

PunchPlatform  
Kafka Primer



PunchPlatform team

 Agenda

- Kafka Role
- Best Practices



## Kafka Role

The PunchPlatform is built on top of Kafka. Kafka is used as intermediary queues to secure the data between the several processing hops. Kafka can scale up to very large clusters (checkout the LinkedIn numbers), takes care of replicating the data, and let several consumers consume the data in real time.

In order to scale, producer and consumers write (read) the data in **multi partitioned topics**. Think of a *topic* as a giant fifo structure, divided in several sub-queues called **partitions**. Several **producers** write logs. Several **consumers** read them in real time. The key factor to scale is the number of partitions. You have one reader for each partition. Each must read fast enough to continuously consume its partition.

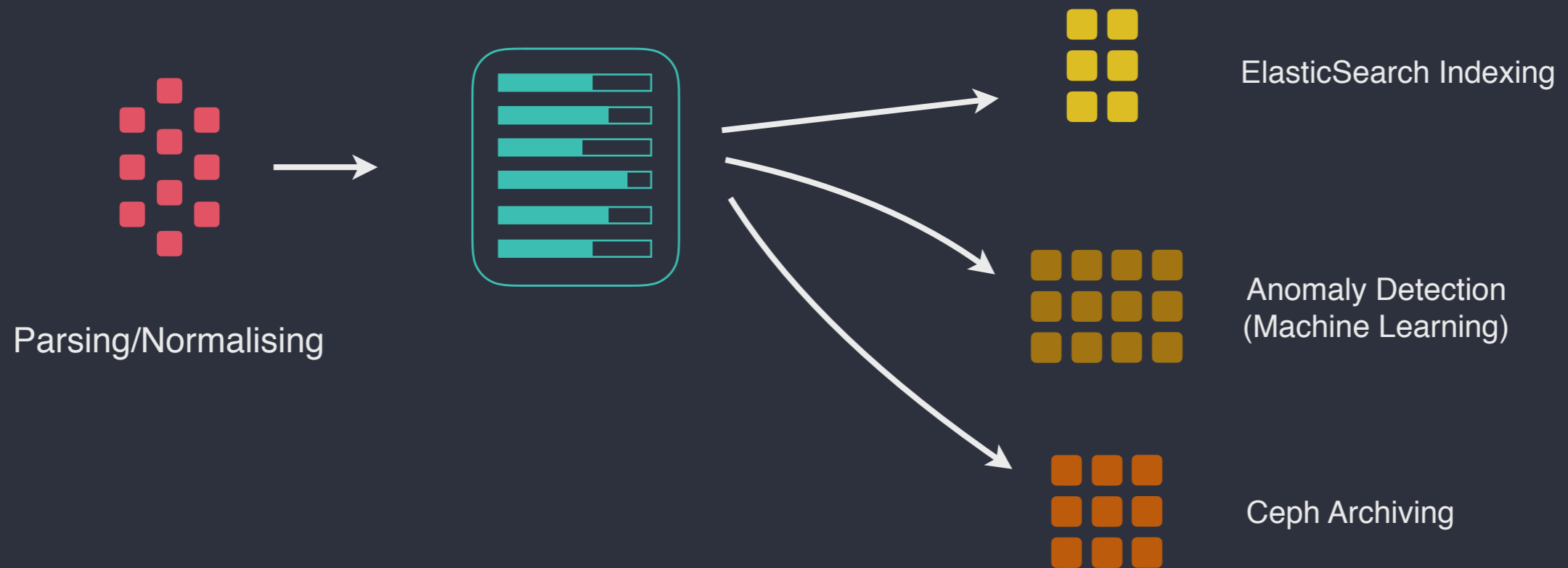




## Kafka Role

Besides saving the data in transit, Kafka is also used to consume the *same data* in different functional pipes. This is what makes the PunchPlatform fulfil the needs of arbitrary uses cases.

Here is a CyberSecurity Platform example.

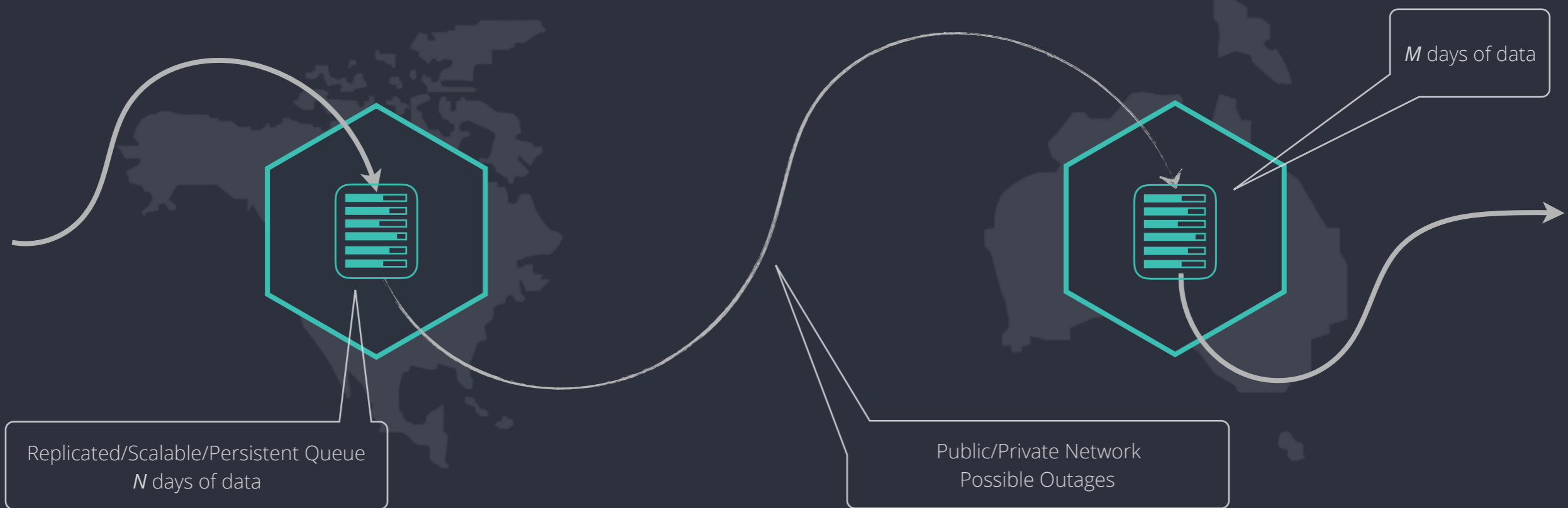




# Kafka Role



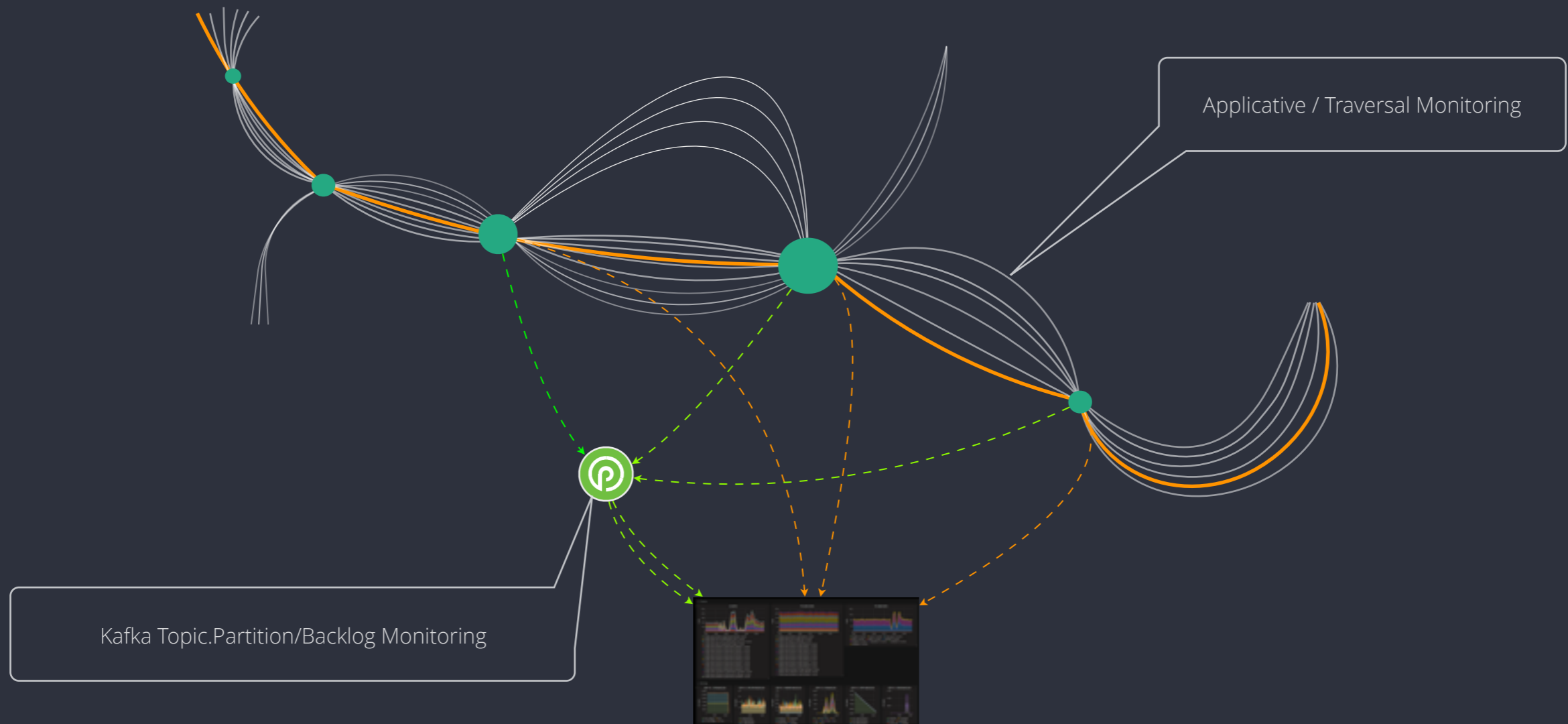
Kafka to Kafka setups make it easy to transfer massive amount of data from one site to another.





# Best Practices

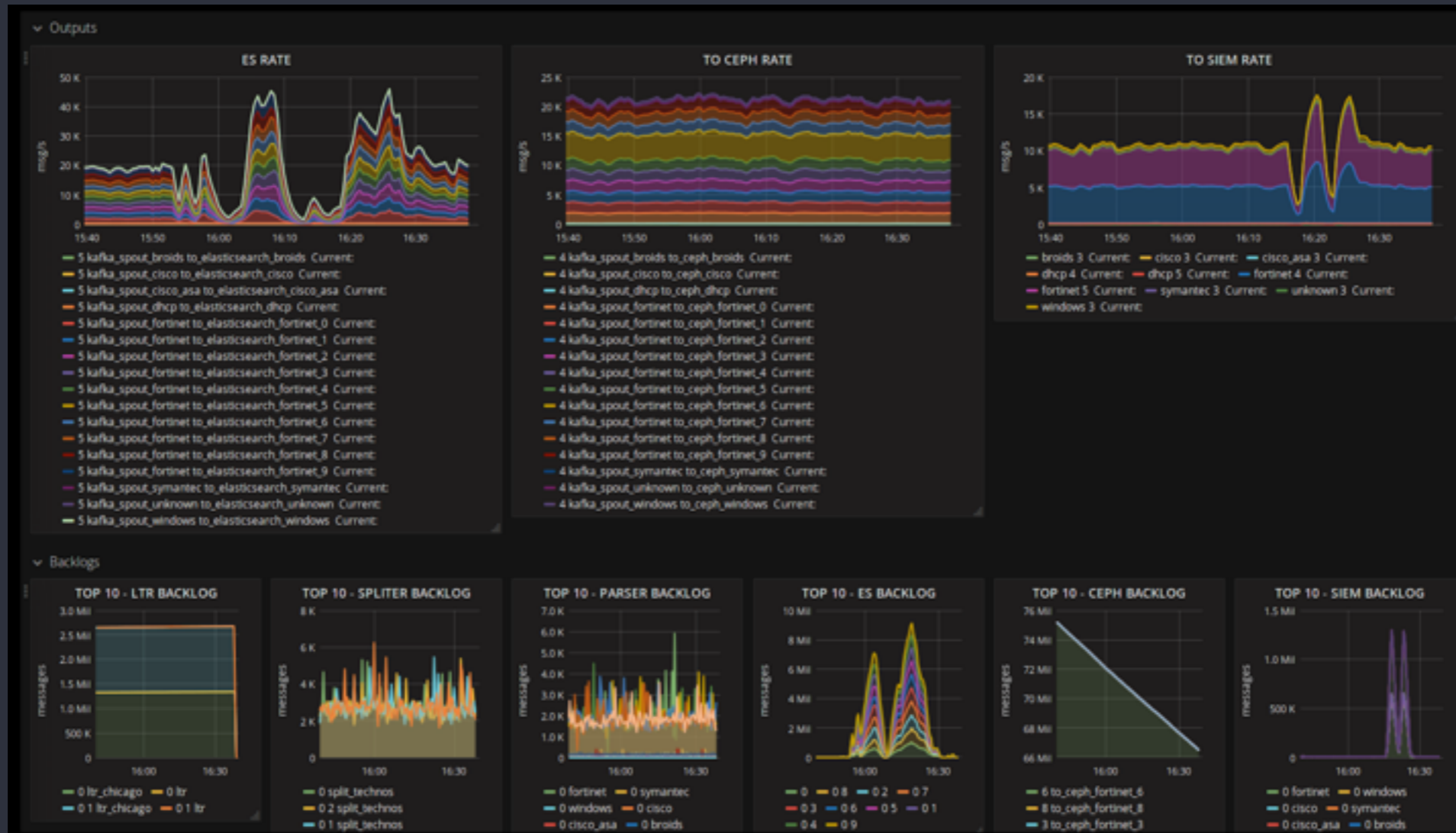
Architect your Real Time Monitoring Solution





# Best Practices

## Monitoring monitoring monitoring ...





Thanks !