



*ticketmaster*<sup>®</sup>

# Doing It Live

Machine learning at scale, and in an instant

# Christopher Smith

- VP of Engineering, Data Science
- [chris.smith2@ticketmaster.com](mailto:chris.smith2@ticketmaster.com)
- FanGraph - Single view of the fan



# Visible Impact

Sec 9-H \$132.00 ea  
Row 73, Seats 107-108 [Buy](#)

[Details](#)

[Get Tickets Instantly](#)

Sec 14-H \$133.00 ea  
Row 58, Seats 101-102 [Buy](#)

[Details](#)

[Get Tickets Instantly](#)

Sec 14-H \$133.00 ea  
Row 58, Seats 113-116\* [Buy](#)

[Details](#)

[Get Tickets Instantly](#)

Sec 14-H \$133.00 ea  
Row 59, Seats 105-108\* [Buy](#)

[Details](#)

All resale tickets are verified by Ticketmaster.  
Prices subject to change anytime.  
\*You'll get 2 adjacent seats. Exact seat numbers provided on purchase confirmation page.  
Additional fees are applied at Checkout.

**Be in the know – get a text if more tickets become available.**

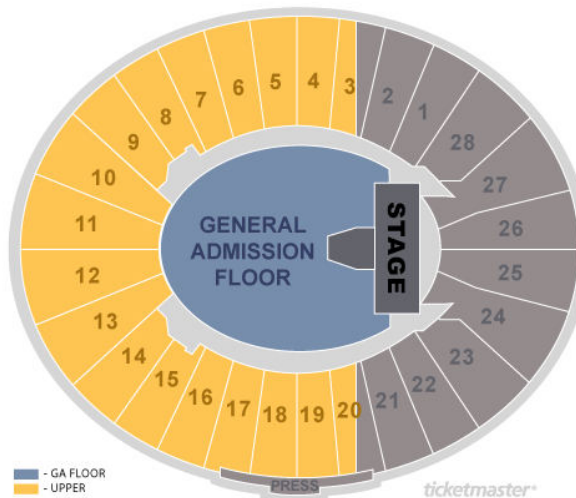
555-555-5555



[I Want In!](#)

Need accessible seats? Click **"Request Accessible Tickets"** at the top right of this map.

You authorize us to use automated technology to send messages to this number.  
Consent not required for purchase. Reply STOP to stop messages. Msg & data rates may apply. [Privacy Policy](#)  
[Terms](#). Help [texthelp@ticketmaster.com](mailto:texthelp@ticketmaster.com)



PLEASE SELECT A SEAT MAP

Metallica

## Rose Bowl

1001 Rose Bowl Drive, Pasadena, CA 91103

[Directions](#) | [Parking](#) | [See All Events at this Venue](#)

Seating charts reflect the general layout for the venue at this time.  
For some events, the layout and specific seat locations may vary without notice.



Sell fun at scale

*ticketmaster*<sup>®</sup>

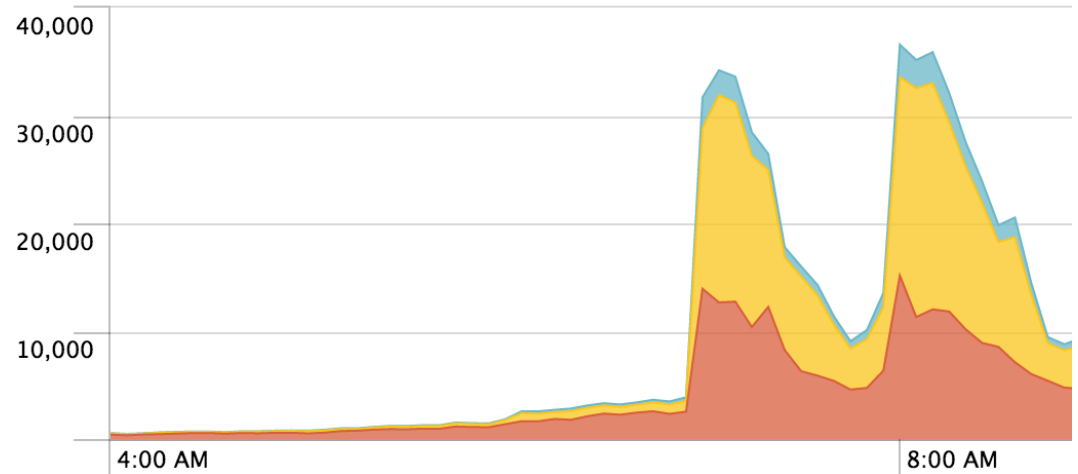


Sell out a stadium  
...in minutes

*ticketmaster*<sup>®</sup>

# Data Science?

- Massive Traffic Spikes
- Real Time Data at Scale
- Separating Fans from Bots

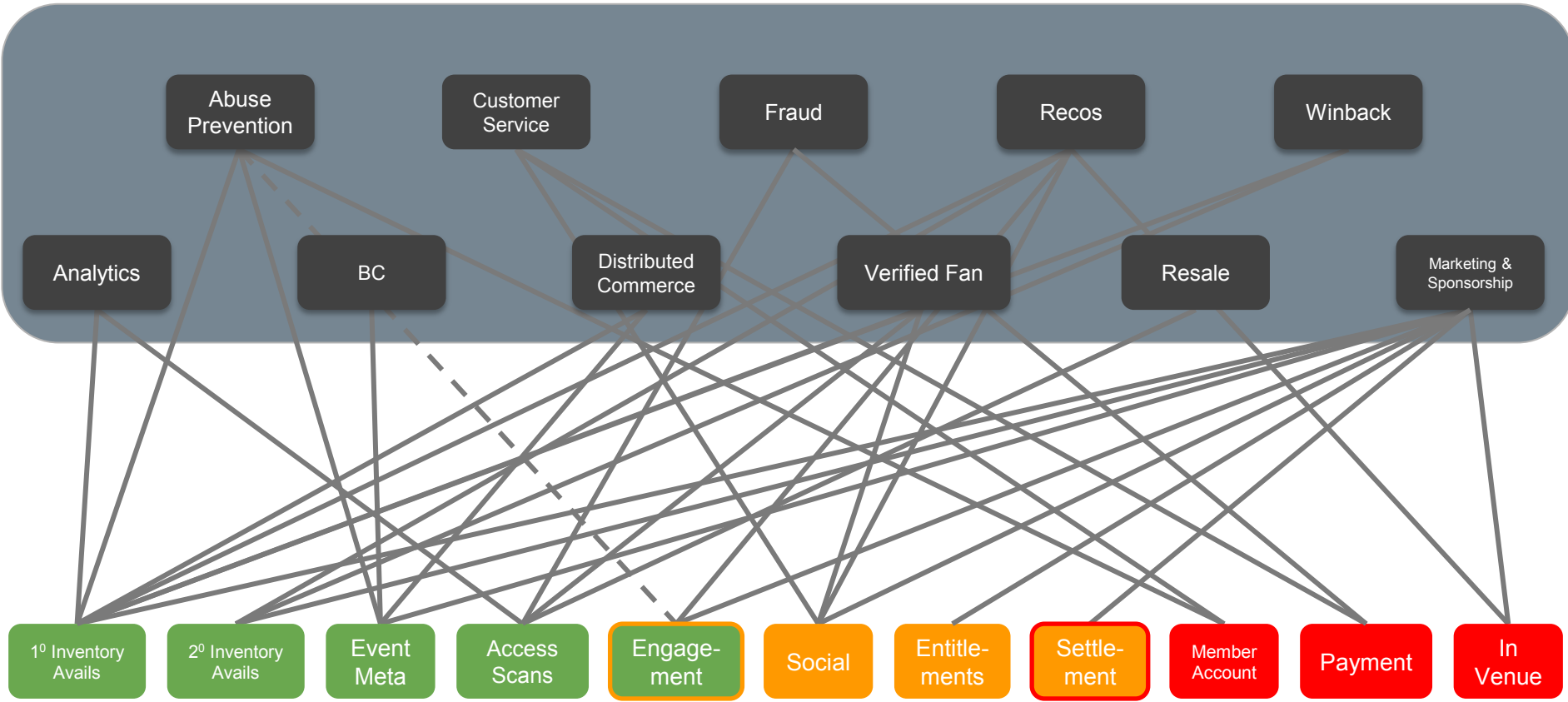




*ticketmaster*<sup>®</sup>

downloaded from pickywallpapers.com

# Data From Everywhere... to Everywhere





# Core Principles

- Source systems have one job: publish mutations/data
- Consumers don't impact producers
- Process data one tuple at a time, in a FP-like fashion
- Data is available everywhere
- Data is always archived
- Organize data per project's needs

# Core Components

- Kafka: reliable, scalable data transit – one cluster per DC
- CKAN: data discovery
- Avro: efficient, extensible serialization w/schema
- Secor: archiving
- Storm/Trident: complex lambda processing
- Vowpal Wabbit: fast, online machine learning
- ElasticSearch: flexible document search

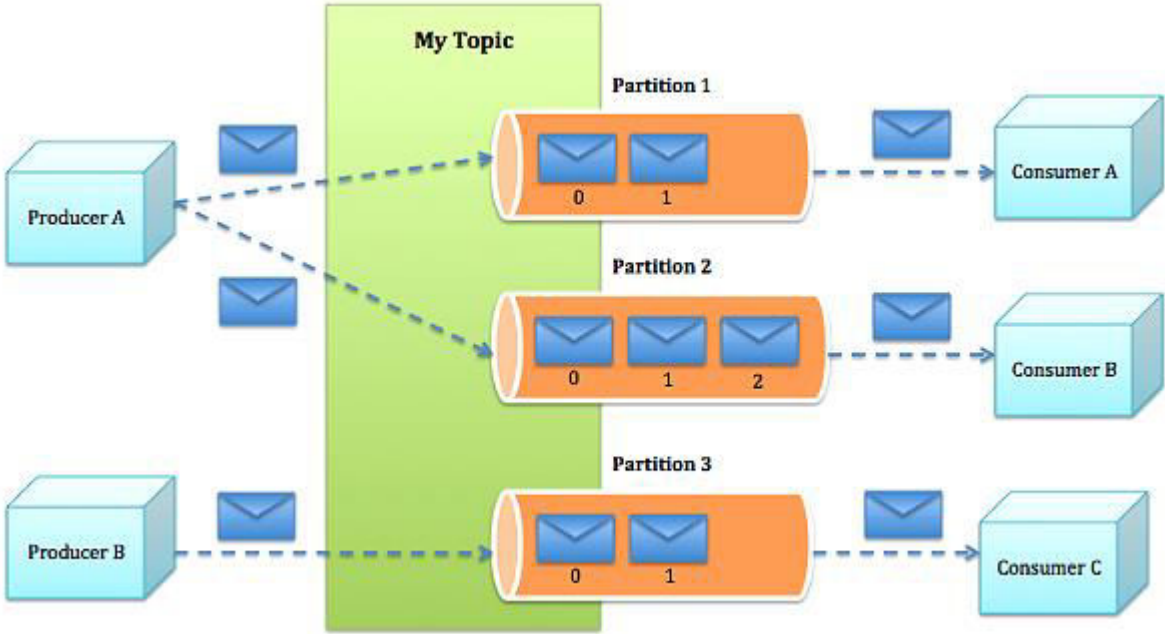
The screenshot shows the CKAN Data Catalog interface. At the top, there's a navigation bar with "DATA CATALOG" and links for "Datasets" and "Organizations". A search bar is present with the text "Search datasets...". Below the search bar, it says "Datasets ordered by Popular".

On the left side, there's a "Filter by location" section with a map of the United States and a "Topics" section with filters for "A-Z" and "1-9".

The main content area displays "195,001 datasets found". Below this, there are three dataset listings:

- Consumer Complaint Database** (536 recent views): Consumer Financial Protection Bureau — These are complaints we've received about financial products and services. Formats: CSV, JSON, XML, API. Location: Federal.
- College Scorecard** (385 recent views): Department of Education — College Scorecards make it easier for students to search for a college that is a good fit for them. They can use the College Scorecard to find out more about a... Formats: HTML, Zipped CSV, CSV, CSV, CSV, CSV. Location: Federal.
- Demographic Statistics By Zip Code** (383 recent views): City of New York — Demographic statistics broken down by zip code. Formats: CSV, RDF, JSON, XML. Location: City.
- Housing Affordability Data System (HADS)** (290 recent views): Department of Housing and Urban Development — The Housing Affordability Data System (HADS) is a set of files derived from the 1985 and later national American Housing Survey (AHS) and the 2002 and later Metro... Format: CSV. Location: Federal.
- ZIP Code Data** (283 recent views): Department of the Treasury — This study provides detailed tabulations of individual income tax return data at the state and ZIP code level. Location: Federal.

# Kafka



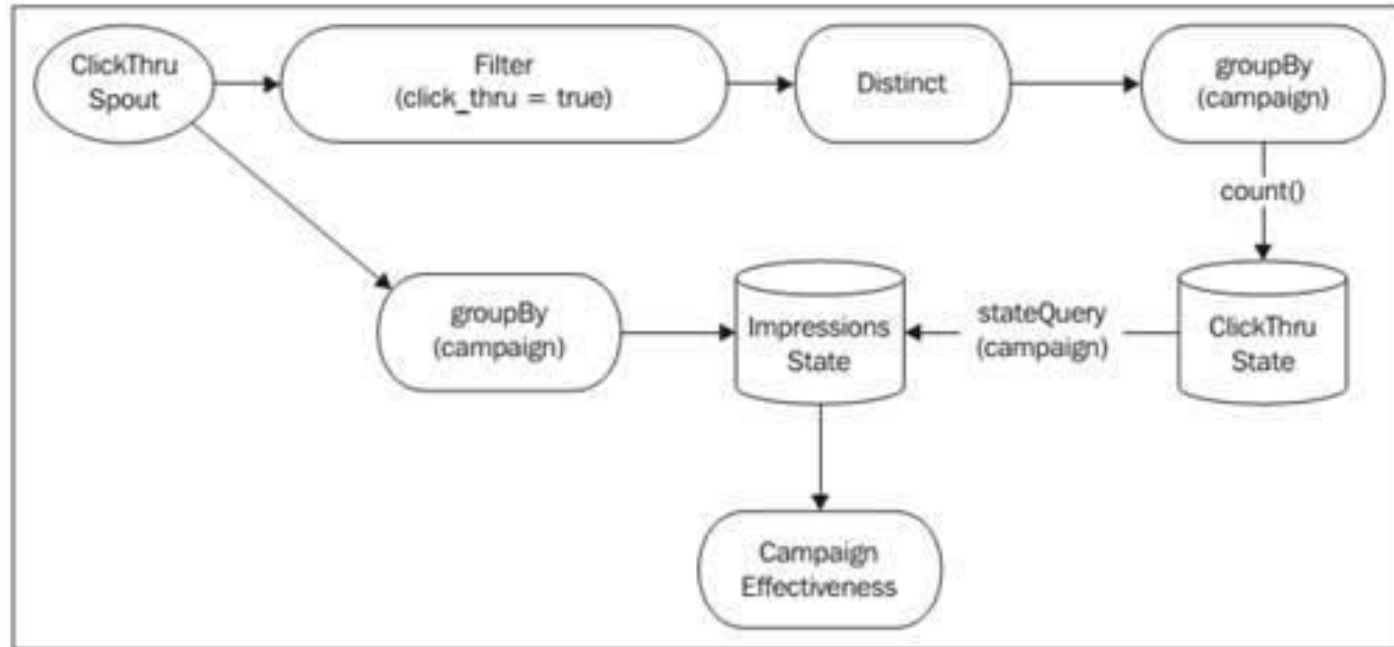
# Avro

```
{
  "namespace": "com.tm.ds.pixel.record",
  "type": "record",
  "name": "TMRecord",
  "fields": [
    { "name": "detectedDuplicate", "type": "boolean" },
    { "name": "detectedCorruption", "type": "boolean" },
    { "name": "firstInSession", "type": "boolean" },
    { "name": "timestamp", "type": "long" },
    { "name": "clientTimestamp", "type": [ "null", "long"], "default": null },
    { "name": "remoteHost", "type": "string" },
    { "name": "referrer", "type": [ "null", "string"], "default": null },
    { "name": "location", "type": [ "null", "string"], "default": null },
    { "name": "viewportPixelWidth", "type": [ "null", "int"], "default": null },
    { "name": "viewportPixelHeight", "type": [ "null", "int"], "default": null },
    { "name": "screenPixelWidth", "type": [ "null", "int"], "default": null },
    { "name": "screenPixelHeight", "type": [ "null", "int"], "default": null },
    { "name": "divoltePartyId", "type": [ "null", "string"], "default": null },
    { "name": "divolteSessionId", "type": [ "null", "string"], "default": null },
    { "name": "pageViewId", "type": [ "null", "string"], "default": null },
    { "name": "eventType", "type": "string", "default": "unknown" },
    { "name": "userAgentString", "type": [ "null", "string"], "default": null },
    { "name": "userAgentName", "type": [ "null", "string"], "default": null },
    { "name": "userAgentFamily", "type": [ "null", "string"], "default": null },
    { "name": "userAgentVendor", "type": [ "null", "string"], "default": null },
    { "name": "userAgentType", "type": [ "null", "string"], "default": null },
    { "name": "userAgentVersion", "type": [ "null", "string"], "default": null },
    { "name": "userAgentDeviceCategory", "type": [ "null", "string"], "default": null },
    { "name": "userAgentOsFamily", "type": [ "null", "string"], "default": null },
    { "name": "userAgentOsVersion", "type": [ "null", "string"], "default": null },
    { "name": "userAgentOsVendor", "type": [ "null", "string"], "default": null },
    { "name": "eventParameters", "type": [ "null", {"type": "map", "values": [ "null", "string"]}], "default": null },
    { "name": "isStagingEnvironment", "type": "boolean", "default": false },
    { "name": "topics", "type": [ "null", {"type": "array", "items": "string"}], "default": null },
    { "name": "sessionId", "type": [ "null", "string"], "default": null },
    { "name": "sessionIdTM", "type": [ "null", "string"], "default": null },
    { "name": "memberId", "type": [ "null", "string"], "default": null },
    { "name": "memberIdPersistent", "type": [ "null", "string"], "default": null },
    { "name": "browserId", "type": [ "null", "string"], "default": null },
    { "name": "destination", "type": [ "null", "string"], "default": null },
    { "name": "pageType", "type": [ "null", "string"], "default": null },
    { "name": "domain", "type": [ "null", "string"], "default": null },
    { "name": "marketId", "type": [ "null", "int"], "default": null },
    { "name": "tmLink", "type": [ "null", "string"], "default": null },
    { "name": "recommendations", "type": [ "null", {"type": "array", "items": {
      "type": "record",
      "name": "Recommendation",
      "fields": [
        { "name": "artistId", "type": [ "null", "int"], "default": null },
        { "name": "slotNumber", "type": [ "null", "int"], "default": null },
        { "name": "strategyName", "type": [ "null", "string"], "default": null }
      ]
    }
    ]
  ]
}
```

# Secor

- Highly scalable
- Aggregate Kafka topic data by time
- Flexible output formats (sequence files, parquet, csv)
- Compression
- Store in s3 (cheap)

# Storm/Trident



# Storm/Trident

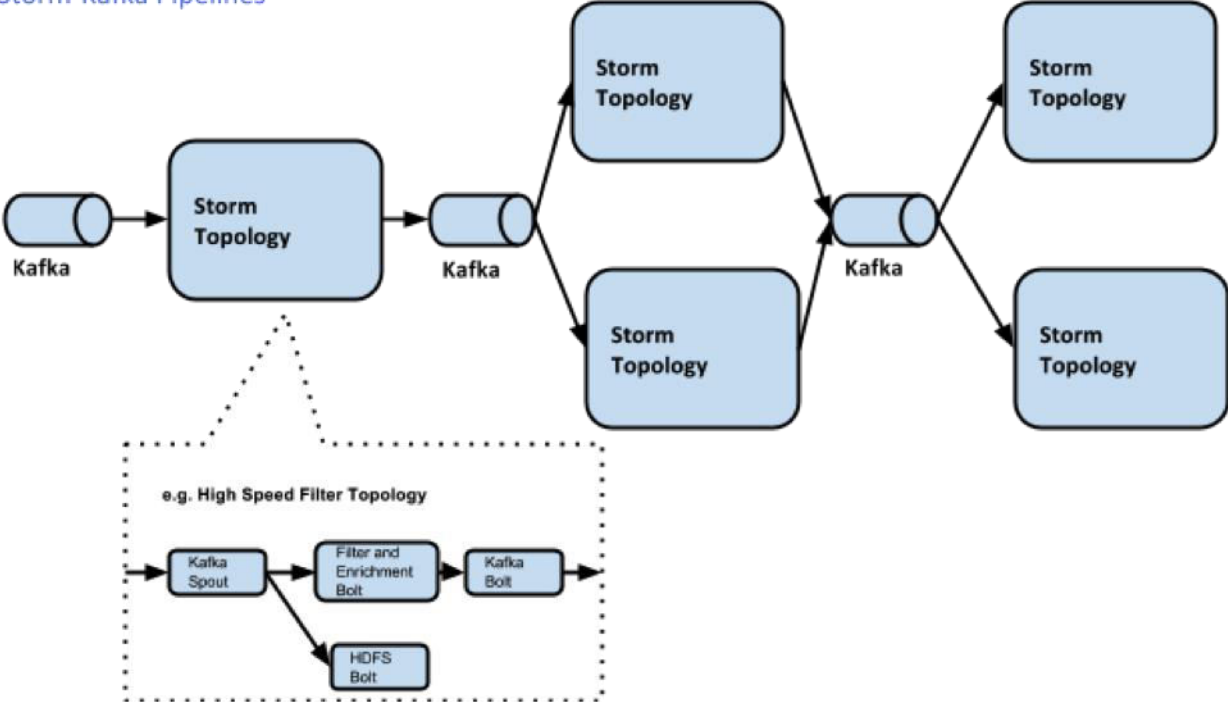
```
public class MyFunction extends BaseFunction {  
    public void execute(TridentTuple tuple, TridentCollector collector) { ... }  
}
```

```
public class MyFilter extends BaseFilter {  
    public boolean isKeep(TridentTuple tuple) { ... }  
}
```



# Storm/Trident w/Kafka

Storm-Kafka Pipelines



# Vowpal Wabbit

- Online/out-of-core learning
- Extraordinarily fast – handles large sparse feature spaces
- Broad selection of algorithms
- Scalable (all reduce)
- Active & reinforcement learning
- Contextual bandit FTW!

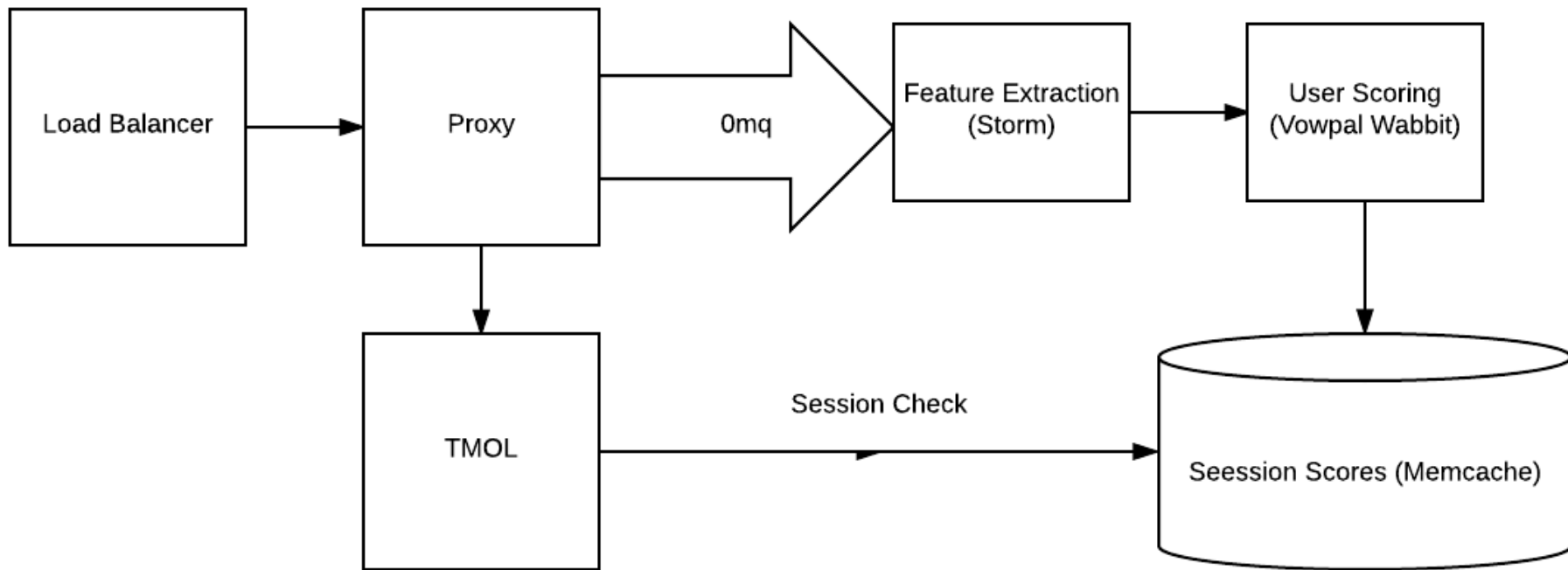
# Elastic Search

- Generic, sharded & redundant search engine
- Real-time indexing
- Distinct master, data, ingest, coordination node roles
- Handles semi-structured data
- Kibana: secret weapon

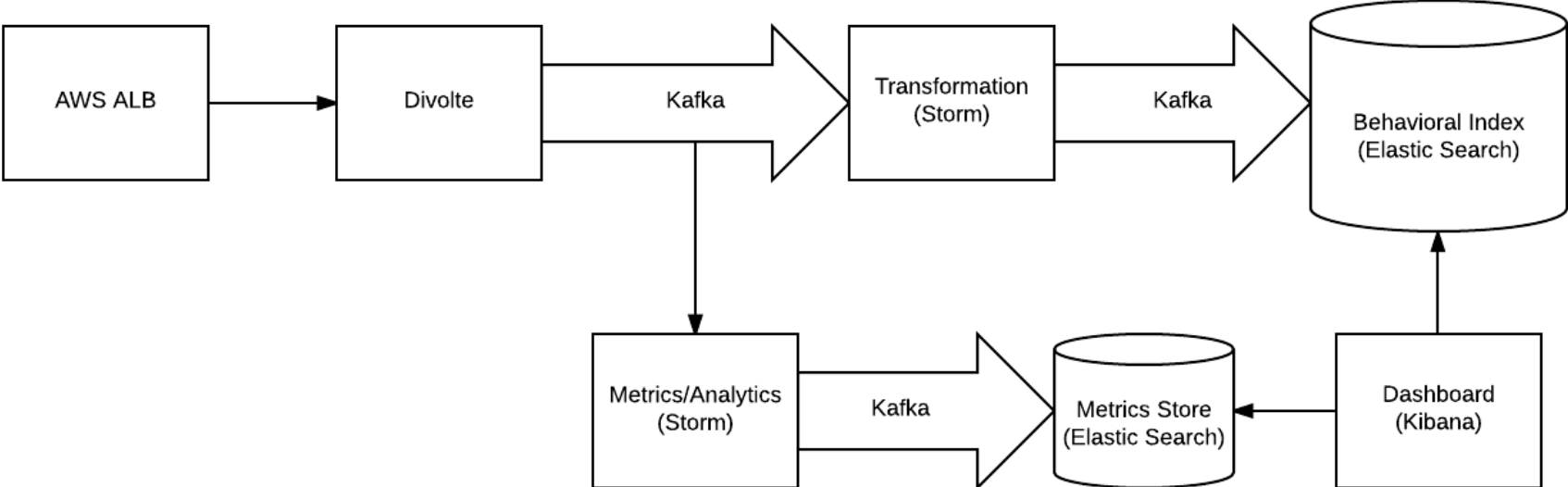
# Kibana



# Abuse Prevention



# Example - Analytics



# Auto Delist on Entry

ticketsnow.com

Welcome to TicketsNow. Log In | Join

Enter Artist, Team or Venue Select City / Area Search

Sports Concerts Theater Sell Your Tickets

US Open Tennis Tickets **G4** | 10 **Event**

US Open Tennis is the conduit for the best in men's and women's tennis to take to the courts in competition! The US Open is a premier Grand Slam event, and you don't want to miss your chance to be part of the US Open action! Purchase US Open Tennis tickets now.

For Grounds Admission, please click here.

- Louis Armstrong tickets and Grounds Admission tickets do not include Arthur Ashe Stadium access.

For Arthur Ashe Kids' Day, please click here.

It is the policy of the US Open Ticket Exchange and USTA to remove each event from the Ticket Exchange 2 hours prior to the beginning of the Session, in order to make sure all Sellers and Buyers have time to complete their transactions.

**Sold by fans. Verified by Ticketmaster.®**  
Tickets are verified by Ticketmaster and reissued in your name.  
**Buy. Download. Print.**

US OPEN

Event	When	Where	Sell Tickets	View Tickets
US Open Tennis Tickets	Mon, Aug. 29th, 2016 11:00 AM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 1				
US Open Tennis Tickets	Mon, Aug. 29th, 2016 11:00 AM	USTA Grandstand (Flushing, NY)	Sell Tickets	View Tickets
Session 1				
US Open Tennis Tickets	Mon, Aug. 29th, 2016 11:00 AM	Louis Armstrong Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 1				
US Open Tennis Tickets	Mon, Aug. 29th, 2016 07:00 PM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 2				

FanGraph

ticketsnow.com

Welcome to TicketsNow. Log In | Join

Enter Artist, Team or Venue Select City / Area Search

Sports Concerts Theater Sell Your Tickets

US Open Tennis Tickets **G4** | 10 **Event**

US Open Tennis is the conduit for the best in men's and women's tennis to take to the courts in competition! The US Open is a premier Grand Slam event, and you don't want to miss your chance to be part of the US Open action! Purchase US Open Tennis tickets now.

For Grounds Admission, please click here.

- Louis Armstrong tickets and Grounds Admission tickets do not include Arthur Ashe Stadium access.

For Arthur Ashe Kids' Day, please click here.

It is the policy of the US Open Ticket Exchange and USTA to remove each event from the Ticket Exchange 2 hours prior to the beginning of the Session, in order to make sure all Sellers and Buyers have time to complete their transactions.

**Sold by fans. Verified by Ticketmaster.®**  
Tickets are verified by Ticketmaster and reissued in your name.  
**Buy. Download. Print.**

US OPEN

Event	When	Where	Sell Tickets	View Tickets
US Open Tennis Tickets	Mon, Aug. 29th, 2016 11:00 AM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 1				
US Open Tennis Tickets	Mon, Aug. 29th, 2016 11:00 AM	USTA Grandstand (Flushing, NY)	Sell Tickets	View Tickets
Session 1				
US Open Tennis Tickets	Mon, Aug. 29th, 2016 11:00 AM	Louis Armstrong Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 1				
US Open Tennis Tickets	Mon, Aug. 29th, 2016 07:00 PM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 2				
US Open Tennis Tickets	Tue, Aug. 30th, 2016 11:00 AM	USTA Grandstand (Flushing, NY)	Sell Tickets	View Tickets
Session 3				
US Open Tennis Tickets	Tue, Aug. 30th, 2016 11:00 AM	Louis Armstrong Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 3				
US Open Tennis Tickets	Tue, Aug. 30th, 2016 11:00 AM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 3				
US Open Tennis Tickets	Tue, Aug. 30th, 2016 07:00 PM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 4				
US Open Tennis Tickets	Wed, Aug. 31st, 2016 11:00 AM	USTA Grandstand (Flushing, NY)	Sell Tickets	View Tickets
Session 5				
US Open Tennis Tickets	Wed, Aug. 31st, 2016 11:00 AM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 5				
US Open Tennis Tickets	Wed, Aug. 31st, 2016 11:00 AM	Louis Armstrong Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 5				
US Open Tennis Tickets	Wed, Aug. 31st, 2016 07:00 PM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 5				
US Open Tennis Tickets	Thu, Sep. 1st, 2016 11:00 AM	Louis Armstrong Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 7				
US Open Tennis Tickets	Thu, Sep. 1st, 2016 11:00 AM	Arthur Ashe Stadium (Flushing, NY)	Sell Tickets	View Tickets
Session 7				



ticketmaster®

# Lambda Architecture FTW!

- Reusable functions
- Dev team autonomy/loose coupling
- Simplifies handling failure
- Simplifies scaling
- Complex computation with low latency
- Even less computationally intensive tasks become easier



*ticketmaster*<sup>®</sup>

