

Real-time Tick Data Processing, Storing and Retrieval System

A key client engaged SAVEN to develop an integrated system to capture the tick data that is disseminating as outbound from various ticker plants, process them real time, store it in hierarchical data format using HDF5 and provide tools for tick analysis using a web GUI visualization and command line utilities. The system also processes for any malformed criteria defined at the backend and flags those malformed data as alerts. The storage of tick data is available from five minutes just before the current real time to the history back depending upon the storage capacity/disk space.

Application Features & Functionalities

- QA support tool to assess the problems in a ticker plant feed processing proactively before delivering their ticker plant releases into production and Operational support tool during the production
- Maintain Historical Tick Servers persistence, accuracy and issue monitoring the downstream,
- Handles Tick data correction, unusual entitlements, category or malformed content other than the ticker plant defined or correct ones.
- Parallel & Multi components processing of tick data - data decoding, data error detection, stats generation, pattern recognition routines on the data
- Persistence of historical data and providing efficient search capabilities in that huge voluminous data. This means storing approximately 350 GB data per day and creating indexes for this data to enable efficient querying
- Researched limitations in RDBMS like MySQL and NoSql like mongo DB with reference to the data insertion & retrieval performance of time series data.
- Researched efficient data storage library, HDF5 and developed specialized indexing algorithms that can take advantage of the properties of the data.
- Implemented Java EE and HTML5 for the web application part for intuitive data query and presentation

Technologies:

- C/C++, HDF5, Perl, Bash scripting, libpcap , kernel module pfring, Java/JavaEE, HTML5 and JQuery
- Ticker plants process- North American Equities, North American Futures, European equities, Fareast, and Energy content like Argus, SNLISO, AMEREX, EOX, USDA, EIA, OPIS, NGX, CAED etc.

Next Goals

- Scalable and handle voluminous data feeds like OPRA, horizontal scalability as a distributed application with Cassandra as distributed data store.

DISCLAIMER : DATA PROVIDED IS FOR DEMO PURPOSES ONLY . IT IS NOT INTENDED FOR TRADING OR DECISION MAKING PURPOSES . ALL DATA IS PAST OR DUMMY DATA , AND NOT REAL TIME . BY ACKNOWLEDGING THESE TERMS , THE VIEWER UNDERSTANDS AND AGREES THAT THE INFORMATION IN THE WIDGET , OR DERIVED FROM IT IS NOT FOR USE AS REAL DATA .

Contact: SRIDHAR CHELIKANI, SAVEN TECHNOLOGIES, WWW.SAVEN.IN
E MAIL: SRIDHAR.CHELIKANI@SAVEN.IN, PHONE: 847-380-3069, FAX: 847-693-4123