

Papers & Talks

- [Paper](#)
- [Talks](#)
- [Samza Meetups](#)
- [Blog Posts on Samza](#)

Paper

- August 2017 - [Samza: Stateful Scalable Stream Processing at LinkedIn](#)
- January 2017 - [The RADStack: Open Source Lambda Architecture for Interactive Analytics](#)
- June 27 - July 2, 2016, Effective Multi-stream Joining in Apache Samza Framework in 5th IEEE International Congress on Big Data, San Francisco, USA
- May 2016, [SamzaSQL: Scalable Fast Data Management with Streaming SQL](#) presented at [IEEE International Parallel and Distributed Processing Symposium Workshops \(IPDPSW\)](#)
- November 1, 2015 - [A Memory Capacity Model for High Performing Data-Iterating Applications in Samza Framework](#)(BigData Quality Workshop)(Slides)

Talks

- June 2018 - [Stream Processing at LinkedIn with Apache Samza \(Bangalore Kafka Group Meetup\)](#) (Slides)
- Nov 2017 - [Unified Stream Processing at Scale with Apache Samza \(BigDataSpain 2017\)](#) (Slides)
- Sept 21, 2017 - [Unified Batch & Stream Processing with Apache Samza \(Dataworks Summit Sydney 2017\)](#) (Slides)
- August 28, 2017 - [Samza: Stateful Scalable Stream Processing at LinkedIn \(VLDB 2017\)](#) (Slides)
- May 16, 2017 - [What it takes to process a trillion events a day? Case studies in scaling stream processing at LinkedIn - Jagadish Venkatraman \(ApacheCon Big Data '17\)](#) (Slides)
- Nov 9, 2016 - [Scaling up Near real-time Analytics \(QCon'16 SF Slides\)](#)
- June 30, 2016 - [Lambdaless Stream Processing at Scale in LinkedIn \(Hadoop Summit 2016 Video\)](#)
- May 9, 2016 - [Will it Scale? The Secrets behind Scaling Stream Processing Applications \(ApacheCon Big Data NA 2016\)](#) (Slides) (No Video)
- Oct 15, 2015 - [Essential ingredients for real time stream processing @Scale \(BigData 2015 @Spain\)](#) (video)
- September 28, 2015 - [Integrating fully-managed data streaming services with Apache Samza \(ApacheCon BigData EU\)](#)(Slides)
- June 12, 2015 - [Going Realtime with Kafka and Samza at Improve Digital \(GeekOut\)](#)
- February 3, 2015 - [Scalable real-time data processing with Apache Samza \(Jfokus\)](#)
- January 24, 2015 - [Making sense of stream processing \(/dev/winter\)](#)
- November 20, 2014 - [Moving Towards a Streaming Architecture \(Strata EU\)](#)
- November 18, 2014 - [Scalable stream processing with Apache Samza and Apache Kafka \(ApacheCon EU\)](#) (Slides)
- November 3, 2014 - [Samza in LinkedIn: How LinkedIn Processes Billions of Events Everyday in Real-time \(QCon SF\)](#)
- October 28, 2014 - [Staying agile in the face of the data deluge \(Span Conference\)](#) (Slides)
- October 16, 2014 - [Building real-time data products at LinkedIn with Apache Samza \(Strata/Hadoop World\)](#) (Slides) (Video)
- September 18, 2014 - [Turning the database inside out with Apache Samza \(Strangeloop\)](#) (Slides) (Video)
- September 9, 2014 - [Samza: Reliable Stream Processing atop Apache Kafka and Hadoop YARN](#)
- May 27, 2014 - [Samza at LinkedIn: Taking Stream Processing to the Next Level \(Video\)](#) (Slides)
- February 18, 2014 - [Apache Samza: Reliable Stream Processing atop Apache Kafka and Hadoop YARN \(Video\)](#)
- January 14, 2014 - [Apache Samza: Reliable Stream Processing atop Apache Kafka and Hadoop YARN](#)
- November 13, 2013 - [Samza: Real-time Stream Processing at LinkedIn \(Slides\)](#) (QCon SF 2013)
- November 5, 2013 - [Samza: Real-time Stream Processing at LinkedIn \(Slides\)](#) (HUG at LinkedIn)
- October 22, 2013 - [Apache Samza: Reliable Stream Processing atop Apache Kafka and Hadoop YARN \(Video\)](#) (Slides) (London HUG)
- September 27, 2013 - [Introduction to Samza \(Video\)](#) (YARN Meetup)

Samza Meetups

- July 2018:
 - [Beam me up Samza: How we built a Samza Runner for Apache Beam](#) - Xinyu Liu (LinkedIn)
 - [Concourse - Near real time notifications platform at LinkedIn](#) - Ajith Muralidharan & Vivek Nelamangala (LinkedIn)
- March 2018:
 - [Conquering the Lambda architecture in LinkedIn metrics platform with Apache Calcite and Apache Samza](#) - Khai Tran (LinkedIn) [SLIDES]
 - [Building Venice with Apache Kafka & Samza](#) - Gaojie Liu (LinkedIn)
- December 2017:
 - [Stream processing using Samza-SQL@LinkedIn -- Srinivasulu Punuru](#) (LinkedIn) [SLIDES]
 - [Streaming data pipelines@Slack -- Ananth Packkildurai](#) (Slack) [SLIDES]
- September 2017:
 - [Samza at Redfin: Using Streaming to Help Home Buyers and Sellers](#) - Brian Hanks (Redfin)
 - [Real-time Indexing of LinkedIn's Economic Graph](#) - Almog Gavra (LinkedIn)
- May 2017:
 - [Streaming Data Pipelines with Brooklin](#) - Samarth Shetty (LinkedIn)
 - [Managed or stand alone, streaming or batch; Unified processing with the Samza Fluent API](#) - Yi Pan (LinkedIn) [SLIDES]
- Feb 2017:
 - [Asynchronous Processing and Multithreading in Apache Samza](#) - Xinyu Liu (LinkedIn)
 - [Batching to Streaming Analytics at Optimizely](#) - Vignesh Sukumar, Mike Davis, Hao Xia (Optimizely)

- November 2016:
 - [Apache Samza: Past, Present, and Future](#) - Kartik Paramasivam (LinkedIn)
- August 2016:
 - [Nearline Topic Tagging of News Articles on Samza](#) -- Eric Huang (LinkedIn)
 - [How to convert a legacy Hadoop Map/Reduce ETL systems to Samza Streaming](#) -- Louis Calisi (TripAdvisor)
- June 2016:
 - [Air Traffic Controller: Using Samza to Manage Communications with Members](#) – Cameron Lee & Shubhanhu Nagar (LinkedIn) [[Slides](#)]
 - [Scalable Complex Event Processing on Samza](#) -- Shuyi Chen (Uber)
- February 2016:
 - [StatServer-Samza: Near Real-time Analytics](#) -- Tomy Tsai (LinkedIn)
- October 2015:
 - [Benchmarking Apache Samza](#) - Tao Feng (LinkedIn)
 - [New Features in Samza 0.10.0](#) - Navina Ramesh (LinkedIn)
 - [Netflix Keystone Pipeline](#) - Monal Daxini (Netflix)
- July 2015:
 - [Athena - Stream Processing Platform](#) -- Chinmay Soman (Uber)
 - [Harvesting the Power of Samza in LinkedIn's Feed](#) -- Mohamed Mahmoud (LinkedIn)
- May 2015:
 - [System Latency Diagnosis for Microservices with Samza and Druid](#) - Roger Hoover (Quintiply)
 - [Indexing Time Series Streams with Samza and Druid](#) — Gian Merlino ([MetaMarkets.com](#))
 - [Clojure with Samza: application architecture...implementation challenges](#) — Harlan Iverson
- February 2015
 - [Optimizing Streaming SQL Queries](#) -- Julian Hyde (Hortonworks)

Blog Posts on Samza

- Aug 22, 2016 - [Stream Processing Hard Problems Part II: Data Access](#)
- June 27, 2016 - [Stream Processing Hard Problems - Part 1 - Killing Lambda](#)
- January 26, 2016 - [Stream processing with Apache Samza - Current and Future](#)
- August 24, 2015 - [Benchmarking Apache Samza: 1.2 million messages per second on a single node](#)
- [Apache Samza, LinkedIn's Framework for Stream Processing](#) ([thenewstack.io](#))