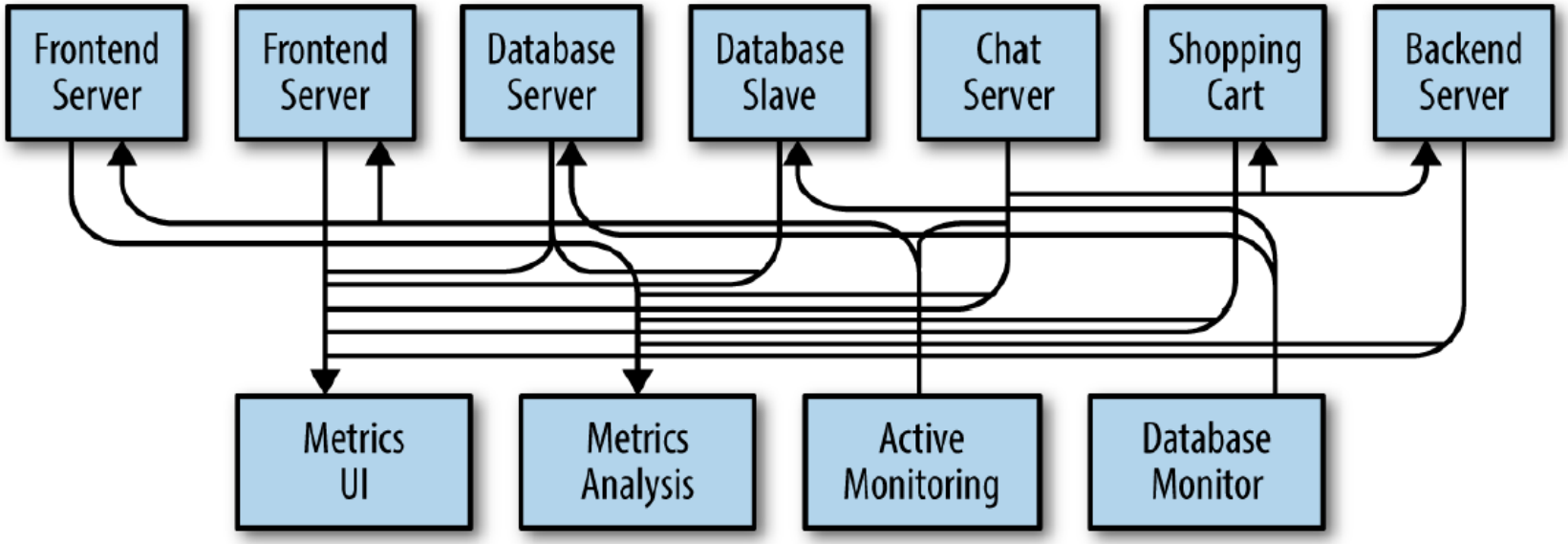
# Apache Kafka

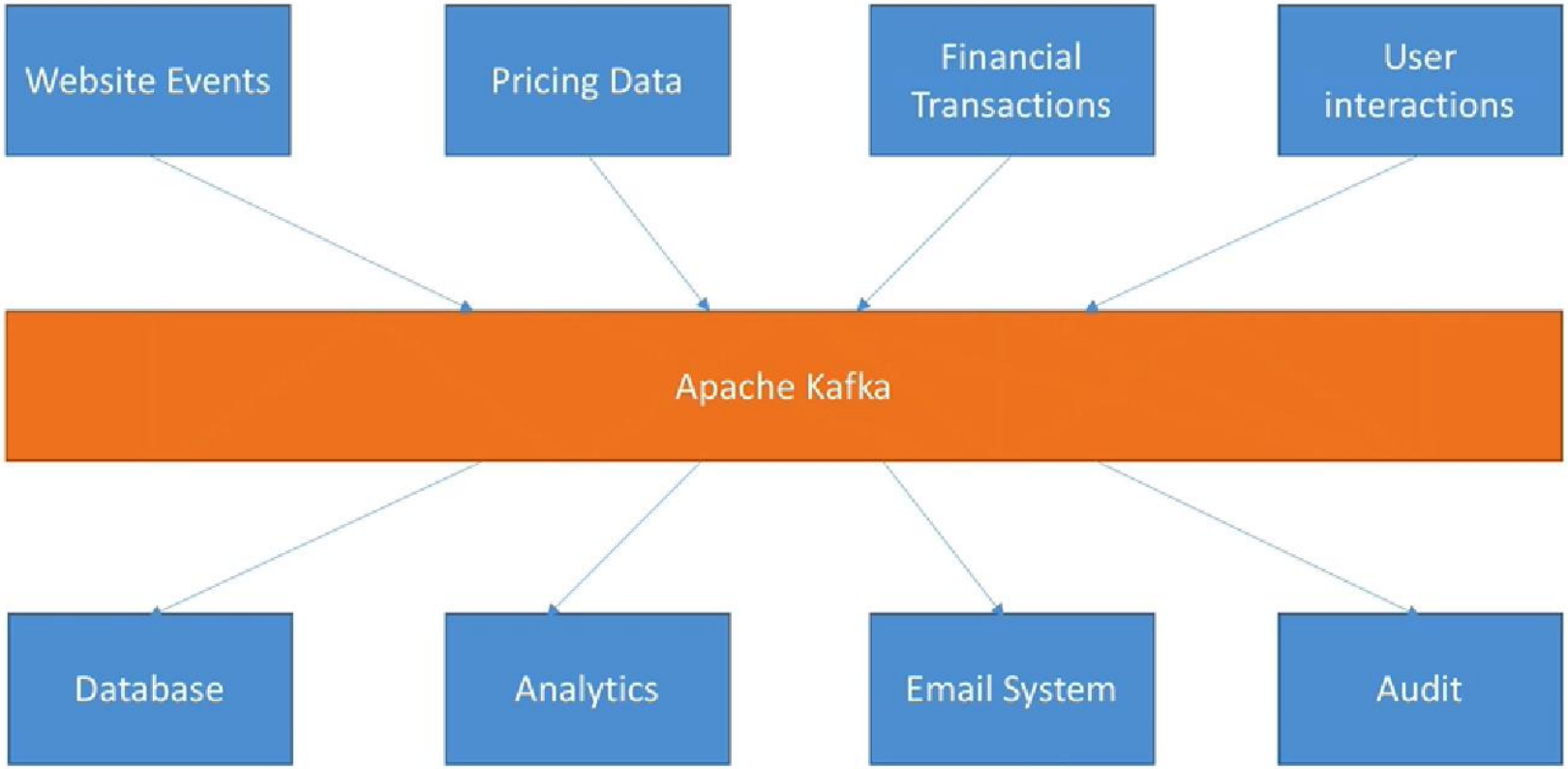
# Publish/Subscribe Messaging

- Current challenges in industry.

Cloudlytics




Cloudlytics




Cloudlytics

- Apache Kafka is an open source, distributed, and replicated commit-log-based publish-subscribe messaging system.
- Resilient architecture & fault tolerant
- Horizontal scalability
- High performance (latency of less than 10ms) –real time
- Used by 2000+ firms

  (LinkedIn, Netflix, Airbnb, Yahoo, Walmart, etc)

Cloudlytics

# Why Kafka ?

- Multiple Producers
- Multiple Consumers
- Disk-based Retention
- Scalable
- High performance

Cloudlytics

# Core Concepts

- Messages and batches
- Topics and Partitions
- Brokers
- Topic replication factor
- Leader for partition

Cloudlytics

# Producers and Consumers

- Producers
  - Write Acknowledgement
  - Message Keys
- Consumers
  - Consumer Groups
  - Consumer offsets
- Role of Zookeeper

Cloudlytics