



Speeding up development with Apache Beam

By Constantin Scacun, Alexander Falca



BEAM
SUMMIT

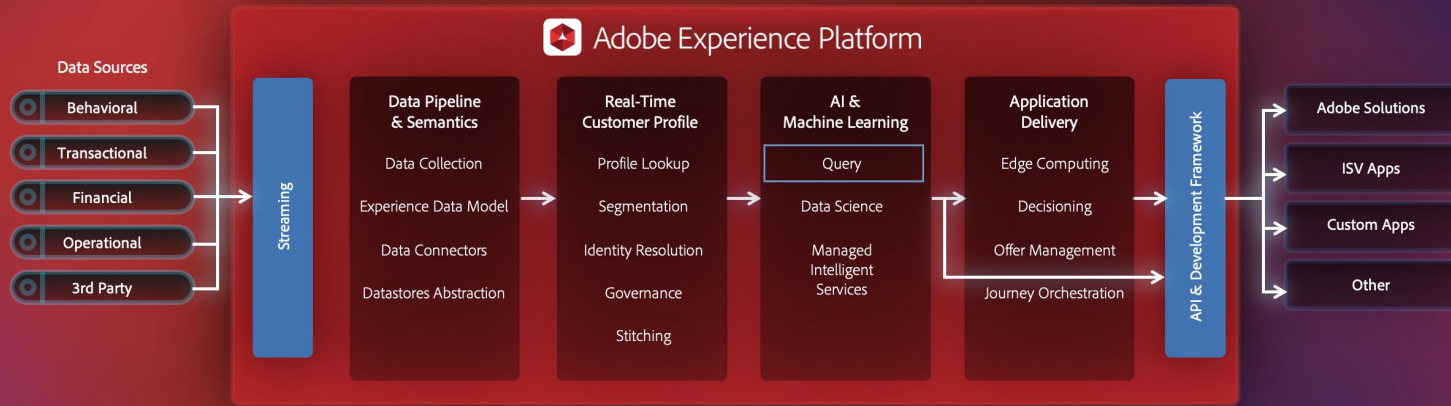
Austin, 2022



Adobe Experience Platform



Why Adobe Experience Platform?



An open, intelligent, and real-time platform for customer experience

Adobe Experience Platform Pipeline



Pipeline Service Locations



13 REGIONS



25 PROD
44 CLUSTERS
OVERALL



310 B MSG/DAY
300 TB/DAY IN
900 TB/DAY OUT

PROD ENVIRONMENT
COLOR KEY
HARDWARE
AZURE
AWS



How Teams Use the Data

- Filter messages (clients might not be subscribed to certain products)
- Validate messages against a schema
- Map/Transform messages
- Export the data out of Kafka
- Custom business logic

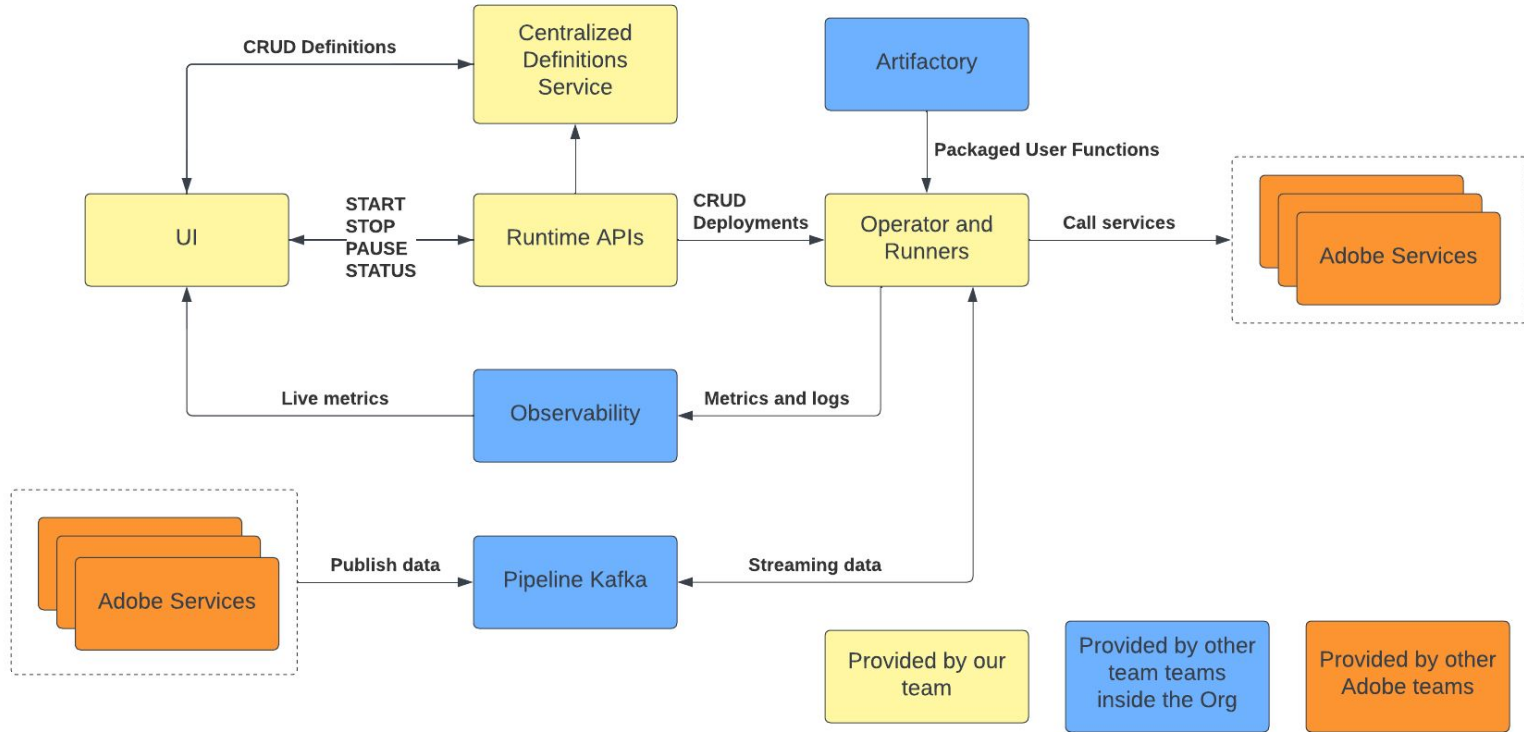
(Most use cases do not even need windowing or aggregations)



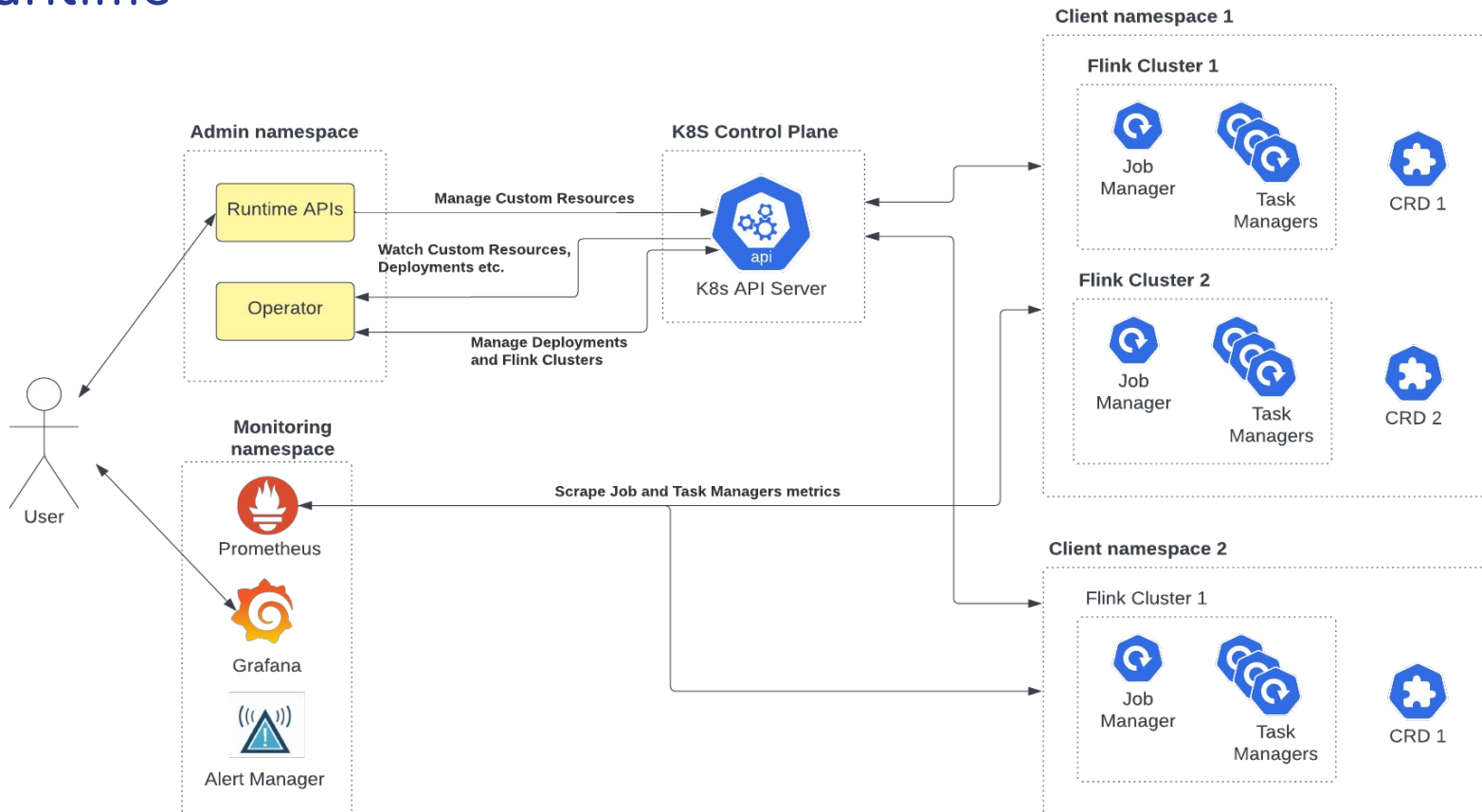
Boiler Plate Code

- Runtime/Deployment
- Observability
- Resiliency
- Fault Tolerance
- Optimizing producers, consumers
- Fine-tuning Apache Beam and Apache Flink
- Scaling on traffic bursts

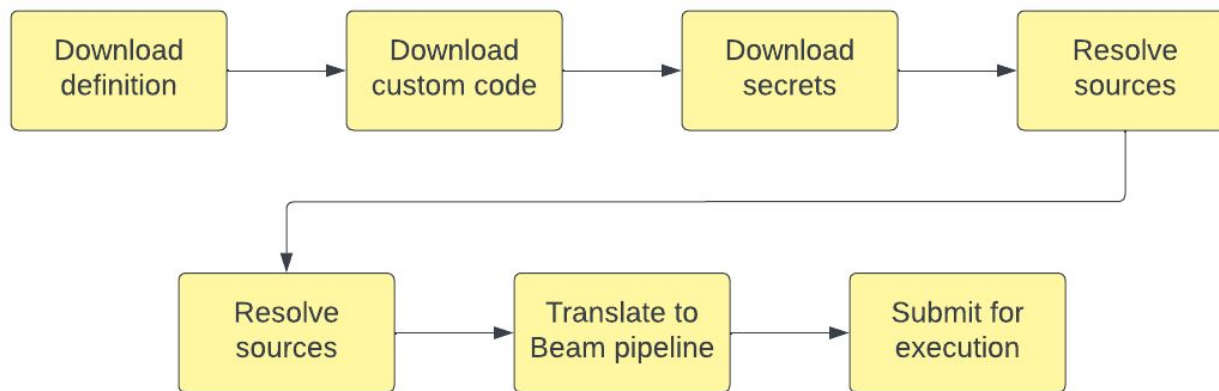
Architecture



Runtime



Execution



Questions?

<https://www.linkedin.com/in/cscacun/>
<https://www.linkedin.com/in/alexfalca/>

<https://medium.com/adobetech/stream-processing-at-scale-within-adobe-experience-platform-909ed502da71>



BEAM
SUMMIT

Austin, 2022