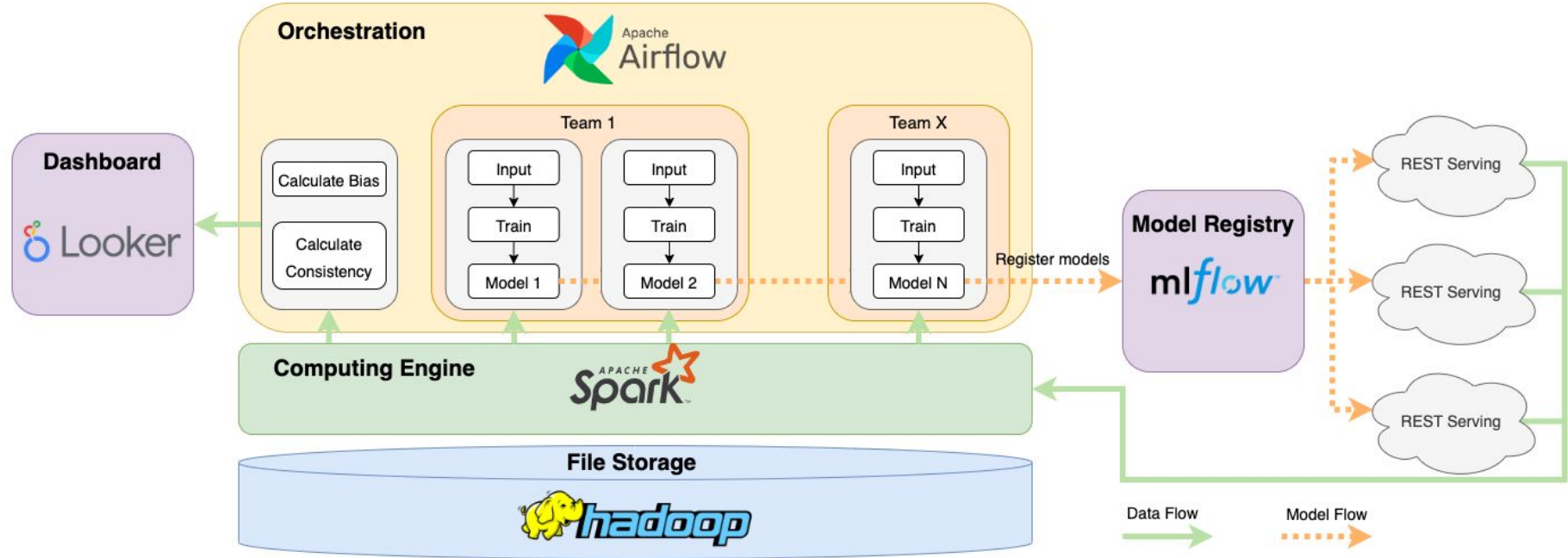


Central Deployment & Serving Platform - Design



Central Deployment & Serving Platform - Discussion

PROS

- Data can be inserted into data warehouse by centralized **ETL processes**;
- Data are **structured** with central management of **schemas**;
- Data lineage can be managed by centralized tooling;
- Data can be shared across teams;
- Features can be calculated and saved in centralized **feature store**;
- Feature **consistency** can be monitored centrally between offline and online stages;

CONS

- The **scheduler** of orchestration frameworks becomes the bottleneck to **scale up** training programs;
- Each team has **less freedom** to experiment with its desired data, as the ownership of data belongs to a central data team;
- Each team has **less exposure** to semi-structured data, unstructured data or **real-time** data;
- Data **transformation** on a daily basis is tangled with 1) **business logic**, and 2) data **schema**, as they all evolve on different rhythms of **versions**;