.
- [USwitch data pipeline](#)
- [InfoChimps Data Platform: Next Gen Real-time Streaming with Storm-Kafka Integration](#)
- [Storm and Kafka example application at Twitter](#)
- [Big Data Right Now: Five Trendy Open Source Technologies](#)
- [A Big Data Trifecta: Storm, Kafka and Cassandra](#)

User group meetings:

- 1st Kafka user group meeting at LinkedIn, Jun. 14, 2012. [video \(part 1\)](#), [video \(part 2\)](#)
- 2nd Kafka user group meeting at LinkedIn, Jun 27 2013. [video](#)
- 3rd Kafka user group meeting at LinkedIn, June 3, 2014, [video](#)
- Kafka meetup at LinkedIn, February, 2015:
 - [Supporting large messages in Kafka](#), Jiangjie (Becket) Qin (LinkedIn)
 - [SignalFX's optimized Kafka consumer](#), Rajiv Kurian (SignalFX)
 - [Bridging the Gap: Connecting AWS and Kafka](#) -- RYanne Dolan & Jason Li (LinkedIn)
 - [Loose Coupling with the Jitney Project](#) -- Alexis Midon & Krishna Puttaswamy (Airbnb)
- Kafka meetup at LinkedIn, November, 2015
 - [Quotas](#) -- Aditya Auradkar (LinkedIn)
 - [Apache Kafka Security](#) -- Harsha Chintalapani & Parth Brahmbhatt (Hortonworks)
 - [Takeaways from Live Site Issues](#) - Joel Koshy (LinkedIn)