Monitoring Anomalies in Experimentation Platform

Deepak Vasthimal & Roopa Penmetsa

Experimentation

- 1500+ experiments

- Process 100s of TBs of data daily over Hadoop.

- 200+ metrics generated daily.

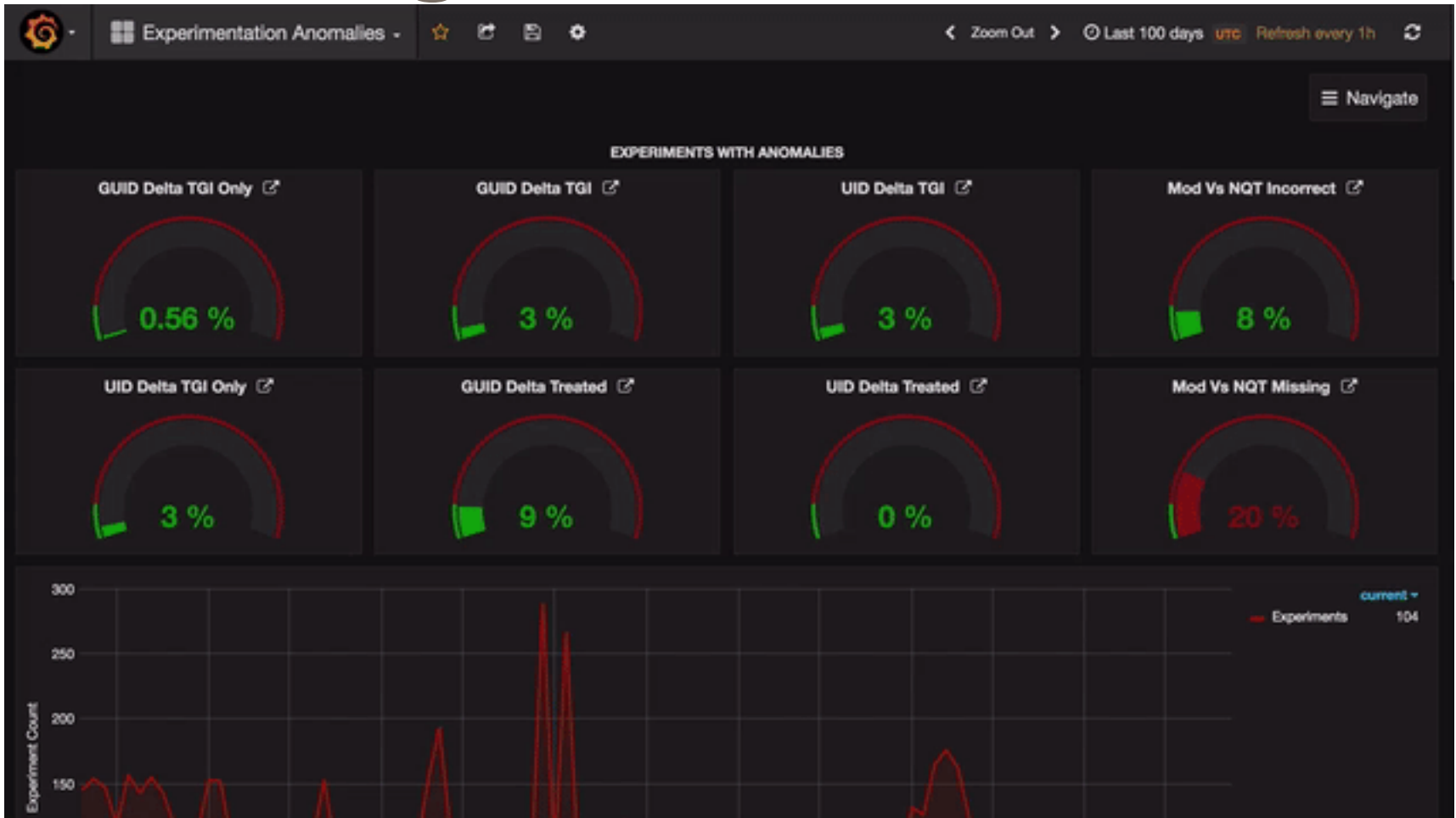- Built using Scala, Scoobi, Hive, Teradata, MicroStrategy, InfluxDB & Grafana.

Anomalies

• Traffic corruption – GUID/UID.

• Tag corruption – Data loss/corrupted during logging & transfer to HDFS.

•GUID Reset – Browser cookie.

•Cache refresh - eBay application servers maintain caches of experiment configurations. A software or hardware glitch
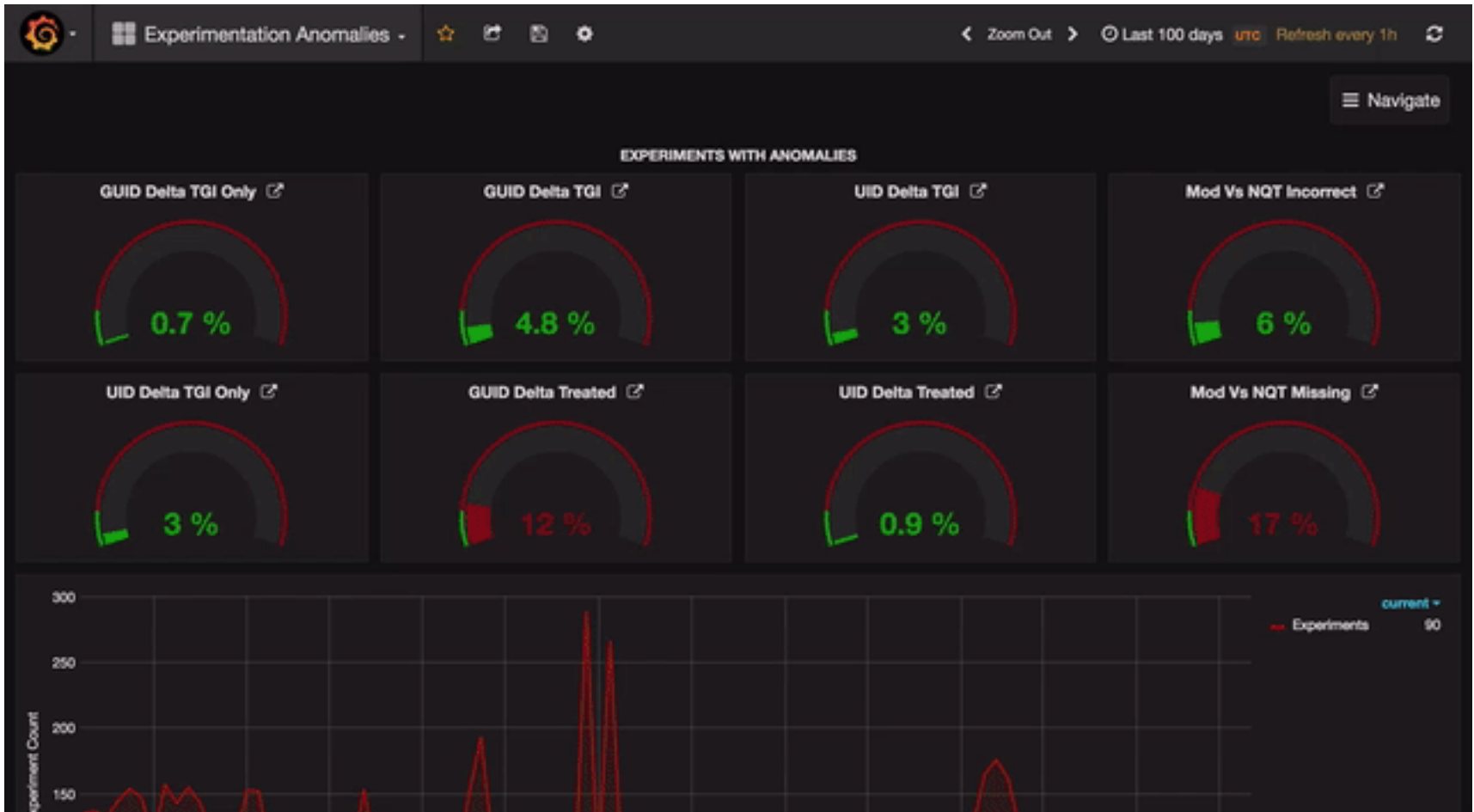
Monitoring Anomalies

• Identify & Categorize anomalies within experiments using Teradata/Hive.

• Store identified anomalies in HDFS and route to InfluxDB (TSDB)
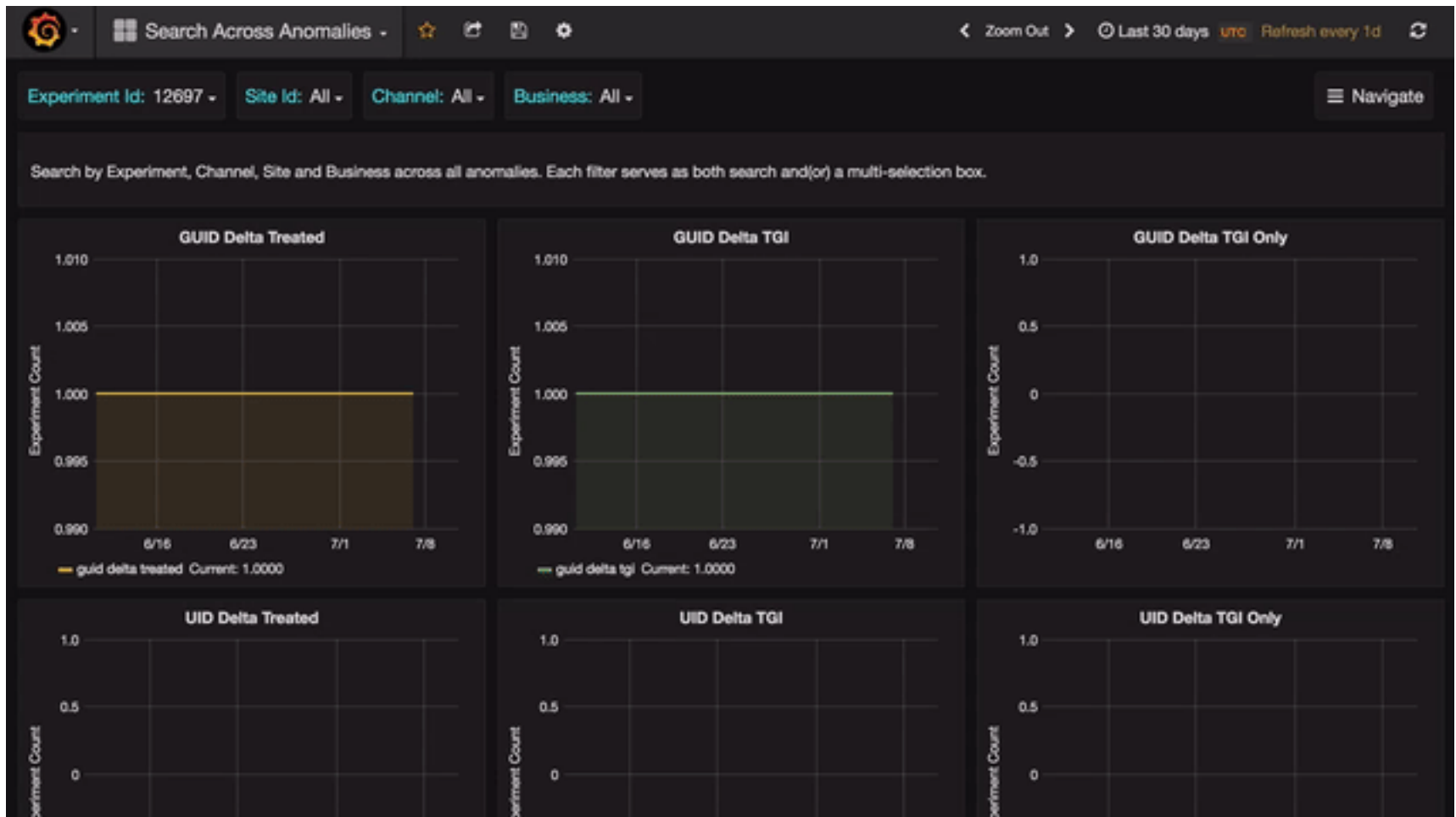
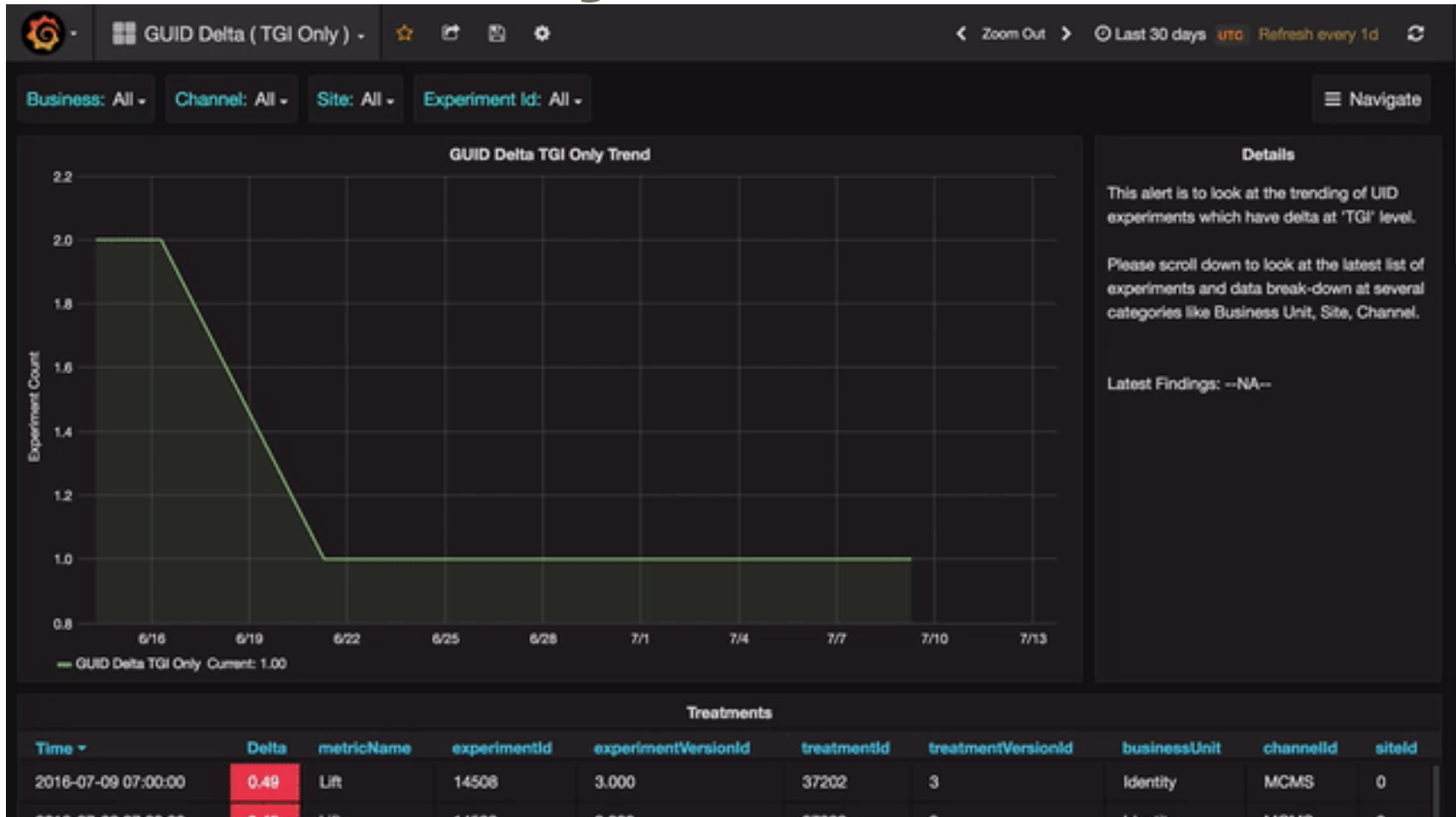• Visualize using Grafana.

Home Page

DrillDown

Search

Share

Built In Query Editor

Questions ?