

# OMNIA

Distributed & Reactive  
platform for data management

*Presented by Daniel Ortiz Esquivel*

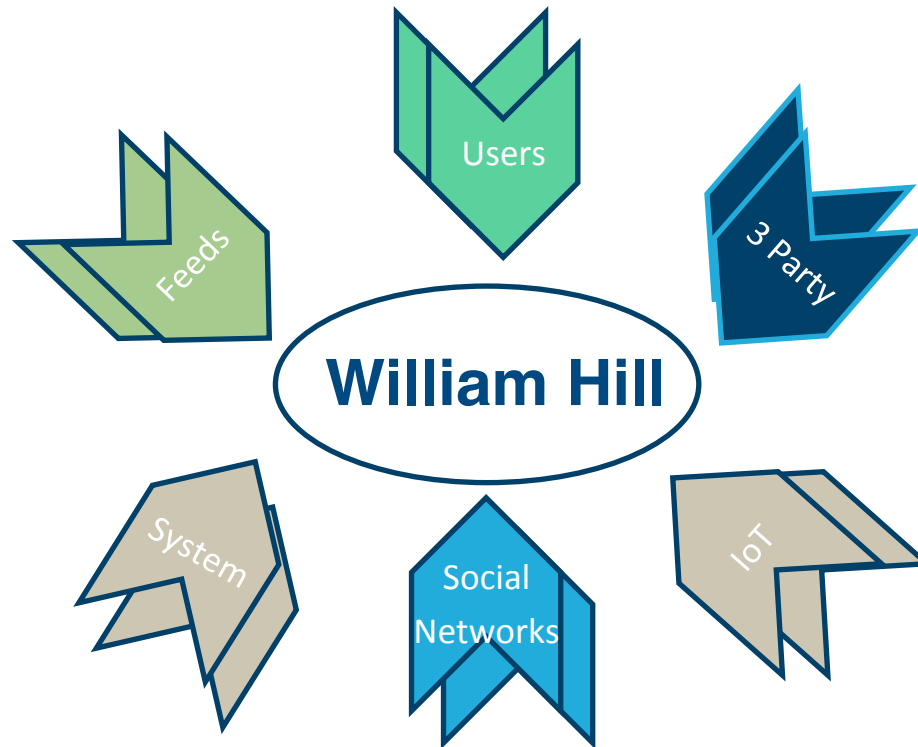
*R&D Software Engineer*

*YAPC::Europe September 2015*

Site: <https://developer.williamhill.com>

Twitter: <https://twitter.com/dortizesquivel>

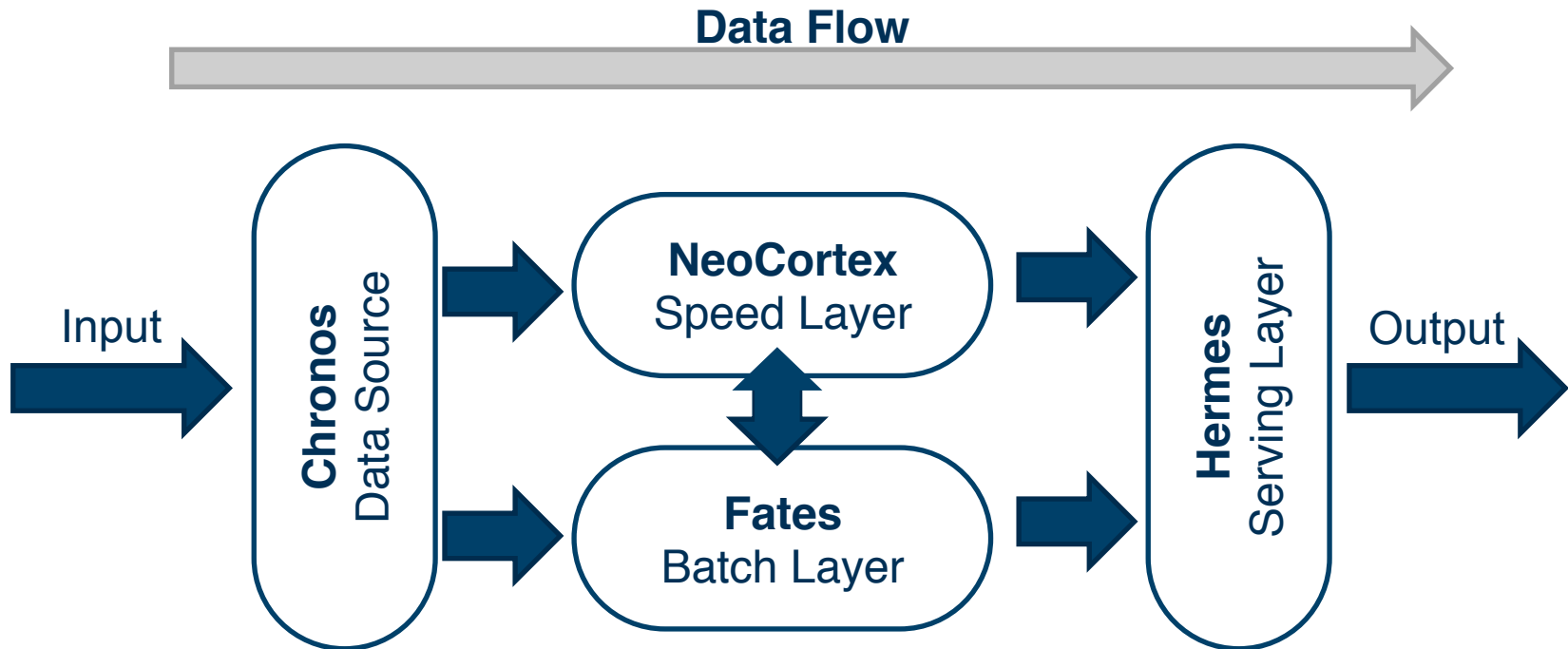
# Motivations



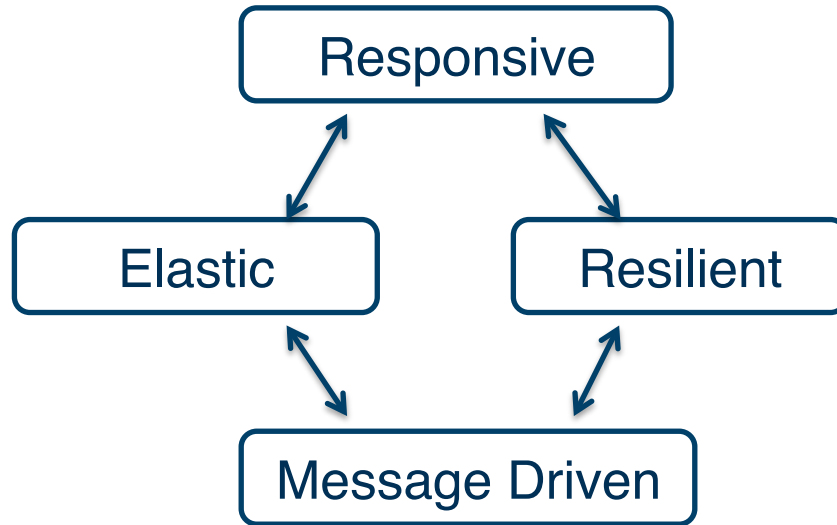
In order to be in a position to innovate we need to control and understand our data

# What is Omnia?

DMP based on the *Lambda architecture* and the *Reactive principles*



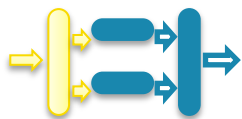
# Reactive principles



The Reactive Manifesto  
<http://www.reactivemanifesto.org/>

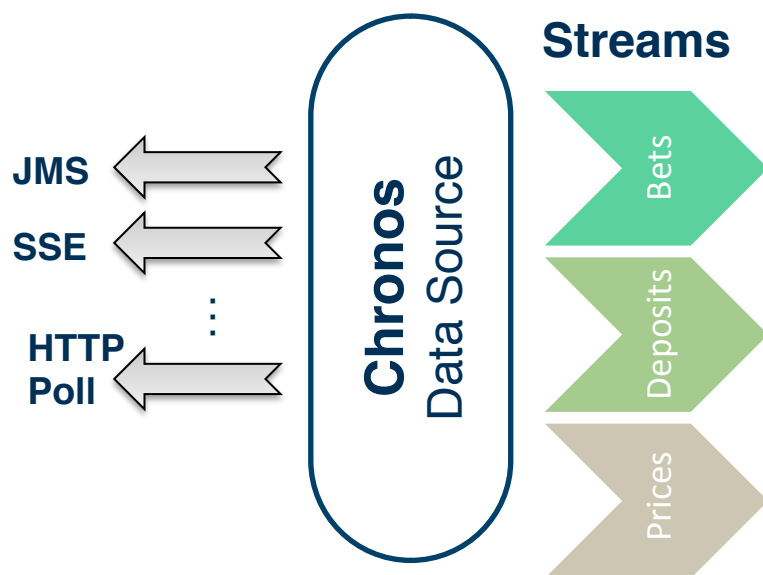
## Resources:

- Reactive Extension for Perl:  
<https://github.com/eilara/Rx.pl>
- Carl Mäsak YAPC:2013:  
<https://www.youtube.com/watch?v=EnIYURSKgk0>



# Chronos: *Data acquisition*

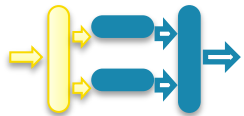
*Chronos* is a reliable and scalable component which collect data from different source and organize them into **Streams of observable incidents**.



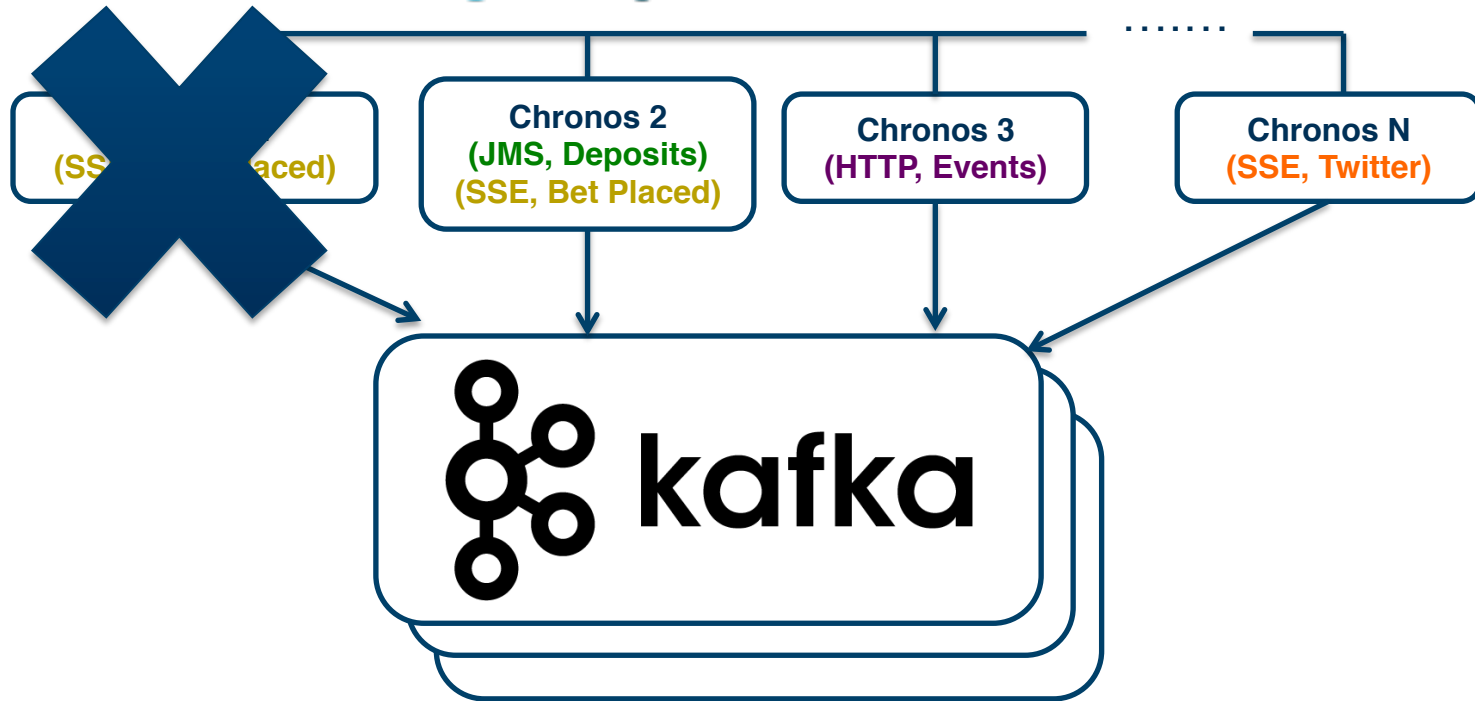
```
Incident: {  
  type: "bet",  
  version: "1.0",  
  time: "2015-09-03 06:00:10",  
  acquisitionTime: "2015-09-03 06:00:06",  
  source: "BetSystem",  
  payload: {... Any valid JSON}  
}
```

**Stream** = Adapter + Converter + Persistence

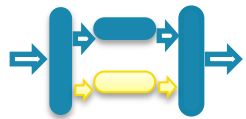




# Chronos: *Data acquisition*

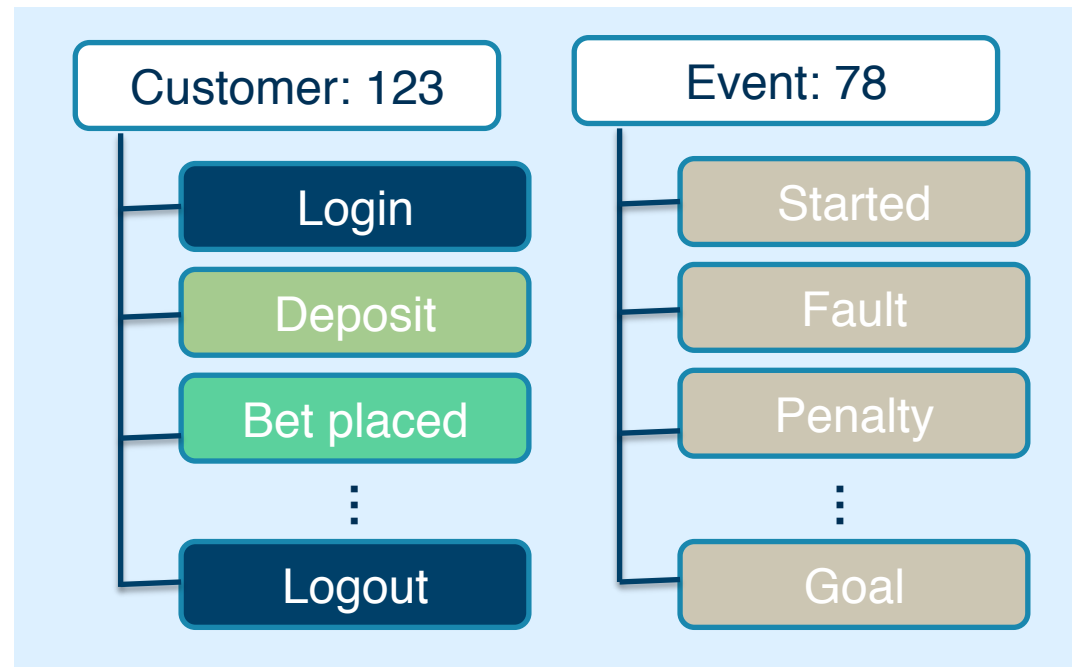
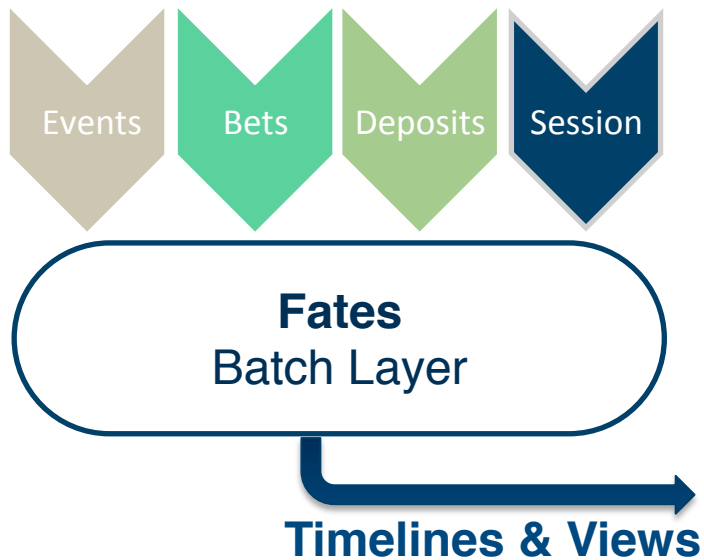


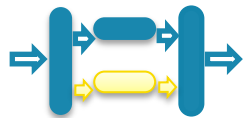
Perl Consumer & Producer Module: <http://search.cpan.org/~sgladkov/Kafka/lib/Kafka.pm>



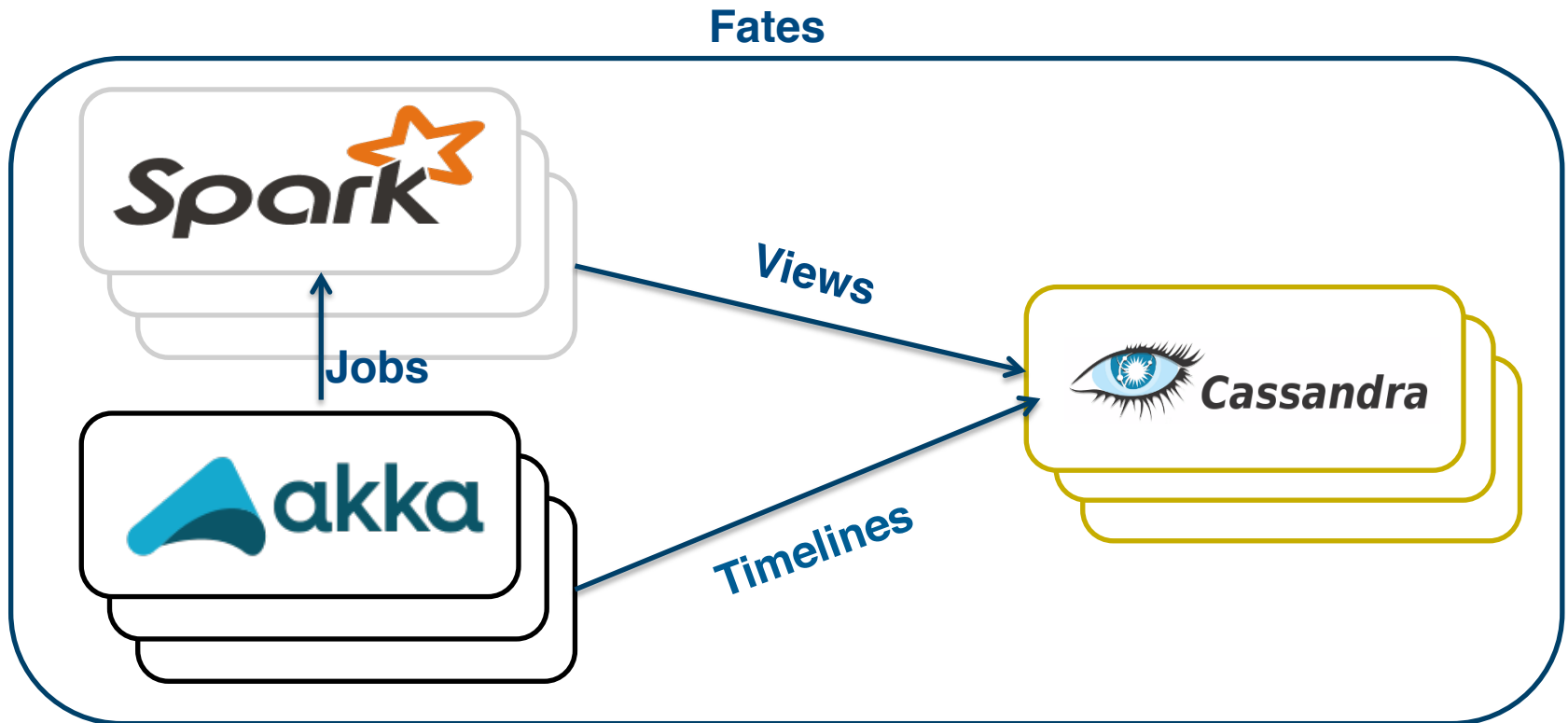
# Fates: *Batch layer*

Fates represents the long term memory of *Omnia*. It organizes the incidents that *Chronos* collected into **timelines** and also elaborates new information as **views** by using machine learning, logical reasoning and time series analysis.



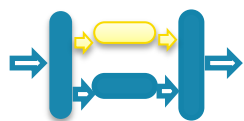


# Fates: *Batch layer*



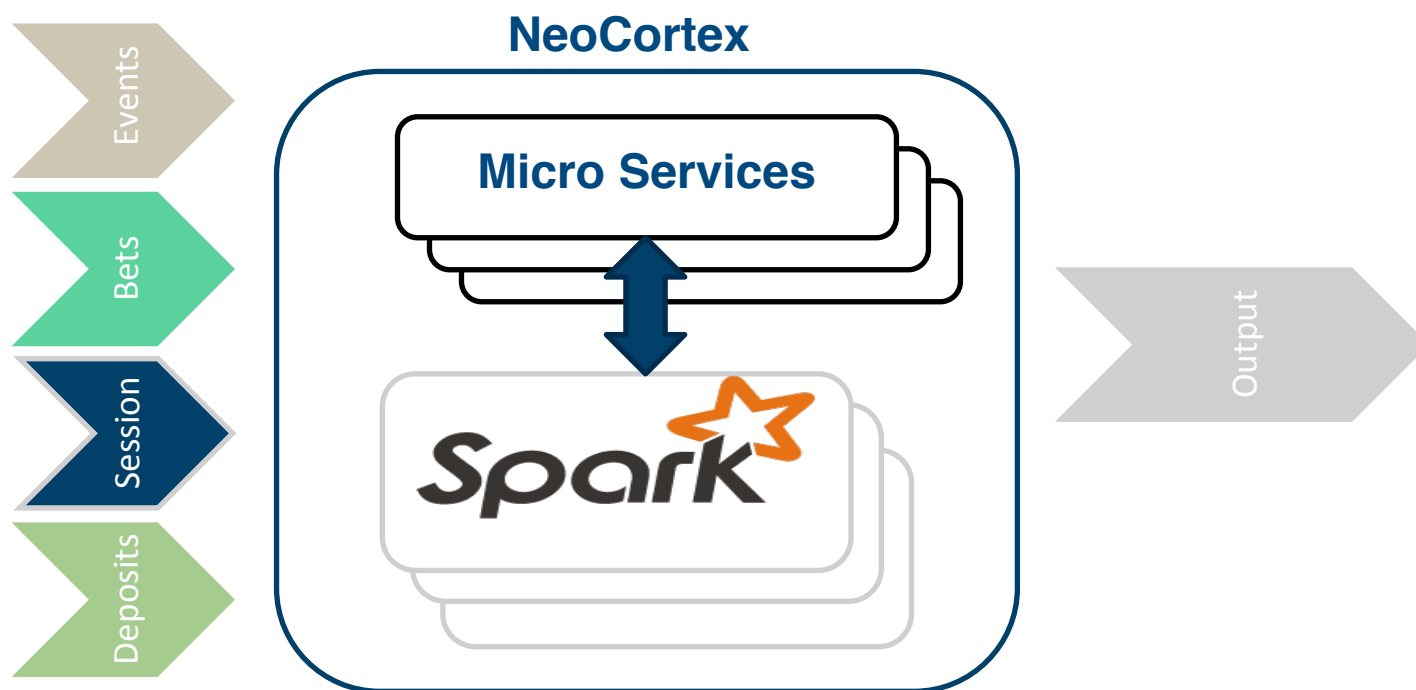
Perl Module for Cassandra: <http://search.cpan.org/~mkjellman/perlcassa-0.041/lib/perlcassa.pm>

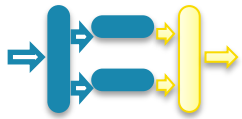




# NeoCortex: *Speed layer*

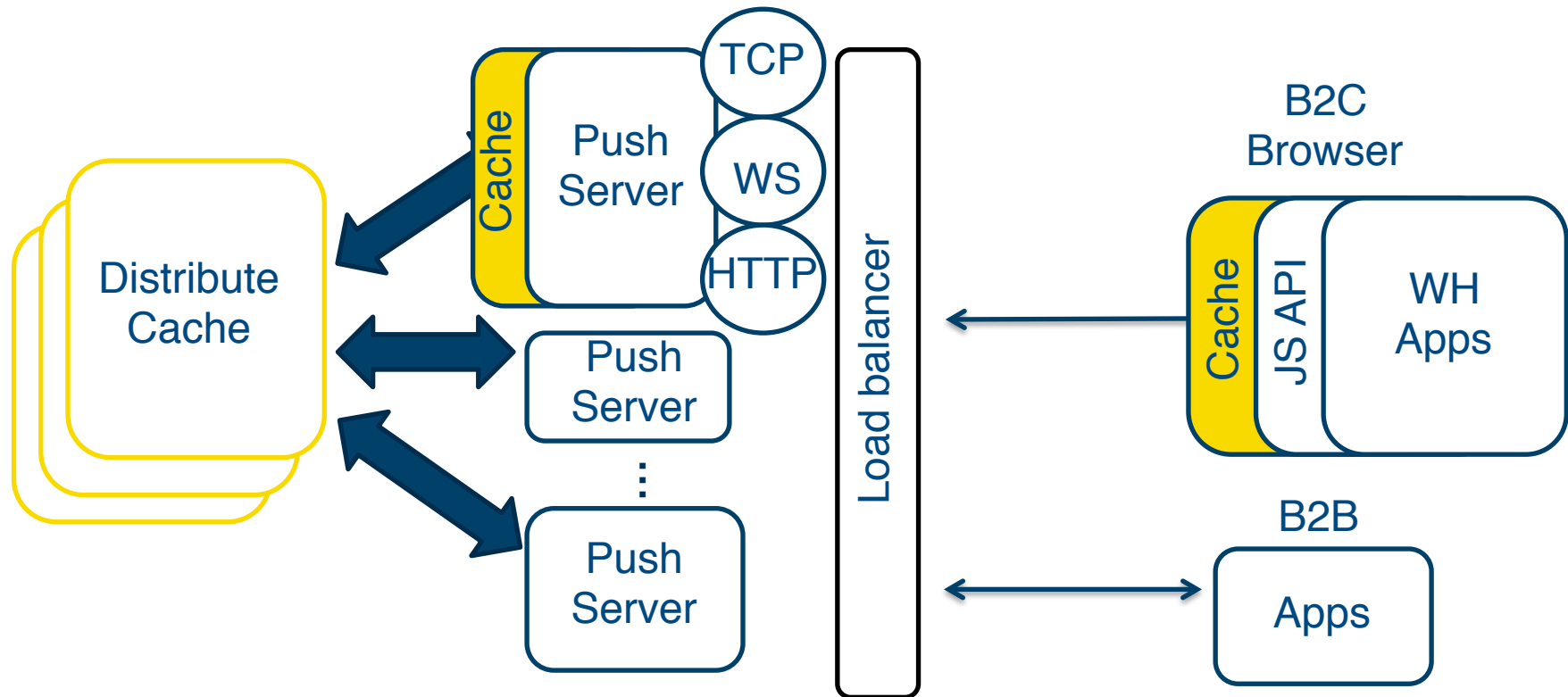
*NeoCortex* represents the short term memory of *Omnia*. It offers a framework to develop **micro services** on top of *Apache Spark*. It performs fast and real time data processing with the data acquired by *Chronos*.



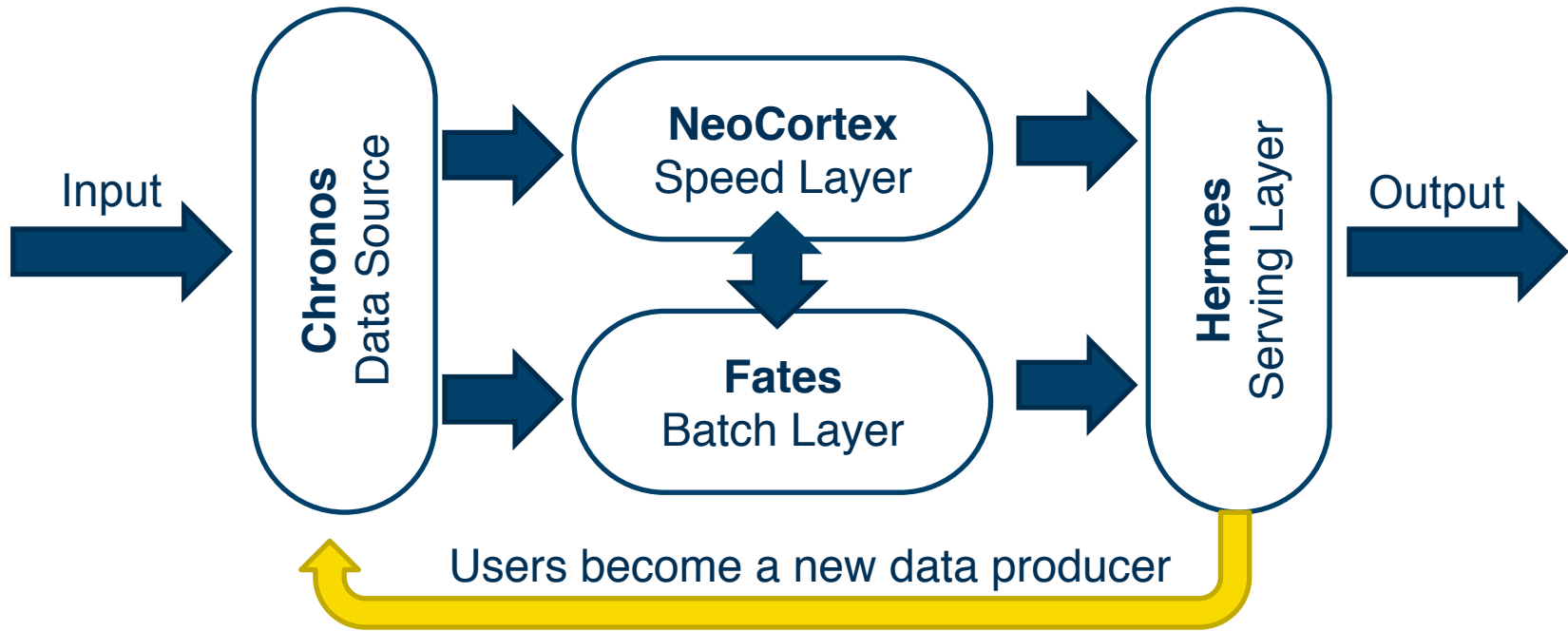


# Hermes: *Serving Layer*

Hermes is a **scalable** and **full duplex** communication for *B2C* and *B2B*.

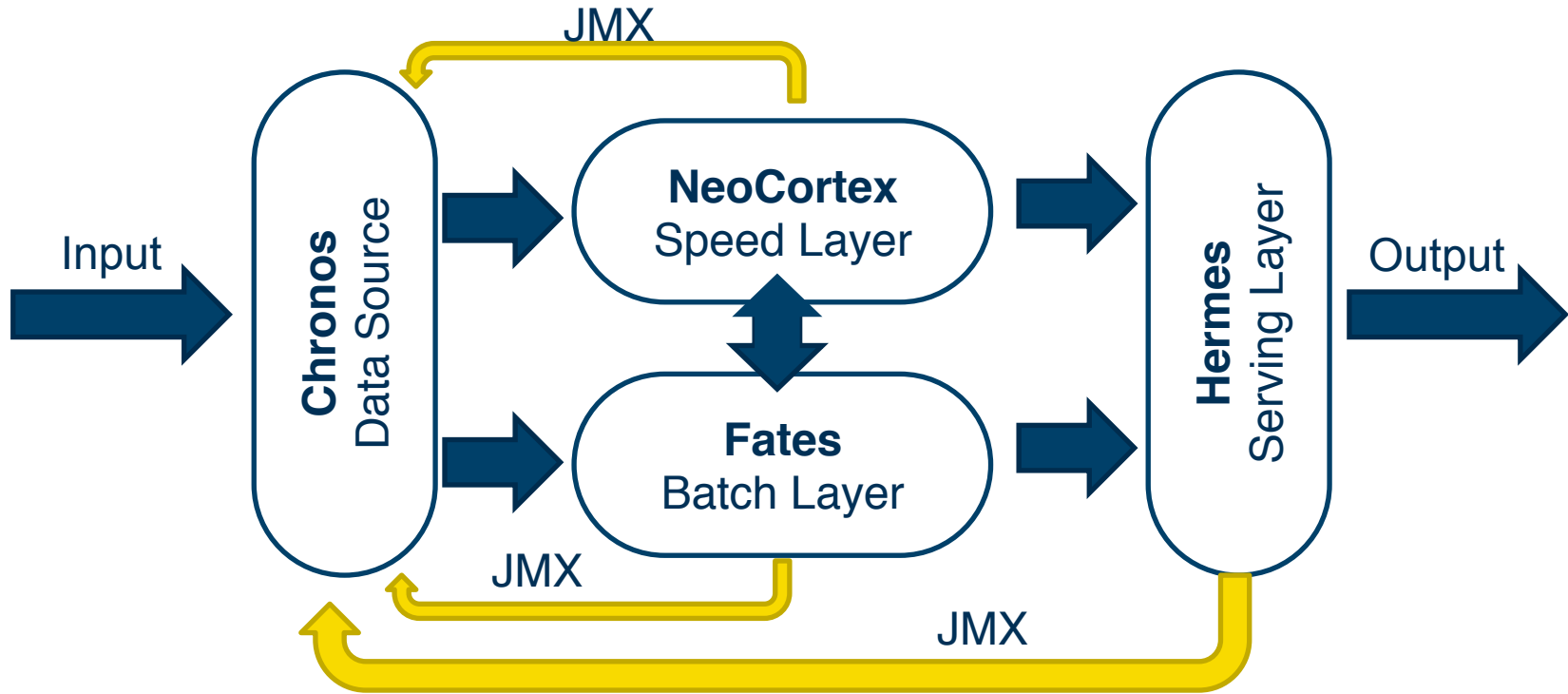


# Omnia Inception



Custom advert, bonus, data load prediction, bot detection...

# Omnia inception +1



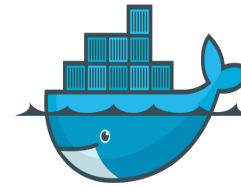
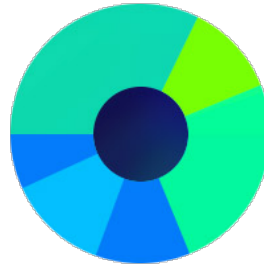
Real time monitoring and elasticity

Docker and Mesos: Scale In&Out based on demand

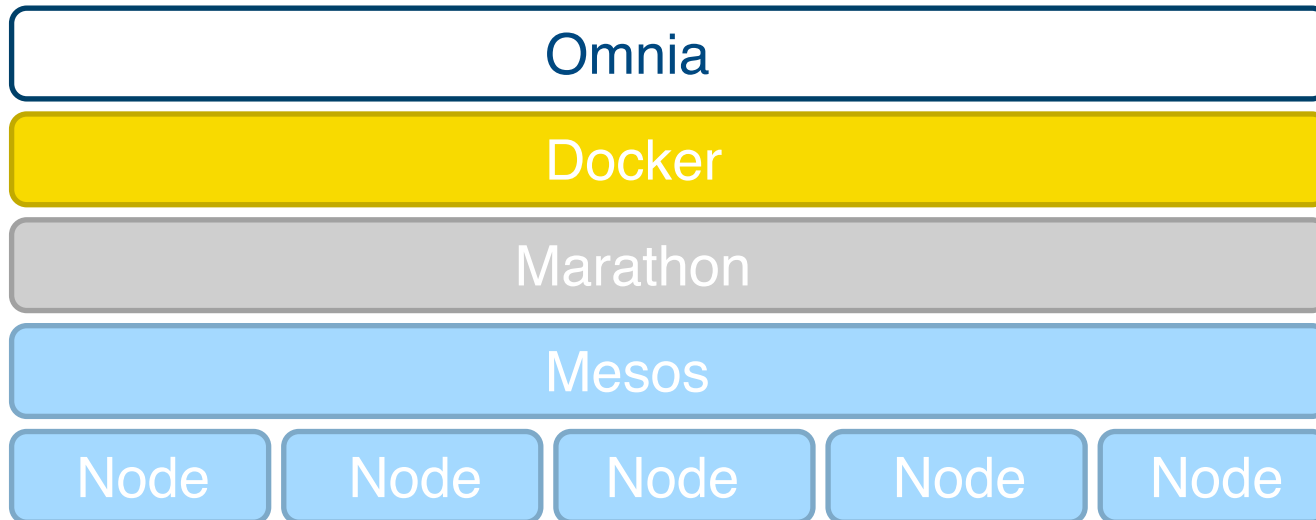
# Omnia infrastructure



MESOS



docker



# Questions



<https://careers.williamhillplc.com/>

Omnia will be released as an open source project

