FREEWHEEL
A COMCAST COMPANY

PROCESSING BILLONS OF AD-TECH EVENTS IN REAL-TIME

MARGI DUBAL

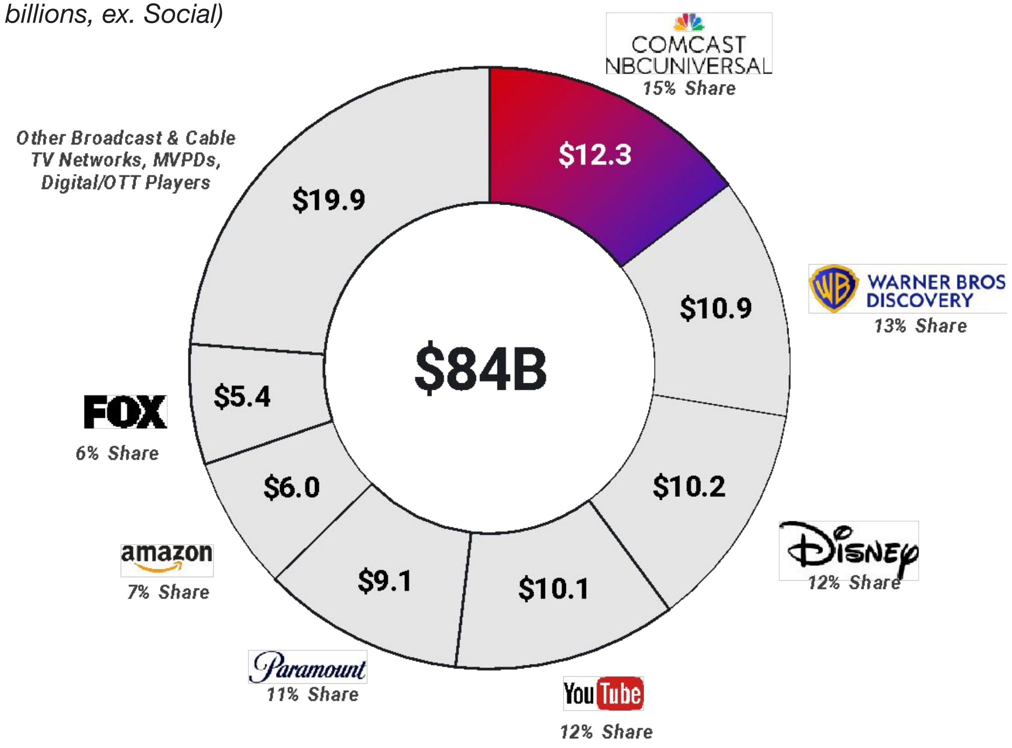
March 30, 2023

WHO ARE WE?

FREEWHEEL A COMCAST COMPANY

COMCAST/NBCU IS #1 SELLER OF VIDEO ADVERTISING IN U.S.

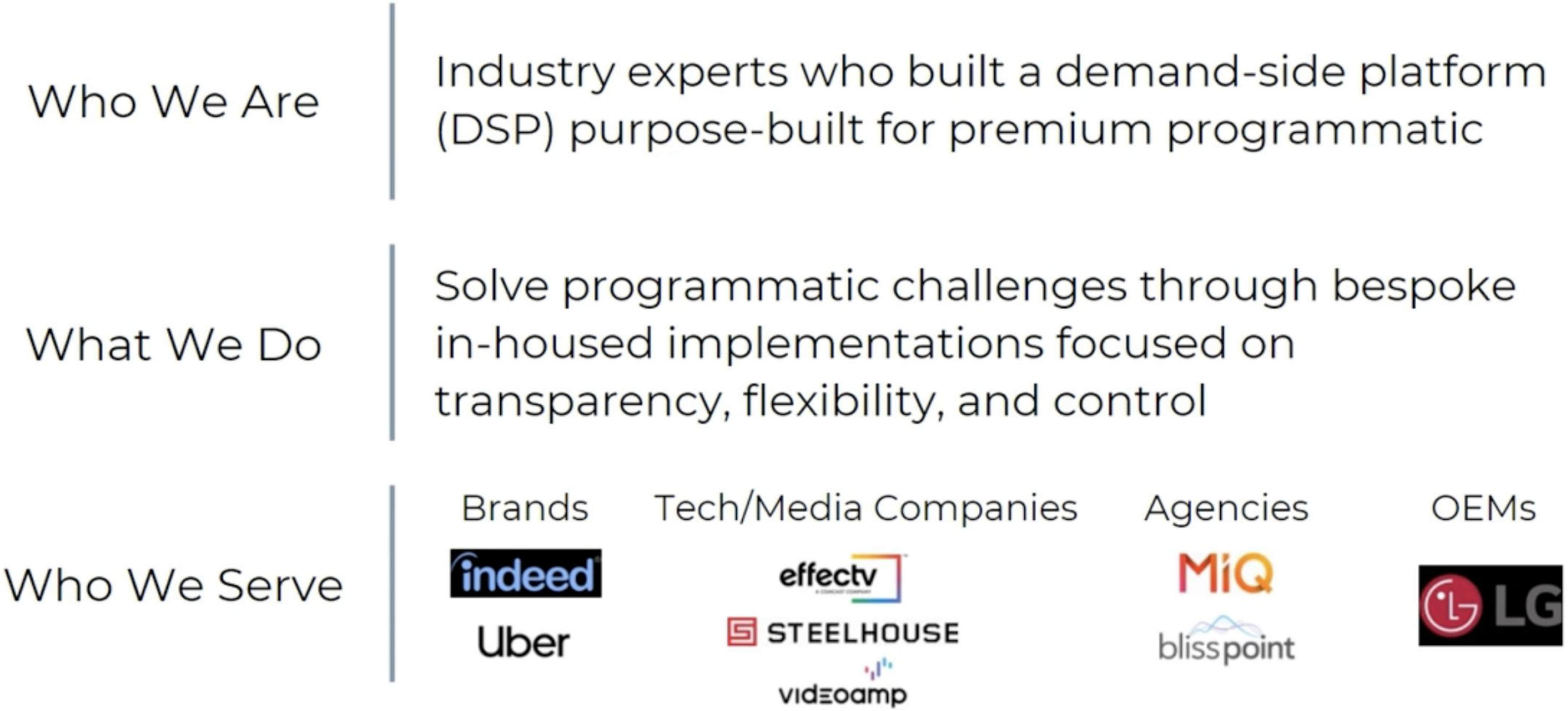
2021 Total U.S. Video Ad Revenue
($ in billions, ex. Social)



FREEWHEEL POWERS ~90% OF PREMIUM VIDEO IN U.S. & EUROPE

NBC, hulu, A+E NETWORKS, dish, AT&T
TV LAND, sky, COX, AMC, WARNER MEDIA, Aol.
CBS, TF1, COMCAST, Charter Spectrum, fios by verizon
univision, ROGERS, Turner, VIACOMCBS, altice
FOX, CRACKLE, Discovery, Frontier COMMUNICATIONS, TNT
MTV, food, HGTV, COMEDY CENTRAL, nickelodeon, PGA TOUR
OXYGEN, Syfy, animal planet, TLC, travel CHANNEL, tbs, fyi,
BET, lifetime, usa, VH1, CMT, bravo, E!

BEESWAX AT A GLANCE



Who We Are	Industry experts who built a demand-side platform (DSP) purpose-built for premium programmatic
What We Do	Solve programmatic challenges through bespoke in-housed implementations focused on transparency, flexibility, and control
Who We Serve	

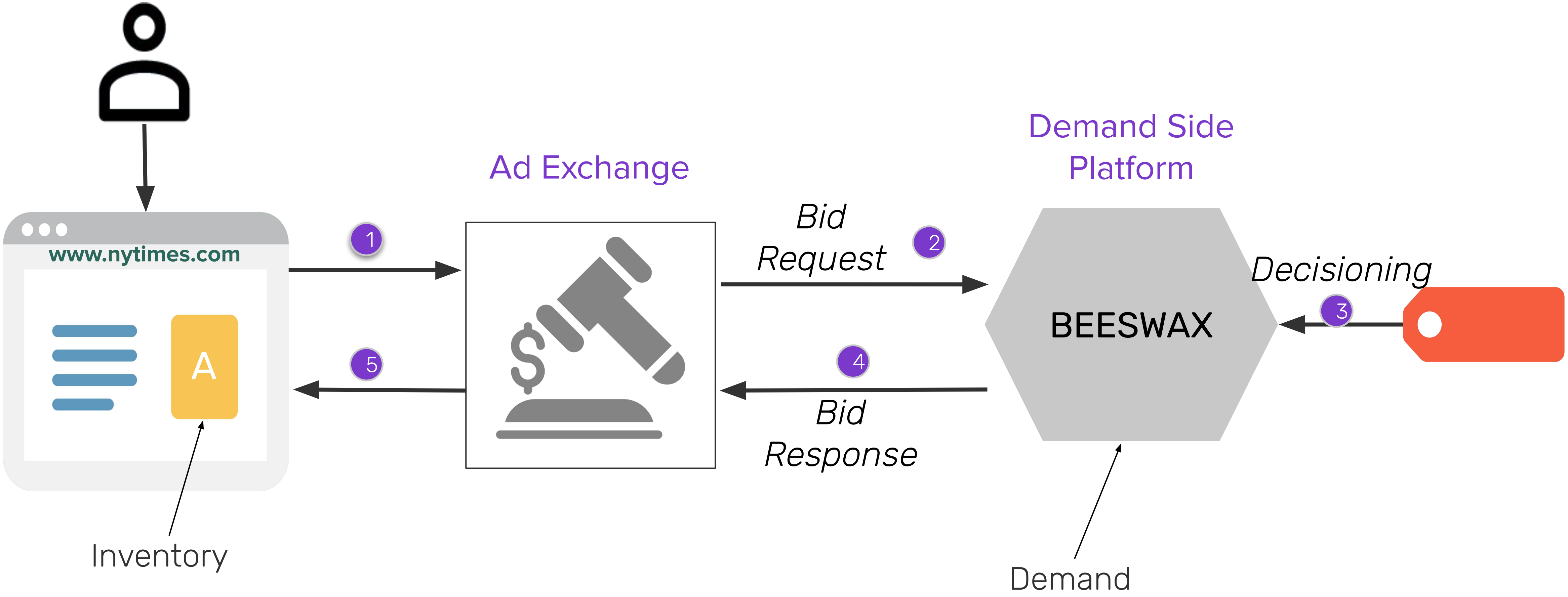
Brands	Tech/Media Companies	Agencies	OEMs
indeed Uber	effectv STEELHOUSE videoamp	MiQ blisspoint	LG

BASICS OF RTB (REAL TIME BIDDING)

Real-time bidding (RTB) is a method of transferring media on an individual impression basis, which is bid on in real time via an ‘auction’.

ADTECH 101 & REAL TIME BIDDING

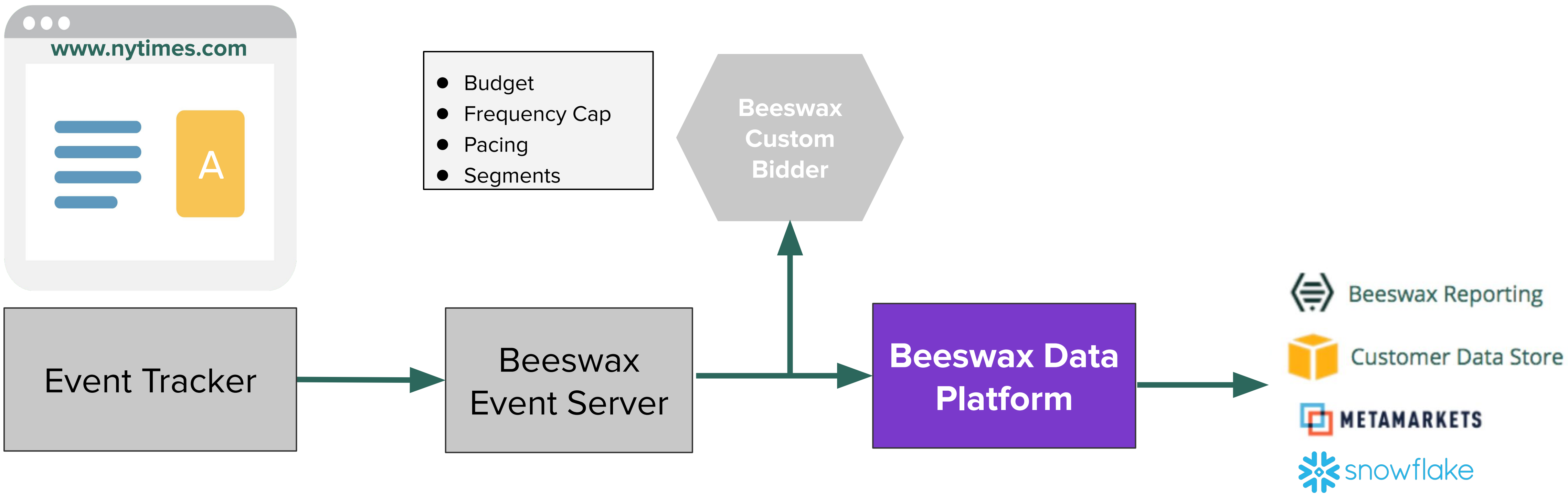
Real time bidding basics



www.nytimes.com
A
Inventory
1
5
Ad Exchange
$
Bid Request
2
4
Bid Response
Demand Side Platform
BEESWAX
Decisioning
3
Demand

WHAT HAPPENS AFTER AD IS SERVED

Real time bidding basics



- Budget
- Frequency Cap
- Pacing
- Segments

Beeswax Custom Bidder

Event Tracker

Beeswax Event Server

Beeswax Data Platform

Beeswax Reporting

Customer Data Store

METAMARKETS

snowflake

EVENT TYPES

Bid Request	Req ID	Creative				
Win Notification	Req ID		Win Price			
Event Tracker	Req ID			Click URL	Time stamp	Event Type
Complete Record	Req ID	Creative	Win Price	Click URL	Time stamp	Event Type

INFRASTRUCTURE HIGHLIGHTS

DATA SCALE

Infrastructure Highlights

Every second

2.5M
Concurrent requests

200K – 500K
QPS to data platform

Every day

10B
Ad Events logged

200TB
Data processed

Every year

400TB
Data for reporting & analytics

6T
Bid requests handled

DATA PLATFORM

COLLECTING, ORGANIZING & DELIVERING DATA

Raw logs delivery & accurate campaign performance reports

50+ Data pipelines are scheduled using Airflow

Elastically scalable infrastructure, optimized for cost

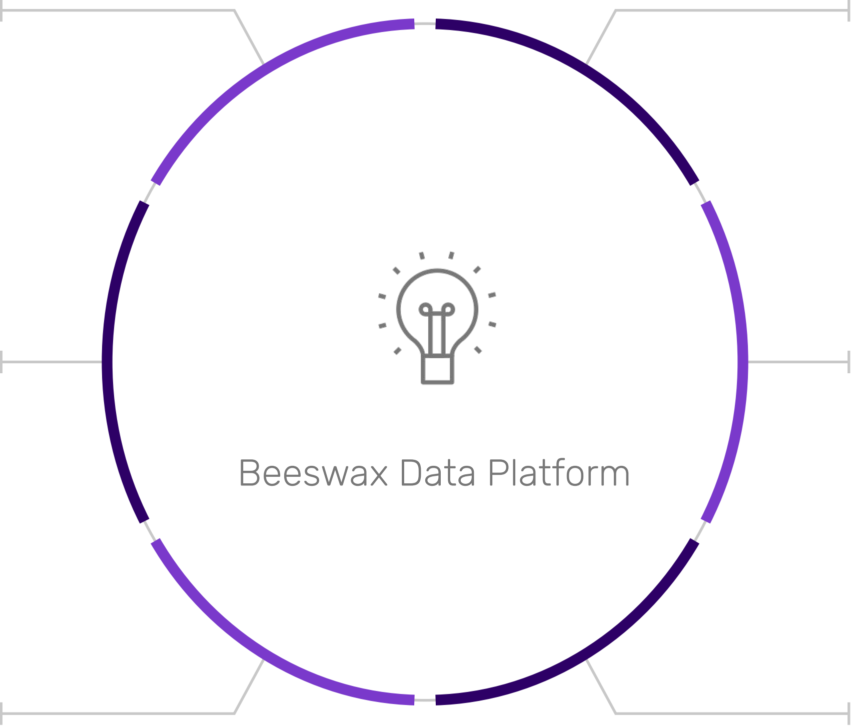
Architected to run on AWS Spot instances and use autoscaling

Diverse set of datastores

Aerospike (KeyValue), Snowflake (Datawarehouse), Kinesis (Stream), S3 (Objects), RDS (MySql/Postgres), InfluxDB (Timeseries)

DATA PRODUCT – DIVERSE DELIVERY METHODS

Infrastructure highlights



STREAMING LOGS

Real-time log delivery.
Support S3, Kinesis and HTTP integrations..

ANTENNA DATA SHARING

Allow customers to access raw log data using Snowflake.

BATCH LOGS

Deliver hourly and daily logs to customers' S3 buckets.

REPORTING UI

14 built-in reports.
Embedded Looker UI provides various visualizations and scheduled reports.

REPORTING API

Leverage Looker API to provide same capability as reporting UI.

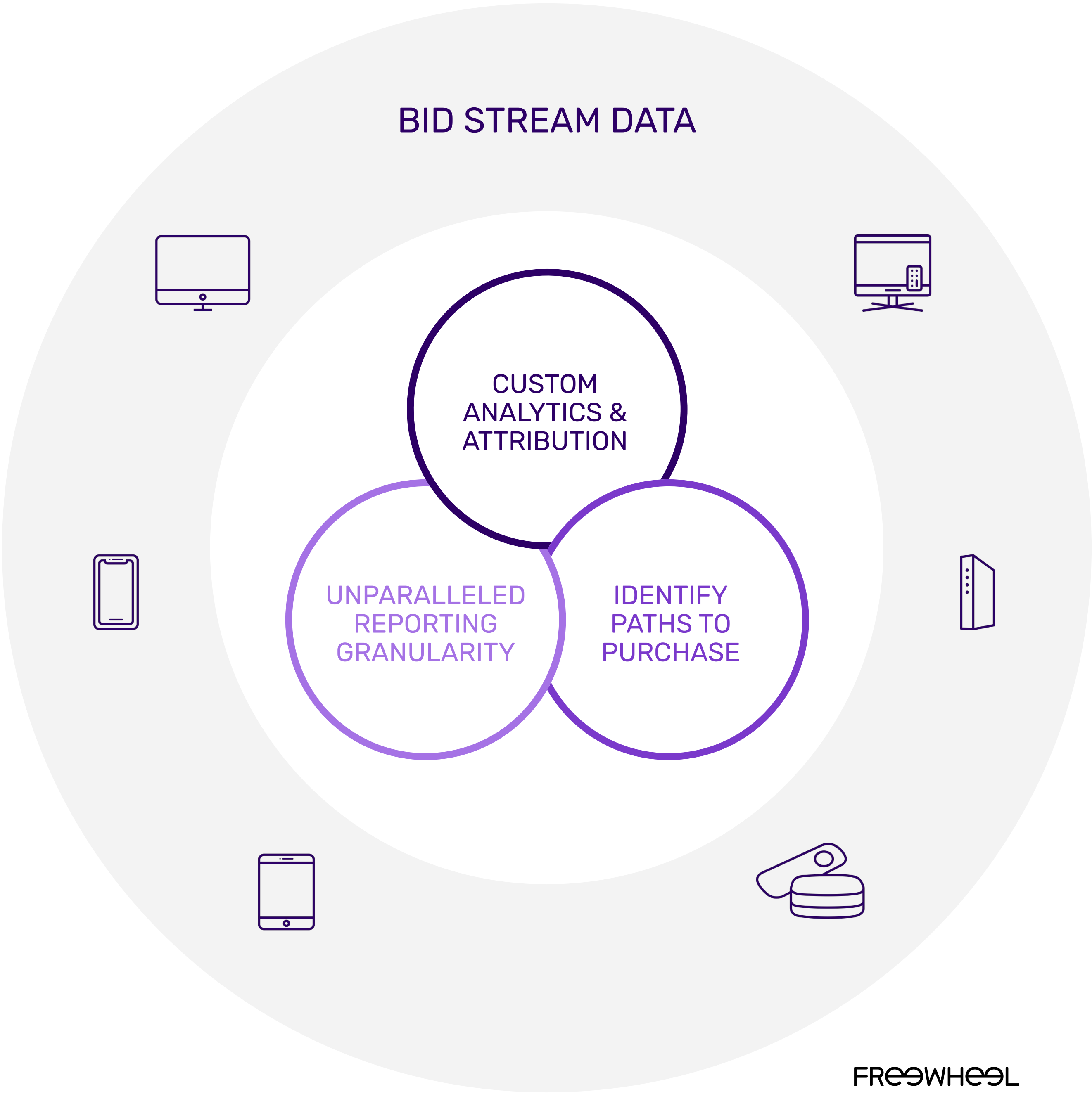
REAL-TIME BIDDING STATS

Visualize the real-time bidding/winning status.

DEEPER INSIGHTS WITH EVENT LOG DATA

Transparency into why campaigns perform and who drives performance

- Complete control and flexibility during reporting
- Join data any way you want to identify custom insights
- Make the most of your spend



BID STREAM DATA

CUSTOM ANALYTICS & ATTRIBUTION

UNPARALLELED REPORTING GRANULARITY

IDENTIFY PATHS TO PURCHASE

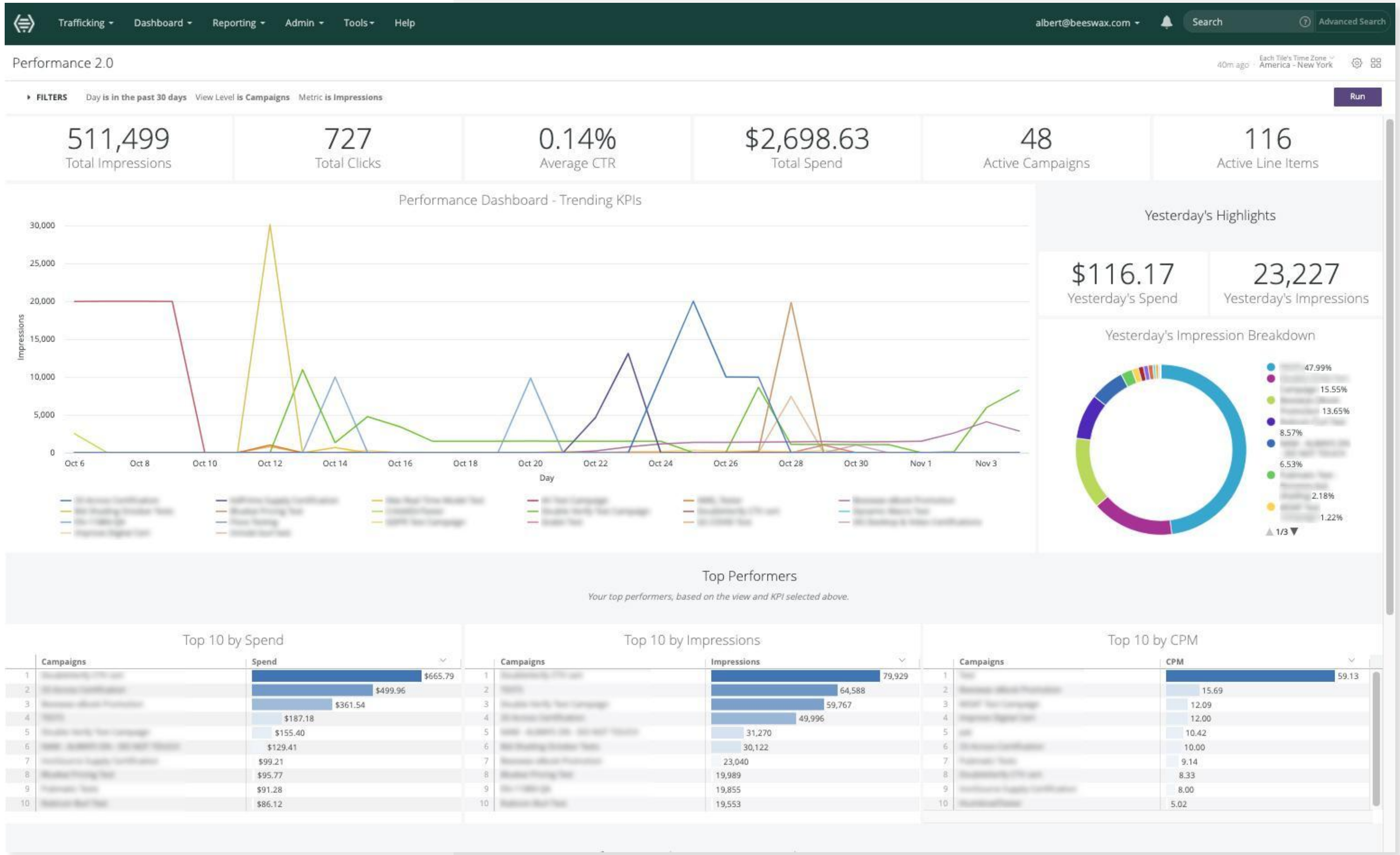
EVENT LOG SUMMARY

Different types of logs, format, delivery cadences

CADENCE	LOG TYPE	DESCRIPTION	DELIVERY	FORMAT
Near real-time	**Auctions**	The auction request from exchange	S3	Gzipped csv
Near real-time	**Bids**	Bids returned from Beeswax to exchange	S3	Gzipped csv
Near real-time	**Conversions**	Conversion events recorded by Beeswax regardless of attribution	S3	Gzipped csv or parquest
8 hours delay	**Attributed Conversions**	Conversion events recorded by Beeswax	S3	Gzipped csv or parquest
Hourly and Daily	**Wins**	The winning auctions, impressions, clicks and video activities	S3	Gzipped csv or parquet
Realtime	**Wins (Joined and Unjointed)**	Stream of events associated with impression	HTTP or Kinesis	Protobuf or Json
1 hour delay	**Losses**	Loss logs provided by limited set of exchanges	S3	Gzipped csv

REPORTING PLATFORM

- Intuitive data visualizations
- Customizable
- Cross-channel reporting


Trafficking
Dashboard
Reporting
Admin
Tools
Help
albert@beeswax.com
Search
Advanced Search
Performance 2.0
40m ago
Each Tile's Time Zone
America - New York
FILTERS
Run
Performance Dashboard - Trending KPIs
30,000
25,000
20,000
15,000
10,000
5,000
0
Impressions
Oct 6
Oct 8
Oct 10
Oct 12
Oct 14
Oct 16
Oct 18
Oct 20
Oct 22
Oct 24
Oct 26
Oct 28
Oct 30
Nov 1
Nov 3
Day
Yesterday's Impression Breakdown
47.99%
15.55%
13.65%
8.57%
6.53%
2.18%
1.22%
1/3

Day is in the past 30 days View Level is Campaigns Metric is Impressions

511,499	727	0.14%	$2,698.63	48	116
Total Impressions	Total Clicks	Average CTR	Total Spend	Active Campaigns	Active Line Items

Yesterday's Highlights

$116.17 Yesterday's Spend

23,227 Yesterday's Impressions

Top Performers

Your top performers, based on the view and KPI selected above.

Top 10 by Spend

	Campaigns	Spend
1	[illegible]	$665.79
2	[illegible]	$499.96
3	[illegible]	$361.54
4	[illegible]	$187.18
5	[illegible]	$155.40
6	[illegible]	$129.41
7	[illegible]	$99.21
8	[illegible]	$95.77
9	[illegible]	$91.28
10	[illegible]	$86.12

Top 10 by Impressions

	Campaigns	Impressions
1	[illegible]	79,929
2	[illegible]	64,588
3	[illegible]	59,767
4	[illegible]	49,996
5	[illegible]	31,270
6	[illegible]	30,122
7	[illegible]	23,040
8	[illegible]	19,989
9	[illegible]	19,855
10	[illegible]	19,553

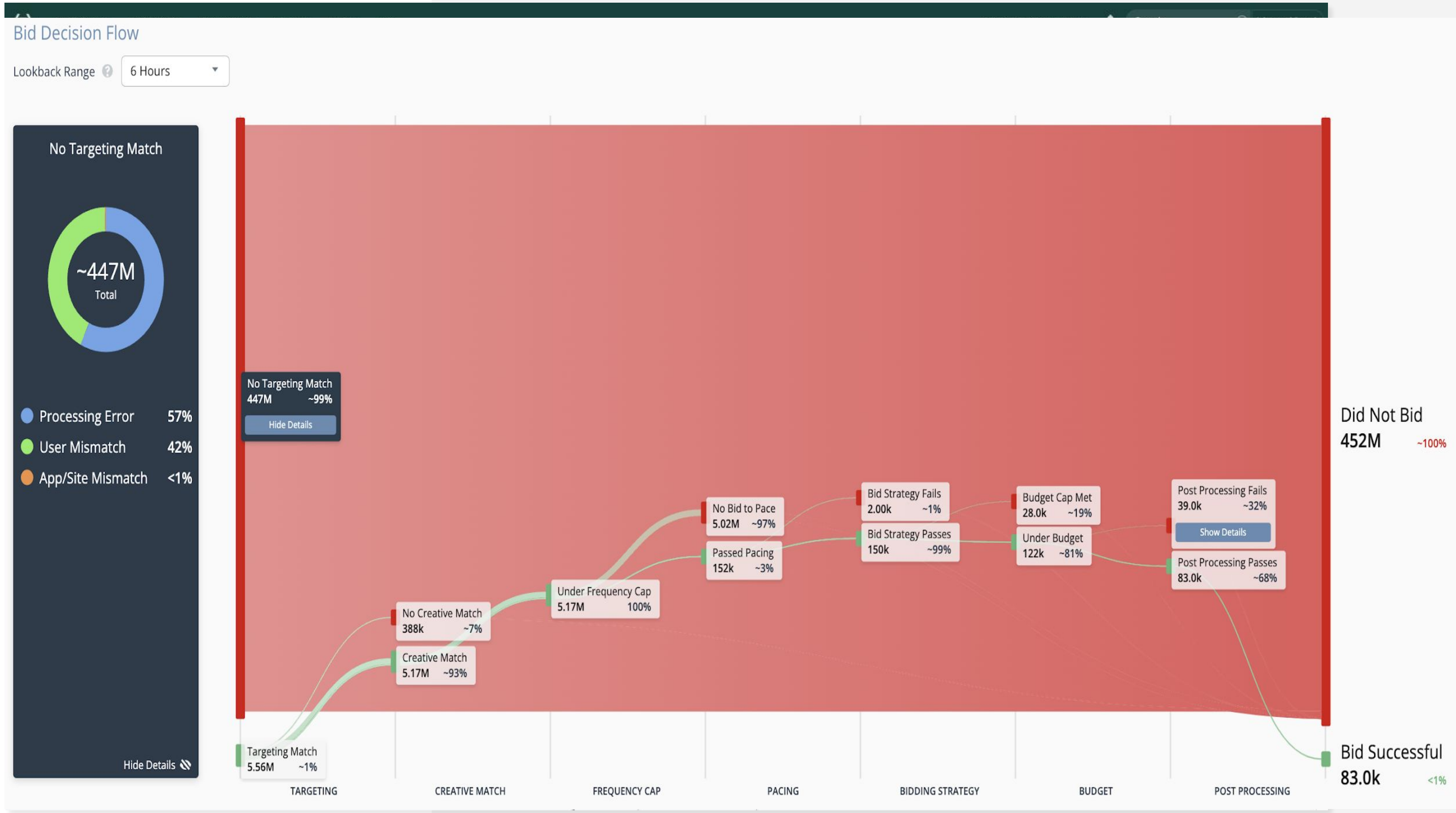
Top 10 by CPM

	Campaigns	CPM
1	[illegible]	59.13
2	[illegible]	15.69
3	[illegible]	12.09
4	[illegible]	12.00
5	[illegible]	10.42
6	[illegible]	10.00
7	[illegible]	9.14
8	[illegible]	8.33
9	[illegible]	8.00
10	[illegible]	5.02

REPORTING PLATFORM

BID DECISION FLOW:

- Do we have enough supply?
- Are we winning enough?
- Are our bids priced right?
- How to improve my bidding strategy?



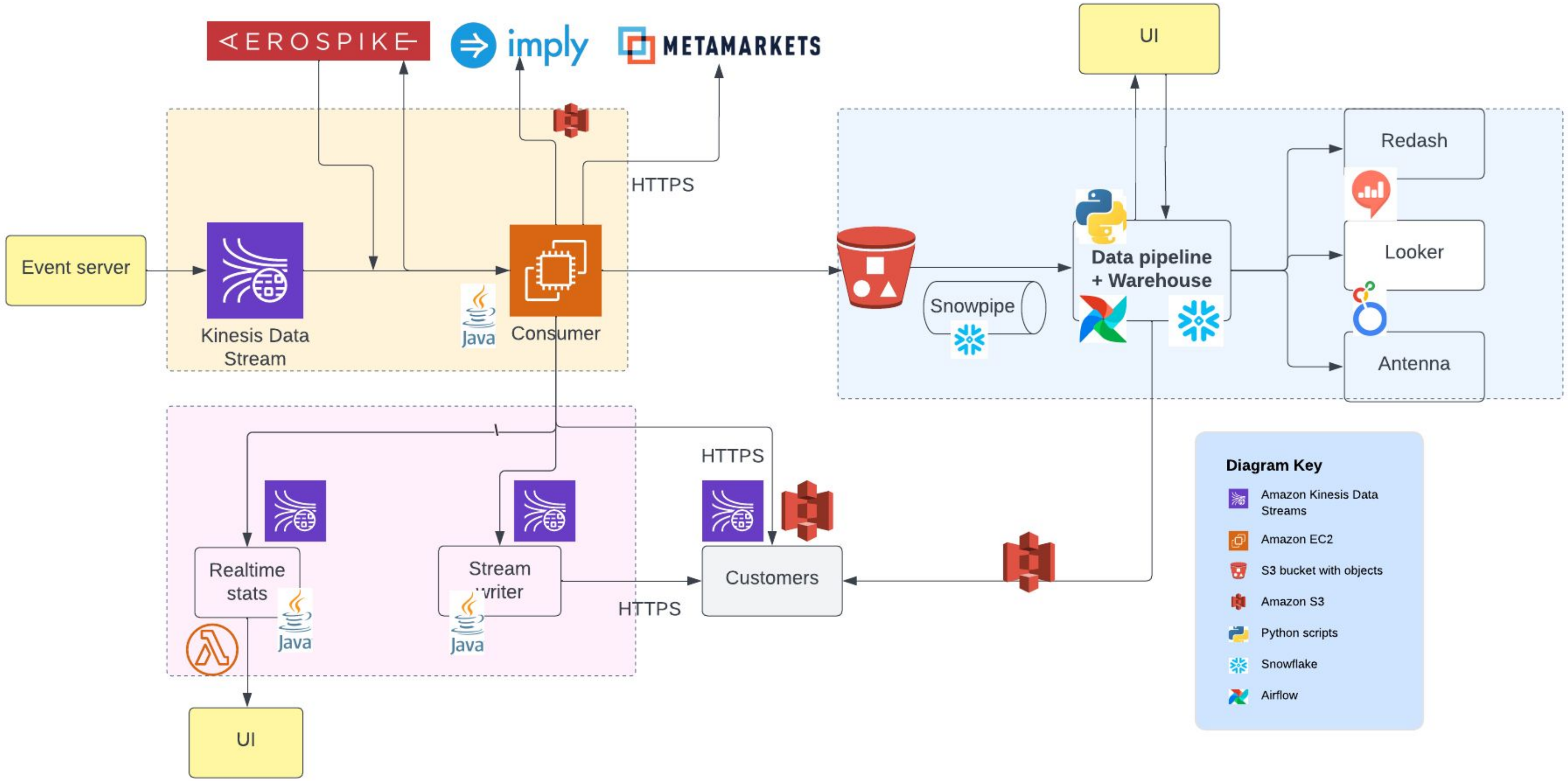
Bid Decision Flow

Lookback Range: 6 Hours

No Targeting Match
~447M
Total
● Processing Error 57%
● User Mismatch 42%
● App/Site Mismatch <1%
Hide Details
No Targeting Match 447M ~99%
Hide Details
Targeting Match 5.56M ~1%
No Creative Match 388k ~7%
Creative Match 5.17M ~93%
Under Frequency Cap 5.17M 100%
No Bid to Pace 5.02M ~97%
Passed Pacing 152k ~3%
Bid Strategy Fails 2.00k ~1%
Bid Strategy Passes 150k ~99%
Budget Cap Met 28.0k ~19%
Under Budget 122k ~81%
Post Processing Fails 39.0k ~32%
Show Details
Post Processing Passes 83.0k ~68%
Did Not Bid 452M ~100%
Bid Successful 83.0k <1%
TARGETING
CREATIVE MATCH
FREQUENCY CAP
PACING
BIDDING STRATEGY
BUDGET
POST PROCESSING

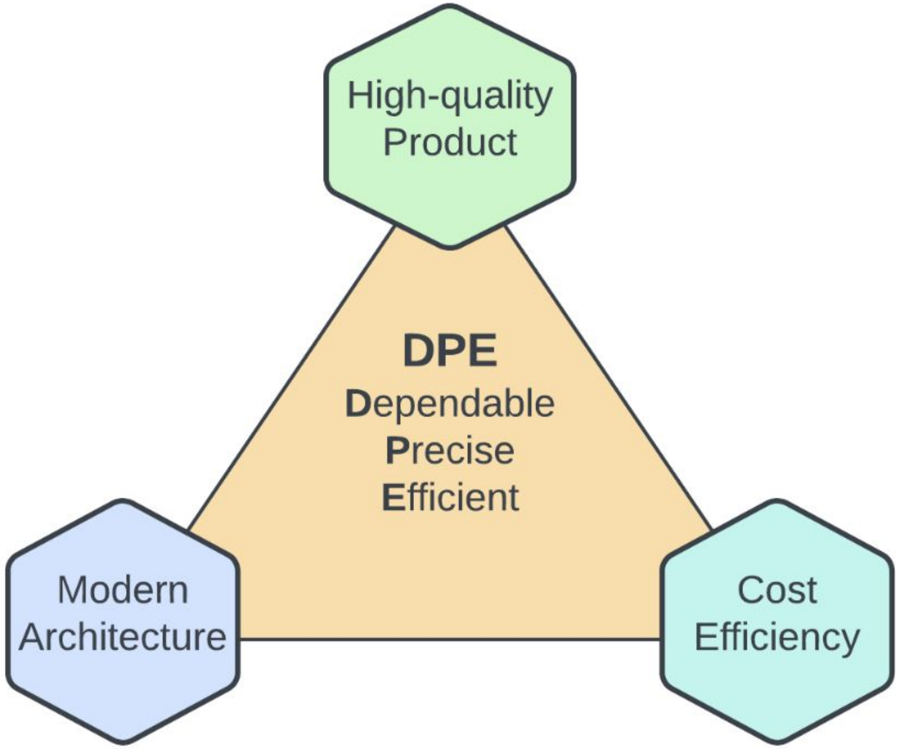
DATA ARCHITECTURE

ARCHITECTURE


AEROSPIKE
imply
METAMARKETS
UI
HTTPS
Event server
Kinesis Data Stream
Java
Consumer
Redash
Data pipeline + Warehouse
Looker
Snowpipe
Antenna
HTTPS
Realtime stats
Java
Stream writer
Java
HTTPS
Customers
UI
Diagram Key
Amazon Kinesis Data Streams
Amazon EC2
S3 bucket with objects
Amazon S3
Python scripts
Snowflake
Airflow

PHILOSOPHY BEHIND THE TECHNOLOGY CHOICES

GUIDING PRINCIPLES


High-quality Product
DPE
Dependable
Precise
Efficient
Modern Architecture
Cost Efficiency

SCALABILITY & EFFICIENCY

Should be able to scale with growth without incurring major cost.

DEFENSIBLE DIFFERENTIATOR

Support customer needs

BUY TECH AND BUILD ADTECH

Focus engineering bandwidth on the differentiators.

TIME TO MARKET

TECHNOLOGY CHOICES

AWS KINESIS

- Fully managed service
- Easy to setup

AEROSPIKE

- Fast real-time NoSQL database
- Exceptional performance & high availability

SNOWFLAKE

- Internal warehouse
- Secure sharing
- Distinguishes compute from storage

LOOKER

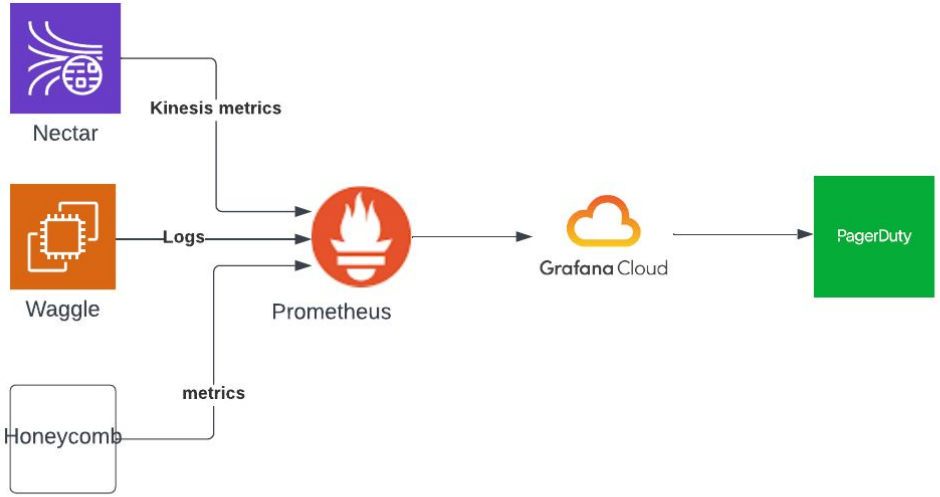
- Analytics and reporting platform.
- Embedded in Buzz App UI.
- Easy to build reports

REDASH

- Democratizing access to data
- Data query & visualization tool

MONITORING & ALERTING

MONITORING & ALERTING



Nectar

Kinesis metrics

Waggle

Logs

Honeycomb

metrics

Prometheus

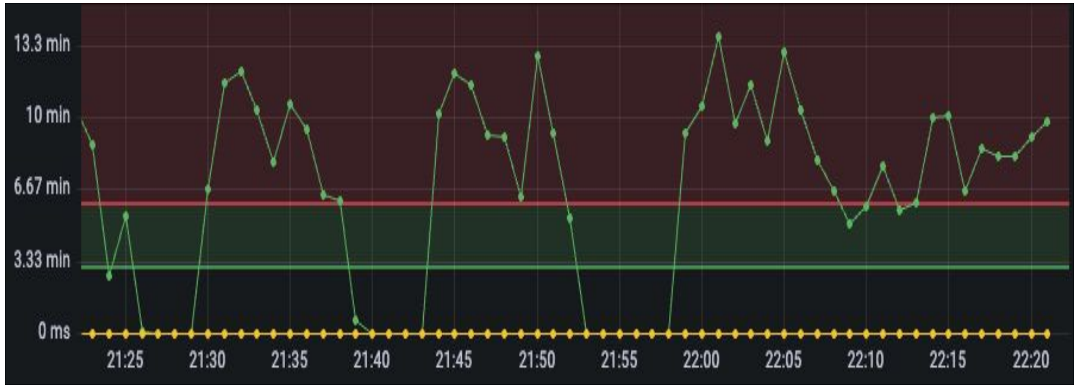
Grafana Cloud

PagerDuty

- 6.5 K metrics
- 8K GB logs
- 50+ alerts configured

Alerts:

- Aging messages on Kinesis stream
- Capacity alerts
- Pipeline failure alerts



ROADMAP

ROADMAP

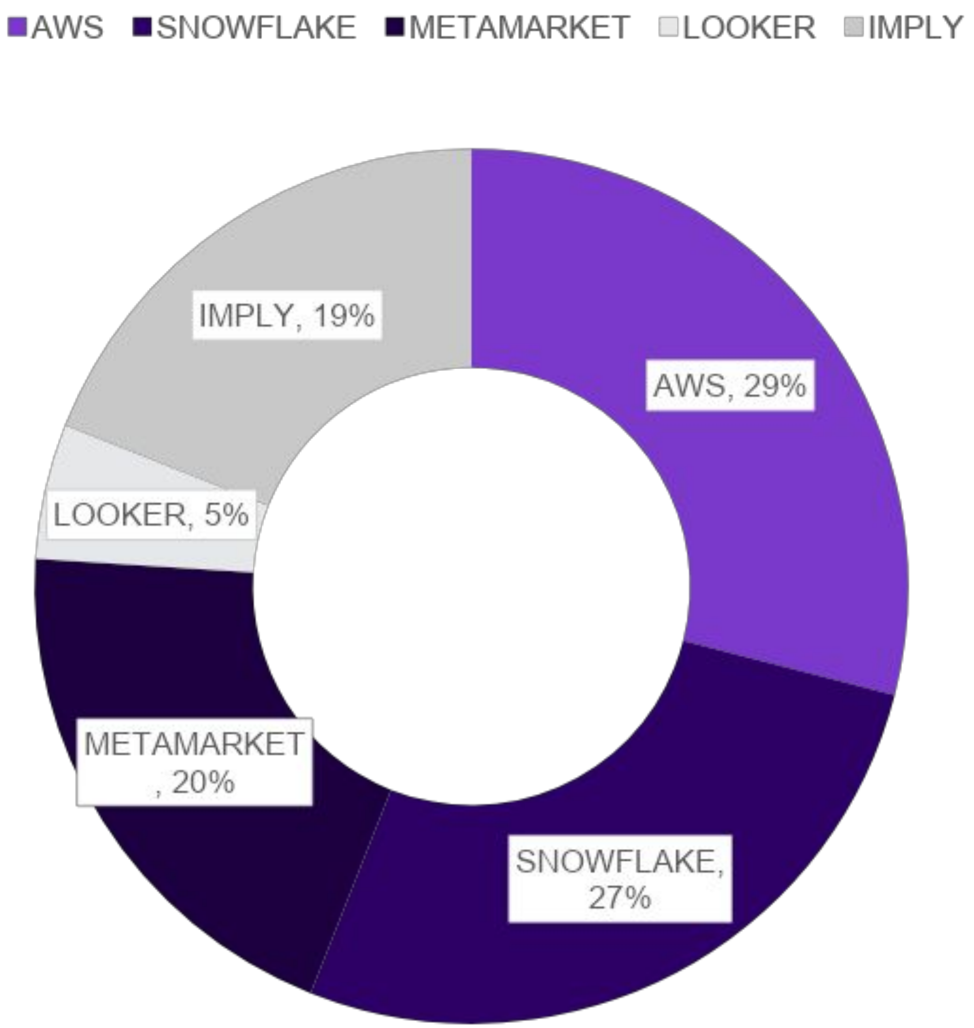
- Increase scalability, performance & cost effectiveness
    - Bring the consumer applications onto Databricks Spark Streaming
    - Rebuild real-time bidding stats using Spark for faster performance
- Improve data quality
    - Great expectation
- Improve data observability
    - data cataloging and lineage

Building a solid data platform is more about using the right technology and focusing on the differentiator for each use case and having a right sized team to support it.

THANK YOU

FREEWHEEL
A COMCAST COMPANY

COST BREAKDOWN



AGENDA

Table of Contents

1 Who we are?

2 Basics of bid request

3 Infrastructure Highlights

4 Data Architecture

5 Philosophy behind technology choices

6 Monitoring & Alerting

7 Cost Breakdown & Roadmap

8 Questions