Harmonyland Project

Bastien Pouëssel, Antoine Boutin, Tom Genlis

Preliminary questions

- 1/ What technical/business constraints should the data storage component of the program architecture meet to fulfill the requirement described by the customer in paragraph «Statistics» ? So what kind of component(s) (listed in the lecture) will the architecture need?

- **Persistence capabilities:** Long duration storage

- **Scalability:** Scale horizontally ("daily report will weight 200Gb")

- The architecture will require **Available** and **Partition tolerant** (AP) Storage

  - **Column oriented databases** (for example Cassandra)

    - Easier add and remove and allow column optimisation/compression

Preliminary questions

- **2/ What business constraint should the architecture need to fulfill the requirement describe in the paragraph «Alert»? Which component to choose?**

- Business constraints:

  - **Low latency** and **real time alert** to Peacemakers

    Usage of a **stream** (for example Kafka)

    (Reporting in the future usage of batch processing)

Preliminary questions

- **3/ What mistake(s) from Harmonystate can explain the failed attempt?**

- Harmonystate created a team of **data-scientists only.**

- They didn't hire any **data engineer** who could have be able to set an efficient scalable program

- It has led to difficulties in term of infrastructure and scalability which made the PoC fail

Preliminary questions

- 4/ Harmonystate has likely forgotten some technical information in the report sent by the drone. In the future, this information could help Harmonystate make its harmonywatchers much more efficient. Which information?

- **Timestamp**

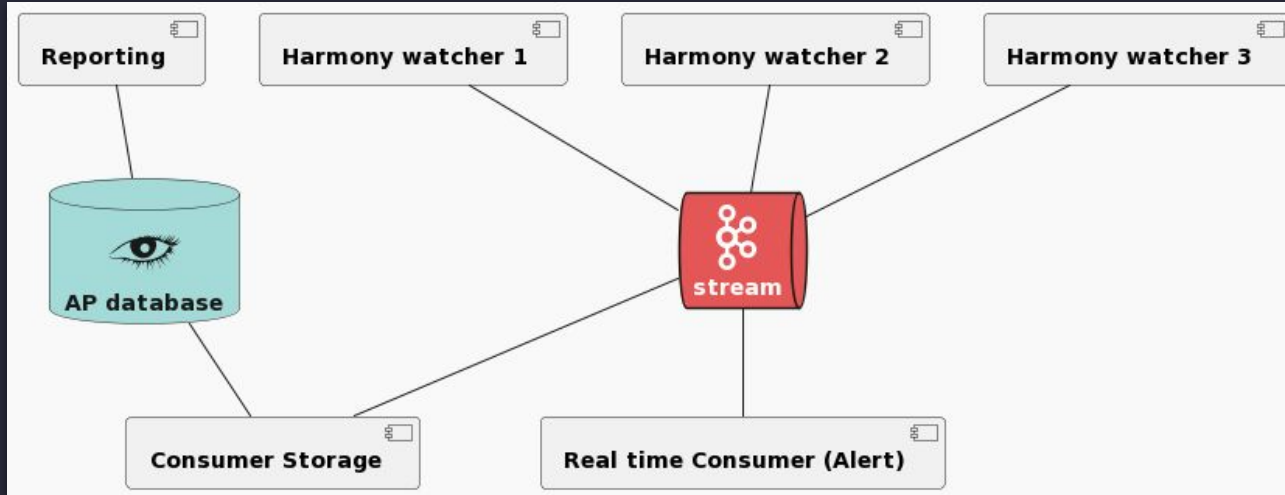 allows to make periodic reports and conduct analysis on time series

  produce any kind of monitoring

  failure handling

Architecture

That's all folks !