Cloud Approach to Machine Learning at the National Renewable Energy Laboratory

David T. Rager
Chief Cloud Technologist
National Renewable Energy Laboratory
June 1, 2012
1. Introduction
2. Cloud Overview
3. ML Focus for Data Science
4. Stratus Environment @ NREL
5. Distributed Workflow Support
6. Summary
Introduction

My background:

Developer / Systems and Data Architect for 25 years
Also a Cloud Solutions Architect for last 12 years

Specialization in cloud does overlap heavily with operations. It is a challenge to separate cloud development support from dev-ops – API access to all functionality blurs the responsibilities/roles.

Approach to cloud:
• A resource for developers and analysts at the lab to solve problems in an agile fashion
• Encourage use of cloud provider managed application and analysis support services to avoid operational overhead, reduce complexity for maintenance and leverage cloud security
• Leverage extensibility of cloud to “plug in” machine learning where cloud supports goals of a project

Multi-cloud tools are hard to find that make typical interactions seamless between environments. API and SDKs are highly opinionated and typically require user specialization to leverage effectively.

Once you embrace the SDK/APIs available cloud computing is empowering.

At NREL we operate the Stratus environment - which encompasses AWS, GCP and Azure to support research.
Cloud Overview: Adoption

Kaggle: State of Machine Learning and Data Science 2021

Survey of data scientists

Credit: https://www.kaggle.com/c/kaggle-survey-2021
Cloud Overview: Product Usage

Kaggle 2020 Survey: Usage of Cloud ML Products by Occupation

Credit: https://medium.com/sbc-group-blog/cloud-ml-in-perspective-surprises-of-2021-projections-for-2022-d0b3b821fa77
Growth is tied to ease of use, but it also leans into scenarios like:

- Events
- Schedules
- Scale
- Pipelines

Look at Netflix pyramid – IaaS – to SaaS

Model Development
Feature Engineering
Model Operations
Versioning
Architecture
Job Scheduler
Compute Resources
Data Warehouse

Credit: https://netflixtechblog.com/open-sourcing-metaflow-a-human-centric-framework-for-data-science-fa72e04a5d9
Stratus Environment

Big Data Analytics
• data warehousing
• data management tools

Containerized Applications
• multiple scheduling systems
• Docker containers at the edge
• Docker serverless functions

Growth In
• IoT support for field experiments
• grid management studies

Ongoing Support for
• data processing / ML workflows
• public web applications / ML inference endpoints
• publishing of large open data sets
Distributed Workflows

**Edge**
- Data Collection
- Cloud (Stratus)
  - Reliable
  - API / SDK Layer
  - On-demand – support for events / schedules / scale

**HPC**
- Data Modeling
- ML Model Creation / Training

Credit: https://link.springer.com/chapter/10.1007/978-3-030-96498-6_13
Example: Use weather data to predict future energy costs.

Integrate AWS Forecast service

OR

A specialize model not provided by AWS using:
• SageMaker Platform
• Dockerized Model

These options allow us to use specialized hardware to optimize the job, which can easily change over time.
Summary

Cloud Computing generally supports our goal of commoditizing reliable machine learning at scale:
• Highly available services
• API controls for operational use of developed models
• Scalable compute for large training jobs or heavy consumption of inference
• Ability to manage and store very large datasets
• Reliable / distributed event system and scheduler
• On-demand model
• FedRAMP security certifications

Challenges:
• Learning curve for using managed services / cloud APIs and SDKs
• Data scientists usually do not leverage cloud platforms, so models are deployed without the benefits cloud can provide when starting with their integrated platforms
"Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program."

The Networking and Information Technology Research and Development (NITRD) Program

**Mailing Address:** NCO/NITRD, 2415 Eisenhower Avenue, Alexandria, VA 22314

**Physical Address:** 490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024, USA Tel: 202-459-9674, Fax: 202-459-9673, Email: nco@nitrd.gov, Website: [https://www.nitrd.gov](https://www.nitrd.gov)