AI/ML with OpenShift

Turning your AI/ML lifecycle into standard software development practices with containers and Kubernetes

Tushar Katarki Al/ML on OpenShift Product Manager Red Hat Robert Sedor Chief Architect Red Hat



Agenda

- Level Setting on AI/ML
- Data Science: Models and Jupyter
- OpenShift for Data Science
- Customer Success Stories
- How to move forward



Slides and recording will be be posted below - maye take a couple of days

https://red.ht/LiveAI-ML

Raffle - Book on ML

(will be announced at the end of the Webinar!)



https://www.amazon.com/Building-Machine-Learning-Pipelines-Autom ating/dp/1492053198



Level setting



The intelligent application journey





Intelligent applications are distributed systems



(Adapted from Sculley et al., "Hidden Technical Debt in Machine Learning Systems." NIPS 2015)



The ML lifecycle and responsibilities





The ML lifecycle in technical detail



8



You have seen this before

It's just software development





What is DevOps?

Where can I buy it?





One platform for all the domains





Data science: Models and Jupyter



The data scientist's domain





Data Science is iterative





Jupyter(Hub/Notebooks) in data science





Everyone does the same thing... differently







Two primary bottlenecks



- Getting started
- Getting into production



The two bottlenecks

The data scientist's dilemma

Getting Started:

- Self service
- Repeatable/shareable environments
- Repeatable/shareable experiments
- Access to specialty resources (GPU, etc)

Getting into production:

- Self service
- Ease of instrumentation
- Access to specialty resources (GPU, etc)
- (Automated) scale-out



OpenShift for data science: Getting started



The ML lifecycle and responsibilities





Containers for models

Reproducible and shareable environments for building, training and serving





Jupyter-Notebook-as-a-Service



- Identity and RBAC
- Project isolation (data, networks)
- Sharing of resources with protection
- Resource isolation
- Resource quotas
- Priority and pre-emption



Open Data Hub: A reference implementation of the possible



Open Data Hub: A community effort





Open Data Hub: Projects and Community



- Monitoring and alerting toolkit
- Used to diagnose problems



- Analytics platform for all metrics
- Query, visualize and alert on metrics



- Deploying machine learning models as micro-services
- Full model lifecycle management



- Unified analytics engine
- Large-scale data access



- Multi-user Jupyter
- Used for data science and research



- Distributed Object Store
- S3 Interface



- Distributed event streaming
- Pub/Sub Messaging



- Container-native workflow engine
- Declaratively deploy ML pipelines and models



Ecosystem Partnerships



Strategic partnerships within AI/ML ecosystem



OperatorHub and certified Operators

- Kubernetes Operators for life deployment and lifecycle management
- Red Hat Marketplace for ISV operators
- OpenShift Operator Certification
- OperatorHub integrated into OpenShift 4

	OperatorHub.io				Q Contribute	2 ~
		0	perato <u>rH</u>	u b		
	Operators deliver the automation advantages of cloud services like provisioning, scaling, and backup/ restore while being able to run anywhere that Kubernetes can run.					
nt	CATEGORIES	48 items			VIEW: 雛 Card 🛩 SORT: A-2	Z ~
	Databases bicovery Key Management Storages	O AWS Cloud Broker	Couchbase Operator	Falco Operator	() influxdb	
	Monitoring PaaS/Container Service Streaming & Messaging Security & Compliance	provided by Amazon Provides native AWS services and simple integration directly within the application	provided by Couchbase Run and manage Couchbase autonomously on Kubernetes.	provided by Sysdig Kubernetes operator for Sysdig Falco that allows developers to manage rules for detecting intruders and backdoors.	provided by InfluxDB The Kubernetes operator for InfluxDB and the TICK stack.	
	OPERATOR MATURITY	*	•			
			Operator Hub Discover Operators from the Kubernetes community and Red Hat partners, curated by Red Hat. Operators can be installed on your clusters to provide optional add-ons and shared services to your developers. Once installed, the capabilities provided by the Operator eppear in the Developer Catalog, providing a self-service experience.			
4	PROVIDER Amadeus (1) Amazon (2) CNCF/Upstream (1) Confluent (1) Cone@S (2)	Home ✓ Projects Status Search Events	Operator Hub Discover Operators from the K optional add-ons and shared s providing a self-service experi	ubernetes community and Red Hat parts ervices to your developers. Once installe ence.	ners, curated by Red Hat. Operators car d, the capabilities provided by the Ope	n be installed on your clusters to provide reator appear in the Developer Catalog,
4	Provinces Anadous (1) Anadous (2) CARCF/Uniteram (1) Confluent (1) Confluent (1) CoreOS (2) Show 3 more	Home × Projects Status Search Catalog ×	Operator Hub Discover Operators from the K optional add-ons and shared s providing a self-service experi All items Kafka Messaging	ubernetes community and Red Hat parts ervices to your developers. Once installe ence. All Items 14 items	ners, curated by Red Hat. Operators care	n be installed on your clusters to provide rator appear in the Developer Catalog.
4	Provinces Amadeus (1) Amadeus (1) ChCF/Upsteam (1) Comfuent (1) Comfuent (1) Show 3 more	Home Projects Status Status Status Status Catalog Catalog Provisioned Services Provisioned Services Operator Management Broker Management	Operator Hub Discover Operators from the K optional add-ons and shared a providing a self-service experi All items Kafha Messaging Middleware Streaming Datatore Openshift Optional Logging Aggregated	ubernetes community and Red Hat part ervices to your developers. Once installe All Items 14 items ang streams provided by Red Hat, Inc. Red Hat AMO Streams is a massively scalable.	ers, curated by Red Hat. Operators car d, the capabilities provided by the Ope atomationbroker provided by Red Hat, Inc. Ansible Service Broker is an implementation of the (Open	n be installed on your clusters to provide rator appear in the Developer Catalog.
4	Provuese Anadous (1) Anadous (1) CACFUptoream (1) CACFUptoream (1) Cacrofount (1) Cacro 5 (2) Show 3 more	Home Projects Status Search Catalog Providence Catalog Provisioned Services Provisioned Services Deparators Provisioned Services U Workloads	Operator Hub Discover Operators from the K optional add-ons and shared a providing a self-service experi All items Kafia Messaging Middleware Streaming Datastore Openshift Optional Logging Aggregated Elasticsearch Kibma	ubernetes community and Red Hat parts ervices to your developers. Once installe All Items 14 items encode ang-streams provided by Red Hat, Inc. Red Hat, AMO Streams is a massively scalable, distributed, and high © Installed	A the capabilities provided by the Operators car d, the capabilities provided by the Operators automationbroker provided by Rich Hat, Inc. Ansible Service Broker is an implementation of the (Open Service Broker AP)] (https://github.com/opera	n be installed on your clusters to provide rator appear in the Developer Catalog.
4	Provinces Amadeus (1) Amadeus (1) ChCF/Upstream (1) Confluent (1) Confluent (1) Stow 3 more	Home Projects Status Status Status Status Catalog Catalog Peveloper Catalog Installed Operators Provisioned Services Operator Hubi Broker Management Broker Management Workloads Networking	Operator Hub Discover Operators from the K optional add-ons and shared a providing a self-service experi All tems Kafka Messaging Middleware Streaming Datastore Openshift Optional Logging Aggregated Elasticsearch Kibana Fluentd Efk	ubernetes community and Red Hat part ervices to your developers. Once installe nence. All Items 14 items many streams provided by Red Hat, Inc. Red Hat AMO Streams is a massive/scalable, distributed, and high © Installed	ers, curated by Red Hat. Operators car d, the capabilities provided by the Ope automationbroker provided by Red Hat, Inc. Ansibio Service Broker is an implementation of the (Open Service Broker API) (https://github.com/open	In be installed on your clusters to provide rator appear in the Developer Cataloo.
4	Provues Anadous (1) CAREF/Unsteam (1) Confluent (1) Confluent (1) Show 3 more	Home Projects Status Sexrch Cotalog Developer C	Operator Hub Discover Operators from the K optional add-ons and shared a providing, a self-service experi All items Karha Messaging Middleware Streaming Datastore Openshift Optional Logging Aggregated Elasticserch Kibana Fluend Erk Tracing Monitoring	ubernetes community and Red Hat parts ervices to your developers. Once installe ence 14 Items 14 Items	ers, curated by Red Hat. Operators car d, the capabilities provided by the Ope end of the Capabilities provided by the Ope second of the Capability of the Capability provided by Red Hat, Inc. Ansible Service Broker is an implementation of the Open Service Broker API (https://gthub.com/open:	In be installed on your clusters to provide rator appear in the Developer Catalog.
4	Provuese Anadous (1) CANEFU Apparam (1) CANEFU Apparam (1) Care (05 (2) Show 3 more	Home Projects Status Search Catalog Provisioned Services Provisioned Services Provisioned Services Provisioned Services Provisioned Service Storage Builds	Operator Hub Discover Operators from the K optional add-ons and shared a providing aeti-service experi All items Katha Messaging Middleware Streaming Datastore Openshift Optional Logging Aggregated Elasticsearch Kibana Fluentd ER Fluentd ER Tracing Monitoring Troubleshooting Distributed	uberretes community and Red Hat parts ervices to your developers. Once installe ments I terms I terms	ers, curated by Red Hat. Operators car d, the capabilities provided by the Operators automationbroker provided by Red Hat, Inc. Ansible Service Broker is an implementation of the (Open Service Broker AP) (https://github.com/operators (https://github.com/operators) Geocheduer provided by Red Hat, Inc. An operator to run the OpenShift descheduer	In be installed on your clusters to provide rater appear in the Developer Catalog. duster-logging provided by Red Hal, Inc. The Cluster Logging Operator for 0K0 provides a means for configuing and meanging your agregated logging war agregated logging gravided by MongoDB, Inc Install Mul-stack monitoring of [OpenShift (usters]

COMMUNITY OPERATORS OPENSHIFT CERTIFIED OPERATORS



Accelerated AI Platform by HPE, NVIDIA, and Red Hat



Backed by <u>Reference Architectures</u> and enterprise support from HPE, NVIDIA, and Red Hat



29

Hewlett Packard Enterprise



Fast Tracking AI projects with NVIDIA NGC & GPUs, and Red Hat OpenShift



Data Pipelines Scenarios



- Many data sources
- Many data pipelines
 and situations can exist
 Machine Learning
- should be tap into all these data pipelines



Covid19 Assessment Pipeline



https://github.com/guimou/dataprep-code/tree/master/examples/xray



Starburst Enterprise for Presto on Red Hat OpenShift

Federated queries across multiple software-defined data sources



https://www.redhat.com/cms/managed-files/pa-starburst-overview-f25521wg-202010-en.pdf
 https://www.youtube.com/watch?v=v5qeXdyRq18



SQL Server Big Data Cluster – Fintech

OLTP, OLAP, and Fraud Detection



© Copyright Microsoft Corporation. All rights reserved Red Hat OpenShift and SQL Server Big Data Clusters More at: aka.ms/sqlworkshops

Cloudera CDP Private Cloud

2

New set of data analytics applications Featuring use-case optimized interfaces



Running on a container cloud Fast provisioning & scaling, efficient, simple



With access to a shared data lake That is secured and governed

CDP PRIVATE CLOUD ON OPENSHIFT ANALYTIC EXPERIENCES



Source-to-Image for models

Automating build, deployment, APIs and instrumentation





Models as stateless microservices

The power of Kubernetes





MODEL

MODEL

MODEL

Monitoring performance and drift

Keeping track of models over time





Intelligent spam detection - app architecture





0

Predictive Decision Modeling - Decisions driven by AI / ML





Example Scenario TOPOLOGY: EVENT CLOUD DECISIONS



AI/ML success stories







Partner to bring COVID-Net radiography screening AI to hospitals

Red Hat Press Release (11/16/20): <u>DarwinAI and Red Hat Team Up to Bring COVID-Net Radiography Screening AI to</u> <u>Hospitals, Using Underlying Technology from Boston Children's Hospital</u>



Boston Children's Hospital and DarwinAl COVID-Net radiography screening Al

The challenge:

COVID-Net, a suite of deep neural networks for COVID-19 detection and risk assessment developed by DarwinAI and University of Waterloo's Vision and Image Processing Lab, is continuing to evolve at a rapid rate as the understanding of the disease evolves. These tools are rather technical and perhaps not very doctor-friendly in their raw form, and because of that, there has been a growing need to make COVID-Net more readily accessible to front-line clinicians around the world.

The solution:

By working with Boston Children's Hospital and Red Hat, the COVID-Net project has been able to use the BCH's ChRIS platform infrastructure and Red Hat OpenShift to build a container-based workflow to easily deploy the tools into the hands of clinicians on the front lines. Work started early Summer 2020. Within a few days, the COVID-Net suite was fully containerized on the ChRIS platform.







11/16/20 Press Release: <u>DarwinAl and Red Hat Team Up to Bring COVID-Net Radiography Screening Al to Hospitals, Using Underlying</u> <u>Technology from Boston Children's Hospital</u>





RBC Machine Learning - The challenge

AI/ML projects took up to 2 months to get off the ground, slowing research and new offerings



self-service access to GPU resources was manual and slow build and required significant resources

requirements were hard to meet



Source:

46

Borealis blog, July 23, 2020: Borealis Al announces new Al infrastructure developed with Red Hat and NVIDIA.

The solution

Accelerated, more secure and flexible AI/ML infrastructure on Red Hat OpenShift, NVIDIA



 \diamondsuit



ML applications and services deployed in containers

Consistent private cloud platform for bank's AI/ML research center NVIDIA Operator simplifies use of GPUs



Source:

47

The Goal

Digital transformation of BMW Group from Cars to Devices and Data

Provide a superior **connected car experience** for 12.5 million drivers

Develop driving algorithms and analyze data to support **autonomous vehicle project**

Adopt a **DevOps culture** and leverage cloud to improve efficiency, application release cycles





The Solution

Cloud native platform lets developers focus on building apps



New ConnectedDrive backend run on Red Hat OpenShift and OpenShift Dedicated (Managed Service on AWS) High performance D3 Platform to analyze massive amounts of global test fleet data Cloud native DevOps platform build on Red Hat OpenShift



Where Next?



The intelligent application journey





AI/ML discovery session

A one day, no-cost planning session



To understand the customer's business drivers and technical use cases to propose an Open Al/ML architecture. DETAIL:

Discussion guided by Red Hat AI/ML Consulting Architect

Attendees from Data Science team, Business decision makers, Engineering, Operations, and Application Development

Deep dive into customer use cases



RED HAT PROVIDES:

Our vision for organization-wide AI/ML adoption

Tailored proposal to solve customer use cases

High-level adoption roadmap

Red Hat AI/ML service offerings



Who attends the AI/ML discovery session?

Customer

- Data scientists
- Data engineers
- IT operations
- Intelligent App developers
- Business sponsors

Red Hat

- Al Center of Excellence (leads)
- Red Hat Services
- Open Innovation Labs
- Solutions Architect (optional)
- Account Team (optional)



RED HAT SERVICES OFFERING

AI/ML Architecture Review

Vendor lock-in and high costs of scaling applications	AI/ML Red Hat Services	Open and flexible architecture for AI/ML applications
Inputs	Activities	Outcomes
 Customer AI/ML use cases AI/ML platform requirements 	 Day 1: Customer use cases deep dive Day 2: AI/ML architecture review workshop Day 5: Live demo in customer lab Mentoring and roadmap creation 	 AI/ML platform strategy AI/ML architecture and roadmap Live demo in customer lab
Scope		
Red Hat OpenShift1 week		
	Key contacts	
	 Neeraj Kuppam - Services Portfolio Team William Benton - Al Center of Excellence 	
		Red Hat



Red Hat Services Solution: Open AI/ML Platform

Why

- Accelerate your AI/ML project using the Open Data Hub architecture
- Create a pilot platform for AI/ML use cases
- Bring data scientists, app developers and operations teams together
- Experience DevOps for machine learning best practices
- Build an AI powered application prototype

What

- AI/ML Consulting + Open Innovation Labs Residency
- Infra: Open Data Hub + OpenShift + Open Innovation Labs tools

Who

• Neeraj Kuppam, Will Benton, Patrick Carney, Eric Murphy









Check out our AI/ML success stories

- <u>RBC</u>
- HCA Healthcare
- Boston Children's Hospital
- BMW Group
- <u>ExxonMobil</u>
- <u>NEC</u>
- <u>KBTG</u>
- Discover Financial Services
- Ministry of Defense, Israel

eBook (8 pages) – <u>Top consideration for building a</u> production-ready AI/ML environment

Websites

- <u>https://openshift.com/ai-ml</u>
- <u>https://learn.openshift.com/ai-machine-learning/</u>
- <u>https://openshift.com/nvidia</u>



Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

- in linkedin.com/company/red-hat
- youtube.com/user/RedHatVideos
 - facebook.com/redhatinc

twitter.com/RedHat

