HOPSWORKS
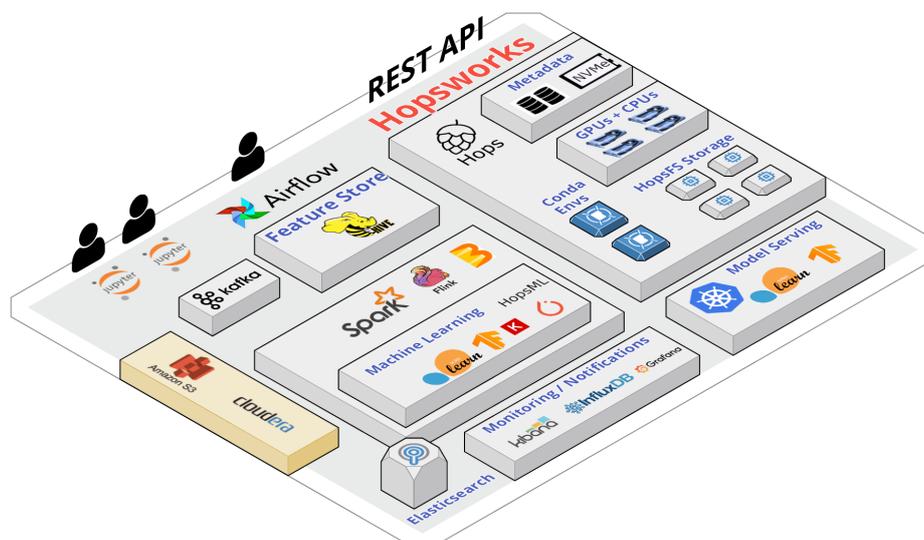
Hopsworks is a full-stack open-source platform for Data Science built around Hops, a next generation filesystem and GPU-aware resource scheduler, and the Feature Store, a data warehouse for your ML features. You can design end-to-end ML pipelines in Jupyter notebooks and run them directly in production ML pipelines. Hopsworks has unique support for teams with project-based multi-tenancy.

MULTI-TENANCY

Hopsworks provides Projects as a privacy-by-design sandbox for data, including sensitive data, and for managing collaborating teams - GitHub repositories for data.

Hopsworks allows Datasets to be shared between projects without copying - like sharing folders in Dropbox. Each Project has its own Conda environment, enabling developers to add/remove Python dependencies using PIP or Conda, without the need to write Dockerfiles.

DEVELOP AND OPERATE ML

Hopsworks is a Python-First platform for Data Science. You can develop end-to-end machine learning workflows fully in Python -- using PySpark for ETL, TensorFlow/PyTorch/SkLearn for training, and Airflow to orchestrate ML workflows. You can even run Jupyter notebooks as stages in pipelines. There is API support for AutoML with parallelized hyperparameter optimization over many GPUs and distributed training. There is provenance, logging and visualization support with Tensorboard, Kibana, and Grafana. Finally, models can be deployed in Kubernetes and monitored in production using Kafka/Spark-Streaming.

FEATURE STORE

Hopsworks' Feature Store is a service that helps compute, store, and manage features for both training and serving machine learning models. Hopsworks' Feature Store allows you to:
• Reuse Feature code/data between projects and teams;
• Control access to Features;
• Cache Feature Data to speed up creating training data in different file formats (.tfrecords, .csv, .numpy, etc);
• Perform feature data validation in either the UI using our API;
• Time travel queries to lookup value of features at times in the past (used to generate new training data).