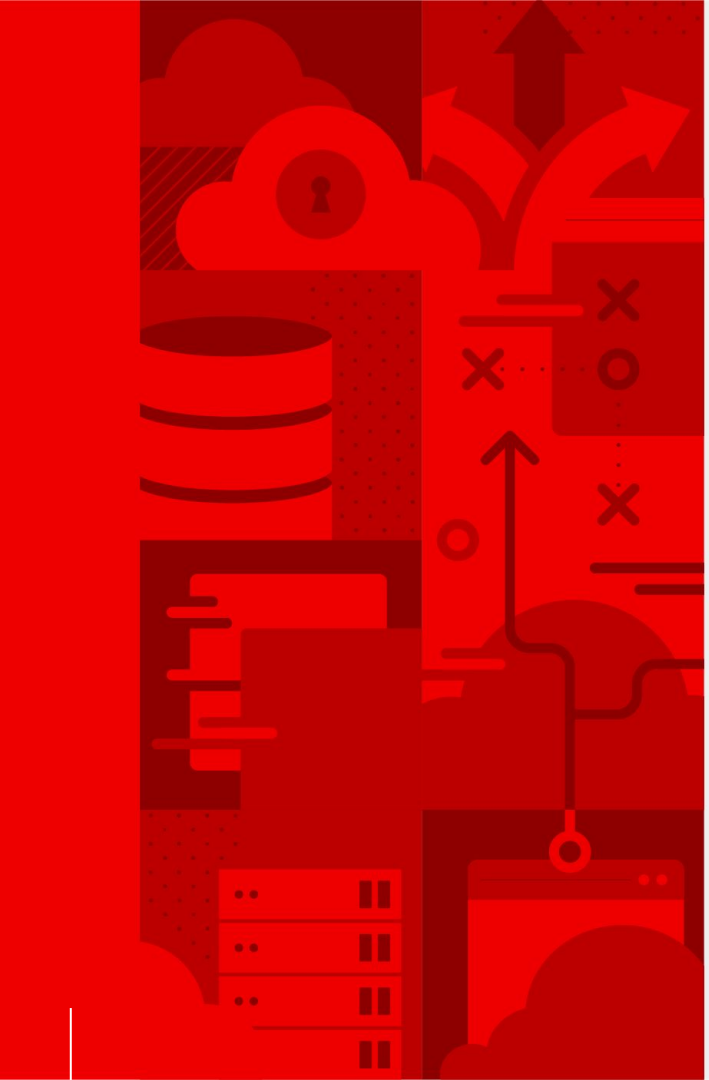


Open Source Use Cases

Rebecca Simmonds rsimmond@redhat.com

Rui Vieira rui@redhat.com

Red Hat

The first instance of open source sharing wasn't related to software at all! [1]

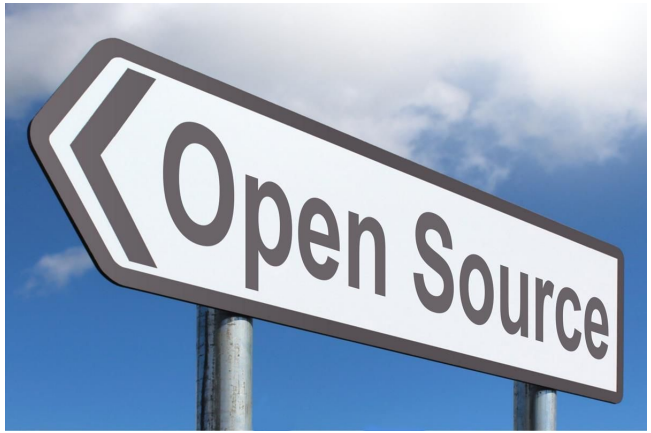
[1] http://redcrackle.com/blog/7-interesting-facts-about-open-source-software

Red Hat

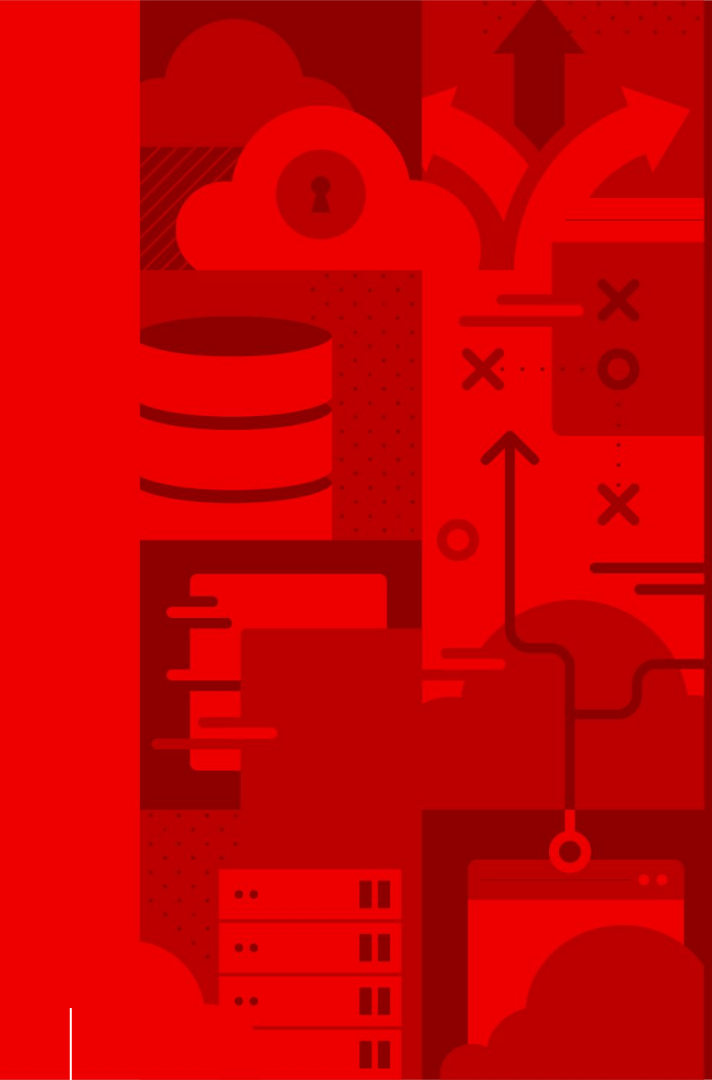
What is Open Source?

Red Hat

The Power of Open



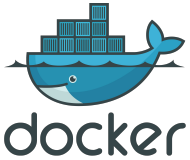
1. Collaboration
2. Feature selection
3. Application direction
4. Community

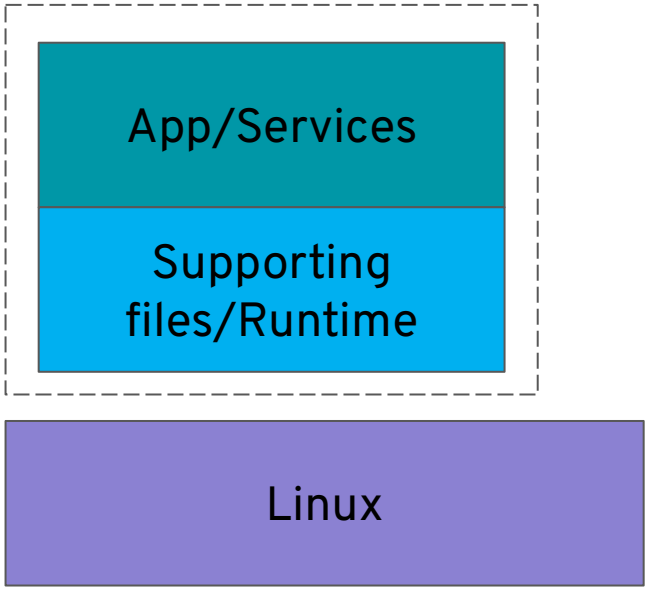


Successful Open Source Projects

Red Hat

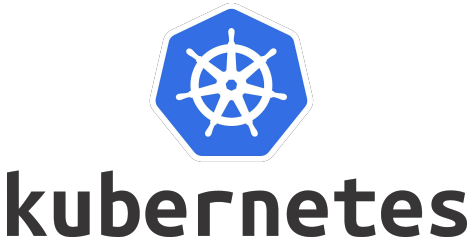
Containers





Red Hat

Kubernetes

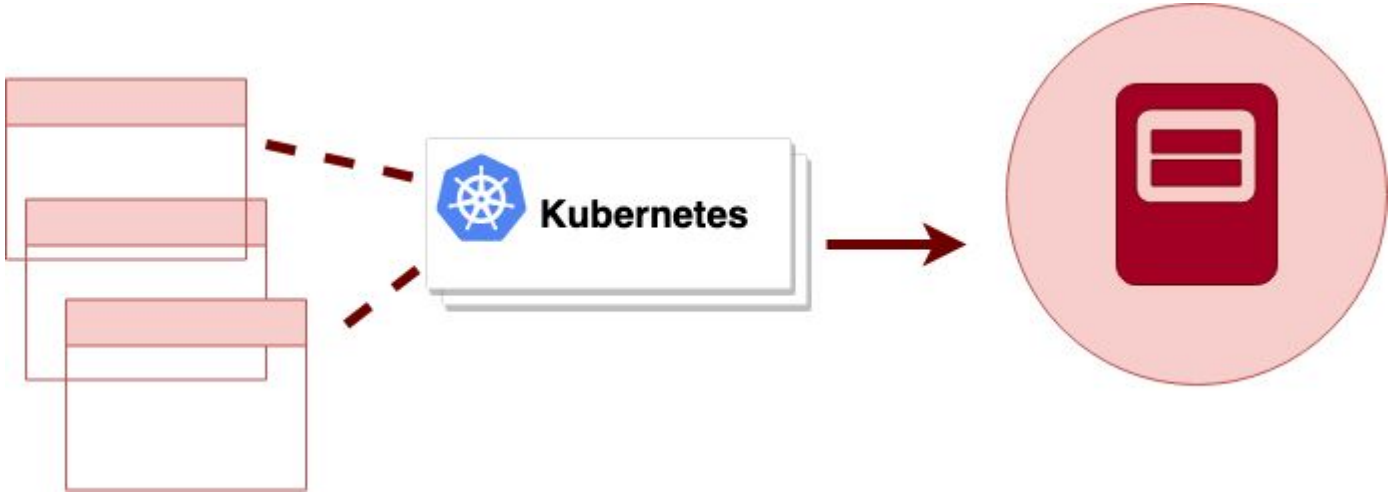


Container orchestration in a clustered environment

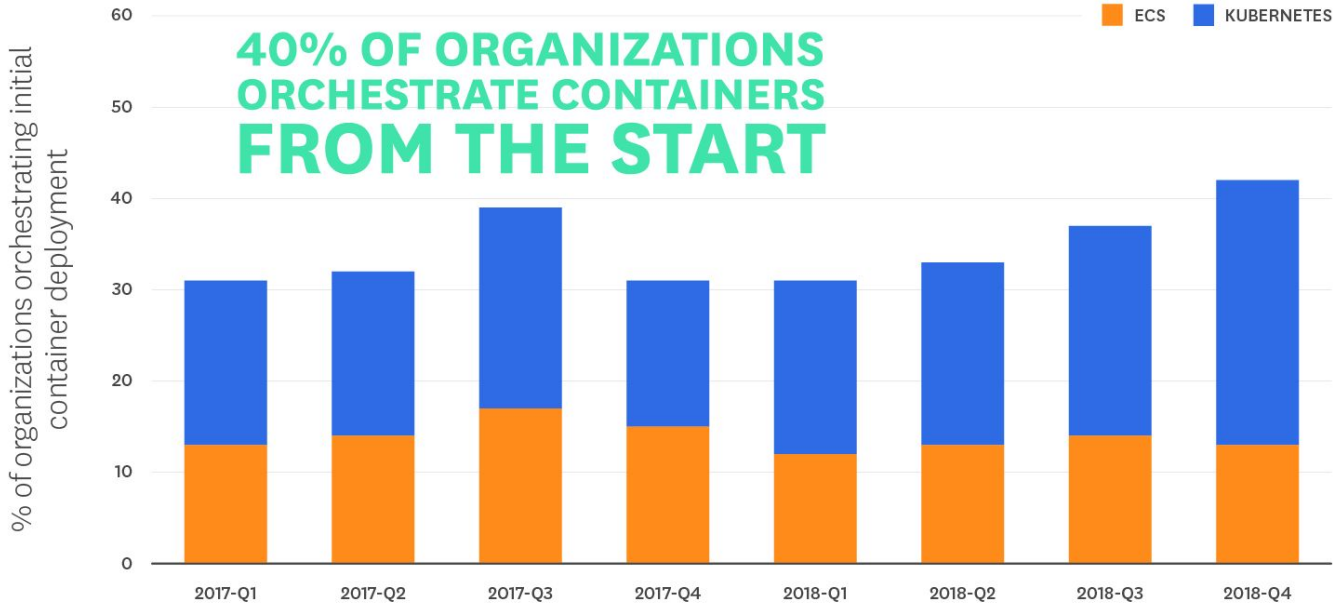
Apache License 2.0

Contributions from Google, Red Hat, Microsoft, IBM, Intel, Rackspace and many more...

Kubernetes cont.



Orchestration Usage at Initial Container Rollout



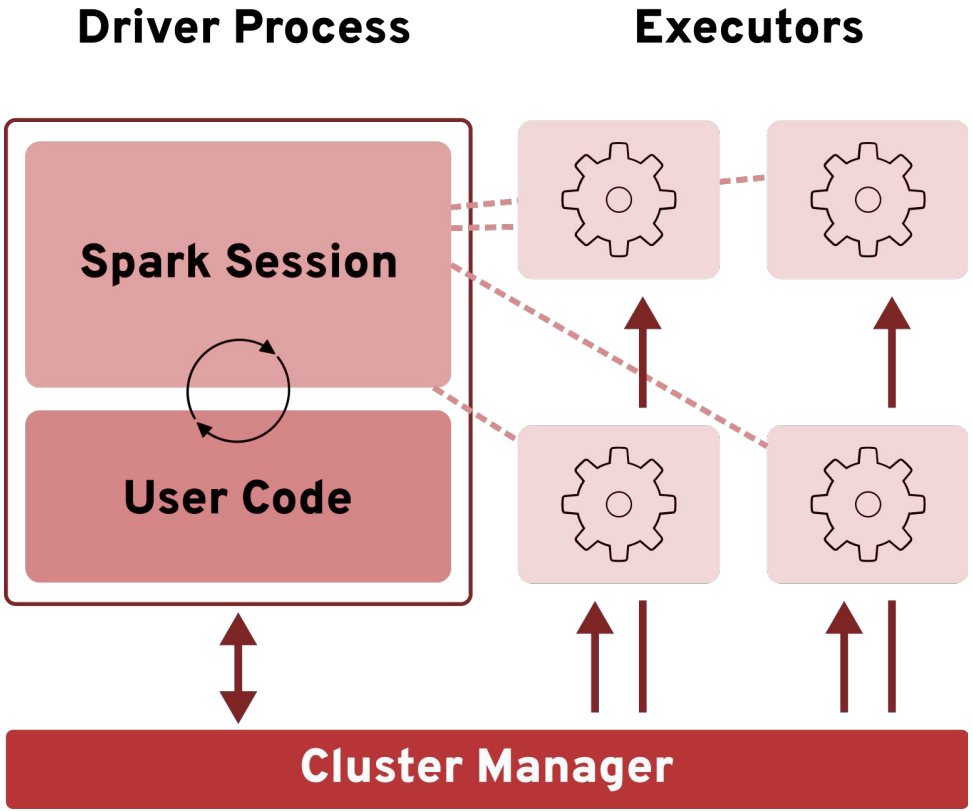
Source: Datadog

Red Hat

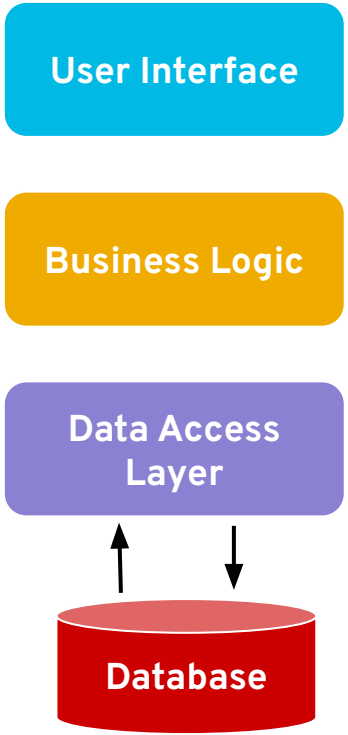
Openshift

Kubernetes Enterprise Distribution

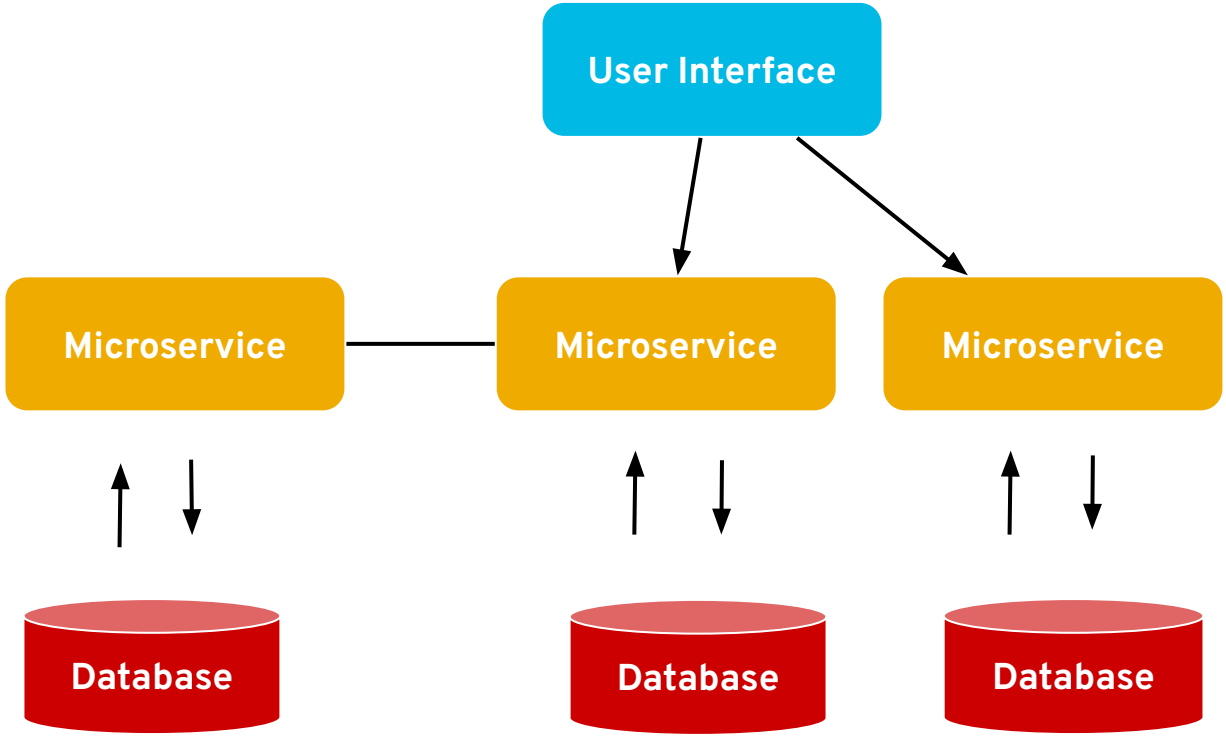
- Container security
- Application delivery and lifecycle
- Validated integrations
- Autoscaling


Driver Process
Executors
Spark Session
User Code
Cluster Manager

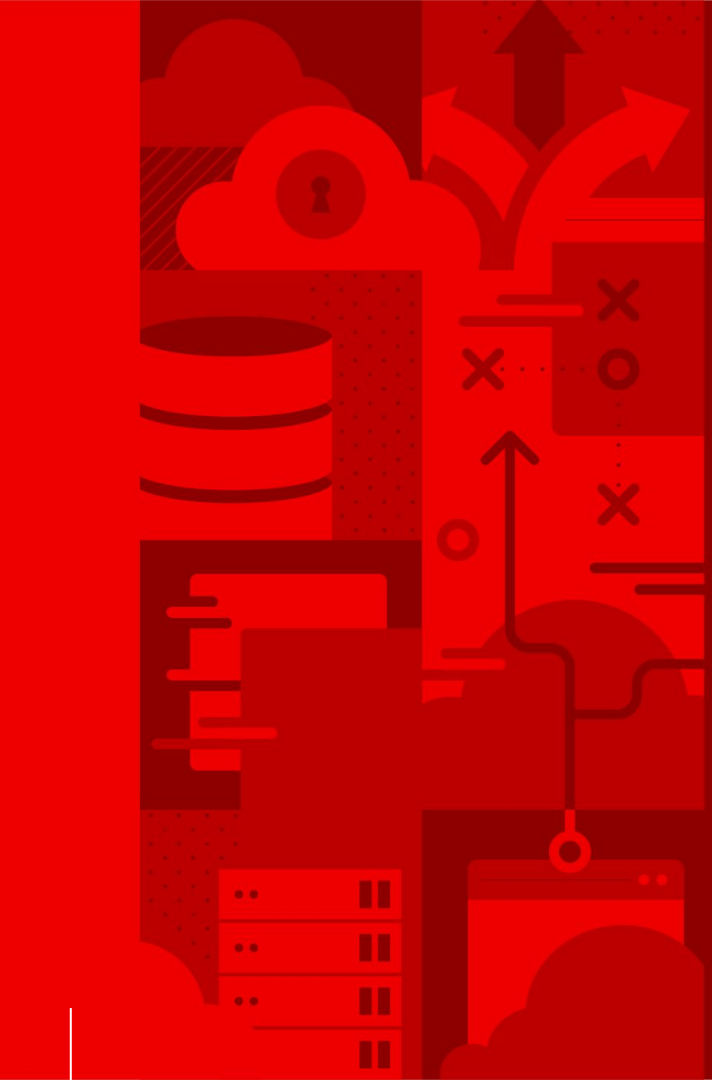
Monolithic Architecture



Microservices Architecture



Red Hat



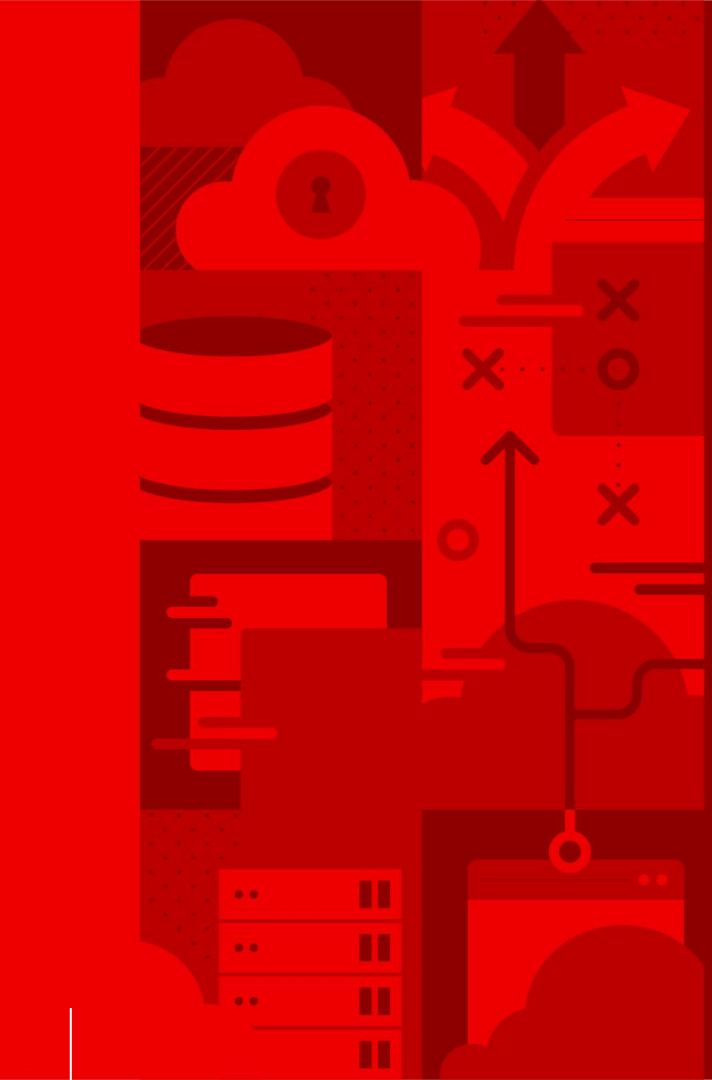
Linux

Red Hat

Linux

- Successful open source project
- Linux kernel
- Operating system
- Red hat Linux/ Fedora
- GPL2

Red Hat



radanalytics

Red Hat

radanalytics

https://radanalytics.io/

Build **intelligent applications** for the cloud

Learning resource

Red Hat

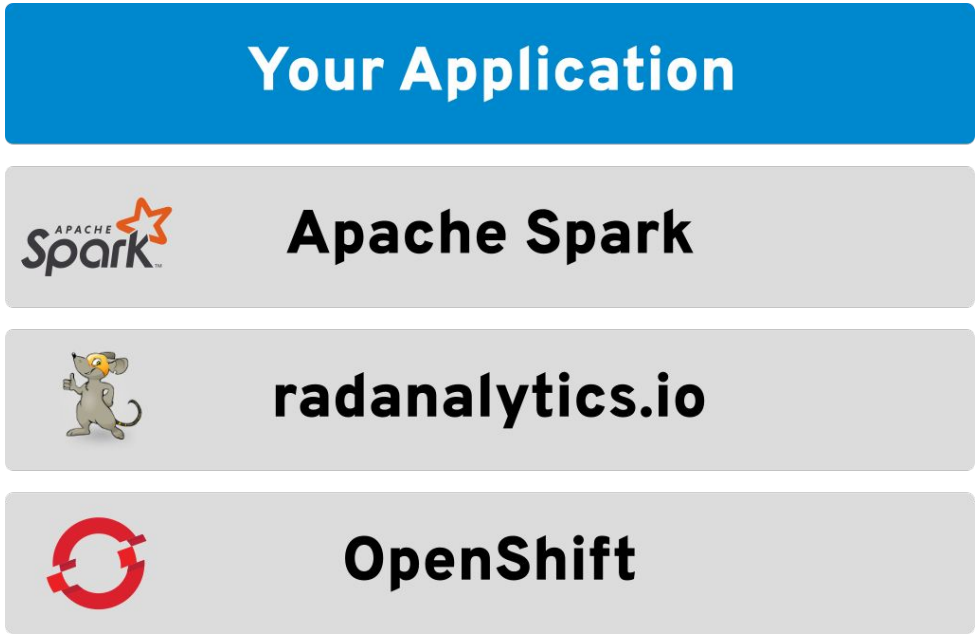
radanalytics

Intelligent applications to collect and learn from data to provide improved functionality with longevity and popularity.

https://radanalytics.io/

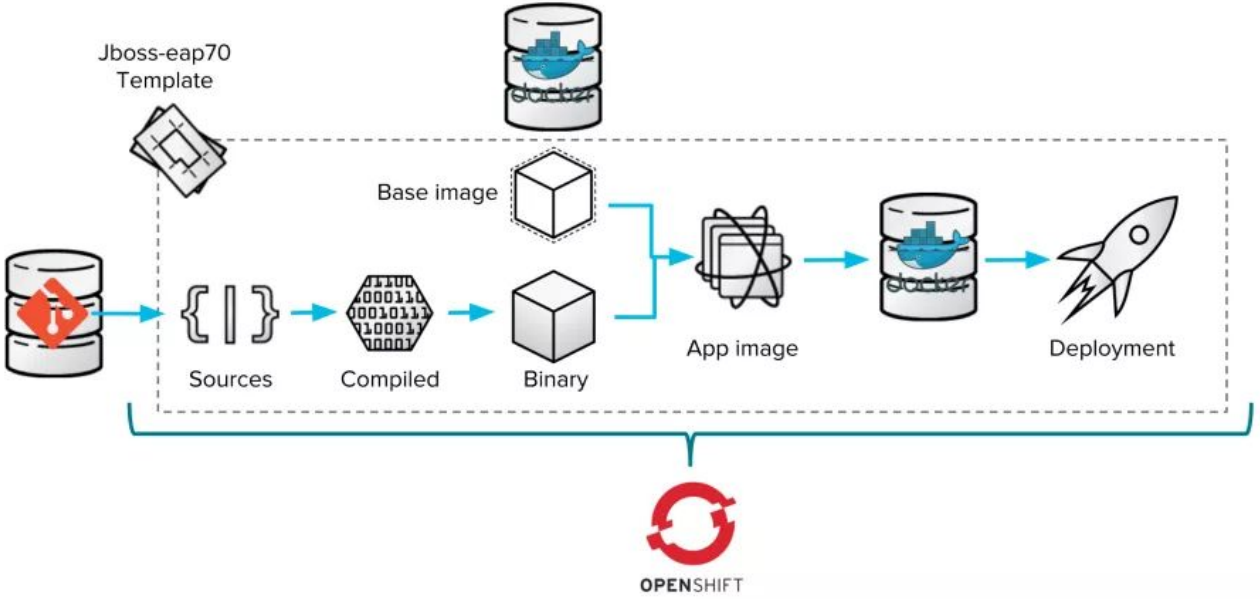
Red Hat

Oshinko



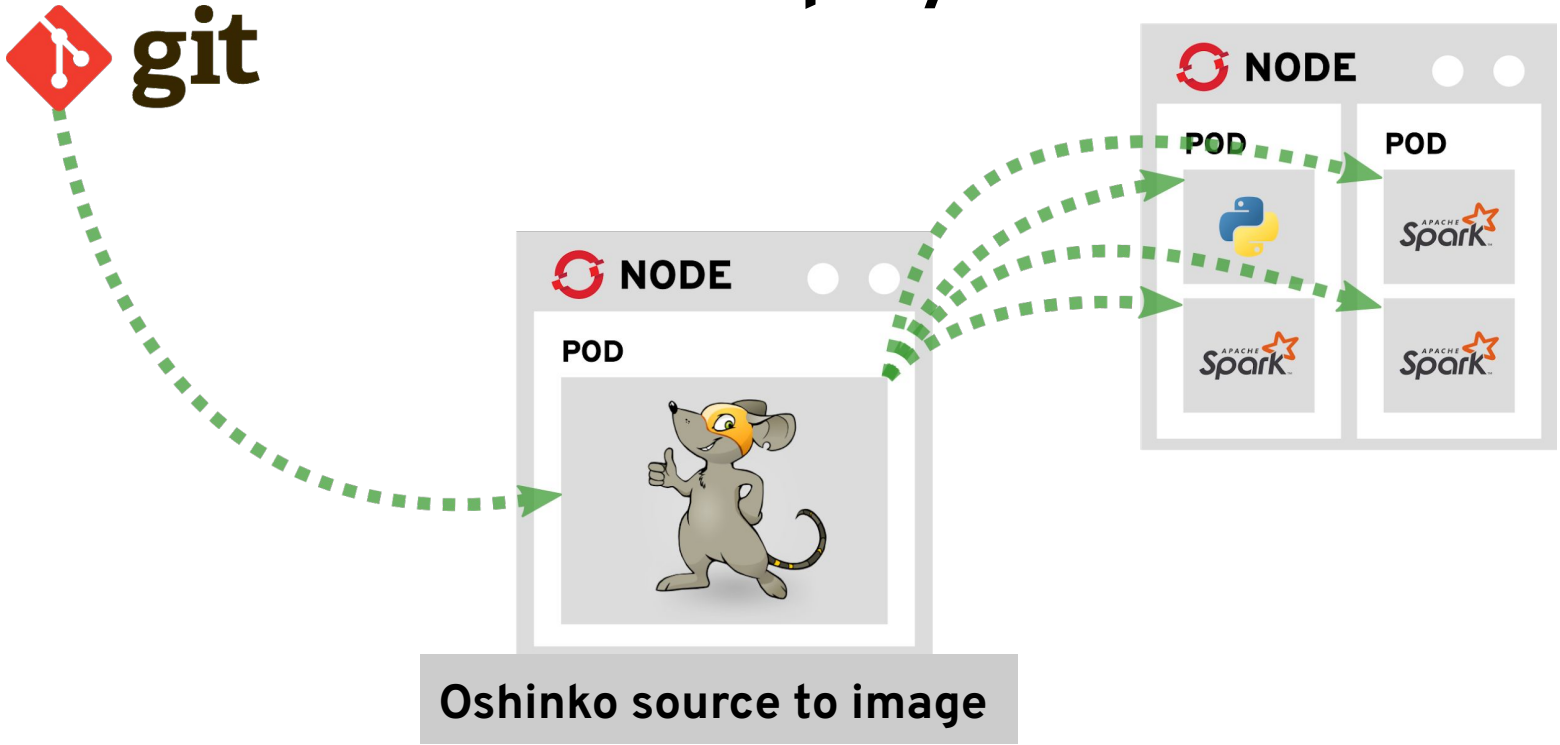
Red Hat

Source to image

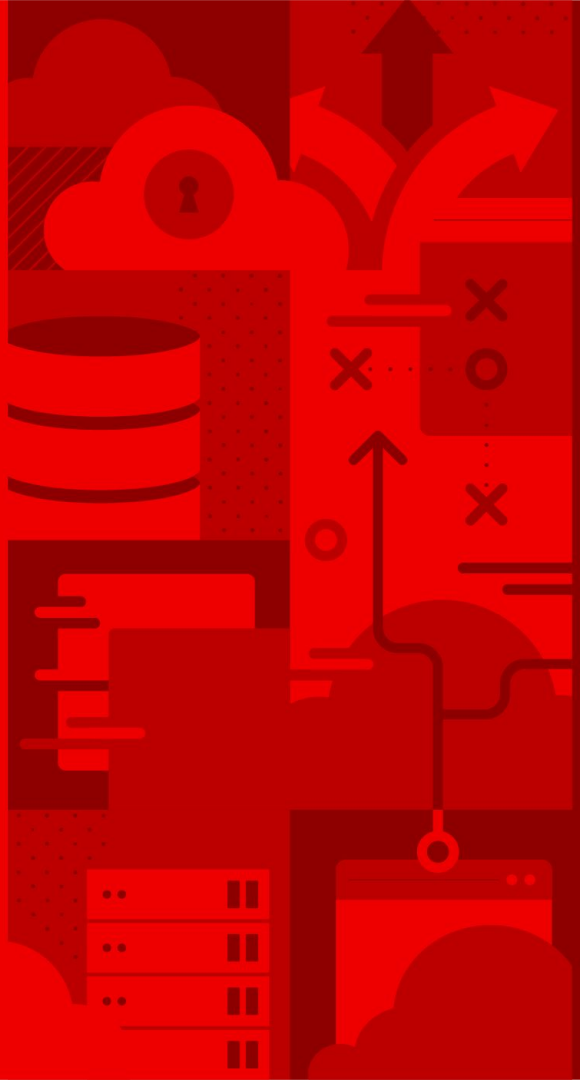


Red Hat

Oshinko Deployment



Red Hat



Open Source Community

Red Hat

What is Community?

Red Hat

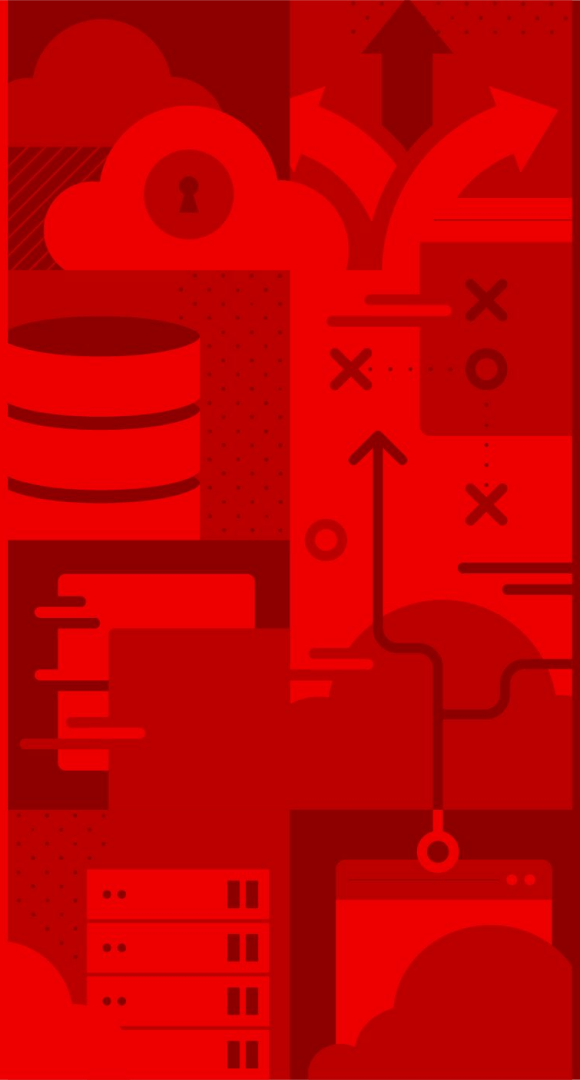
Setting up a Community

- Do you want a large community?
- Selective community, small but focused?
- How will the project be structured? - will you support growth yourself?

Decide this *before* making an open source a project

Example Communities

- Linux and Apache Spark
  - One person’s hobby
  - Grew quickly with interest
- Linux containers
  - Google and Red Hat backed
  - Large community - world wide



Open Source and Innovation

Red Hat

2009

Matei Zaharia class project at UC Berkeley (Mesos)

2009

Matei Zaharia class project at UC Berkeley (Mesos)

2019

Most active Apache Big Data project *
+1000 contributors
Expanded to included Structured Streaming, Machine Learning, …
International conferences

** - Hadoop is classified as a "database project".*

Normally research projects get abandoned after a paper is

published.

There are many components. And if you look back, you can always revise history.

What was different?

Especially if you had success.

First of all, we had a fantastic group of students.

Matei, the creator of Spark and others who did Mesos. And then another great group of **different students**

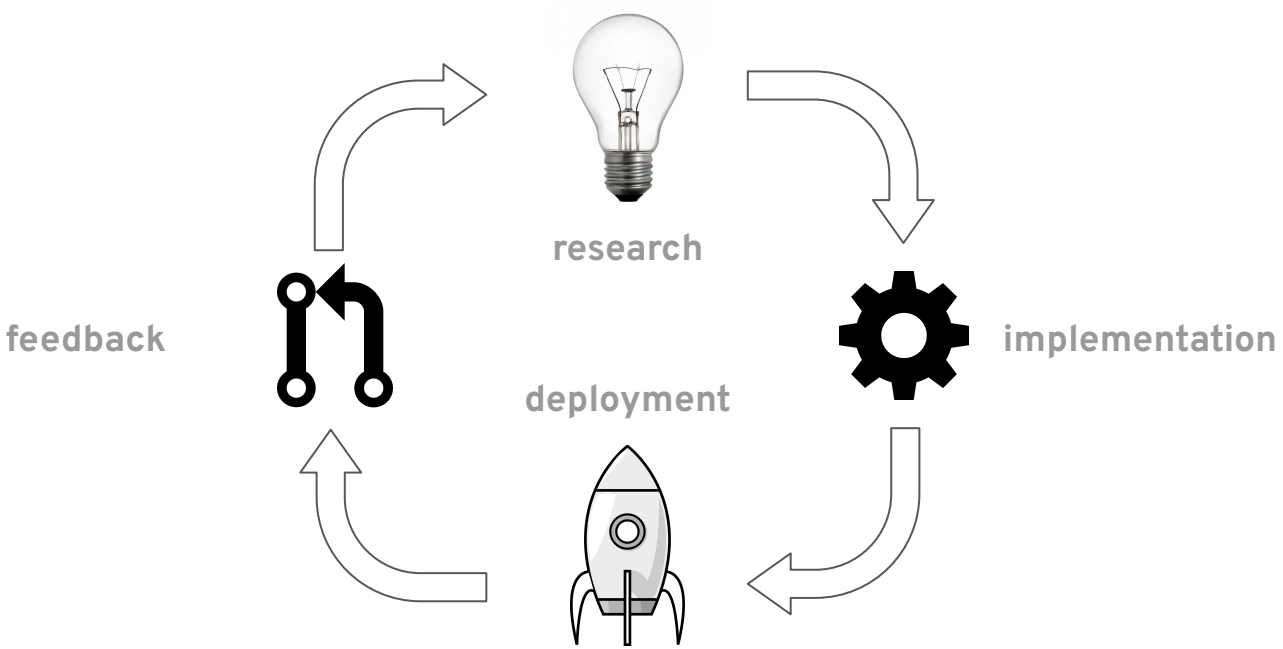
who contributed and built different modules on top of Spark, and made what Spark it is today, which is really a

platform. So, that's one: **the students**.

The other one was a **great collaboration with the industry**. We are seeing first hand what the problems are,

challenges, so you're pretty anchored in reality.

Lifecycle



Use Cases

Project **jiminy**

A cloud-ready, scalable recommendation engine.

- **cloud -ready** - deployable on Kubernetes/OpenShift
- **scalable** - distributed computations supported by Apache Spark
- **recommendation engine** - based on Alternating Least Squares (ALS), a well-known algorithm, winner of the Netflix prize

Red Hat

User Story

As a **developer**, I want a system can be easily deployed from source in a cloud environment. The system should also be easy to tailor or extended to my specific needs.

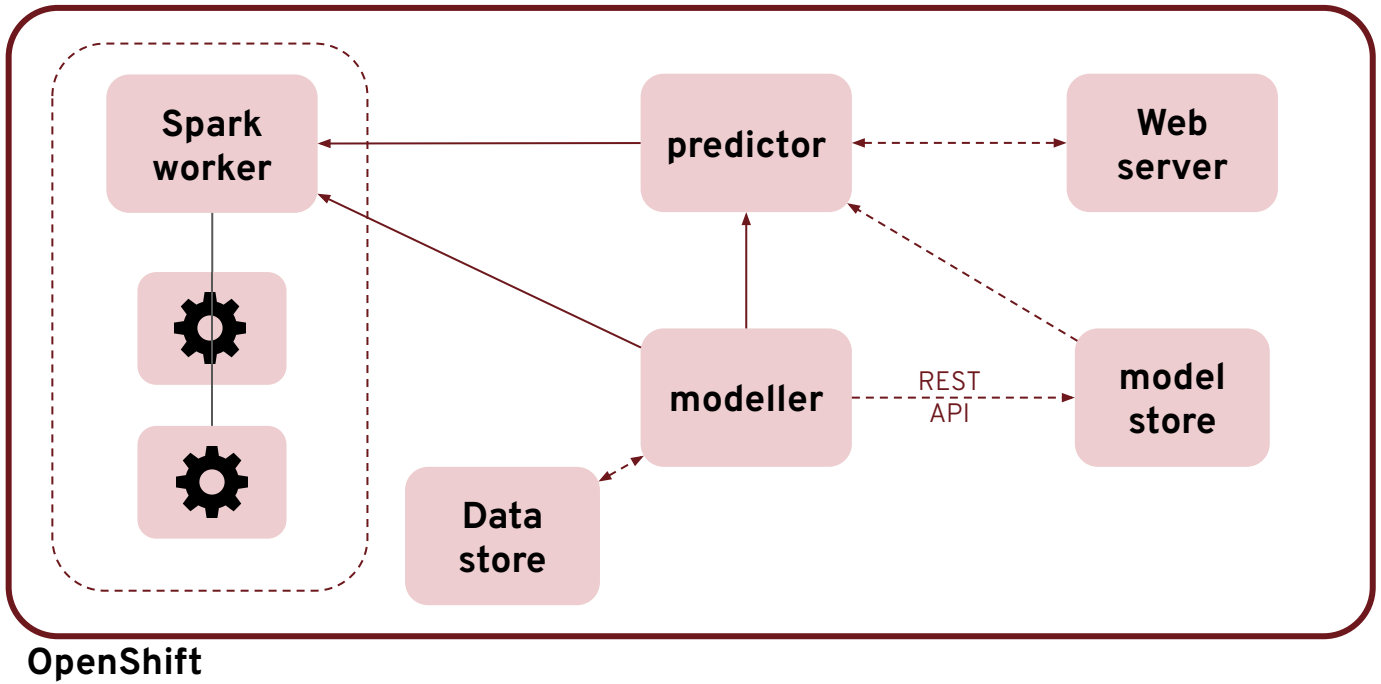
User Story

As a **business**, I want a system which helps maximising revenue by providing users with meaningful new product recommendations.

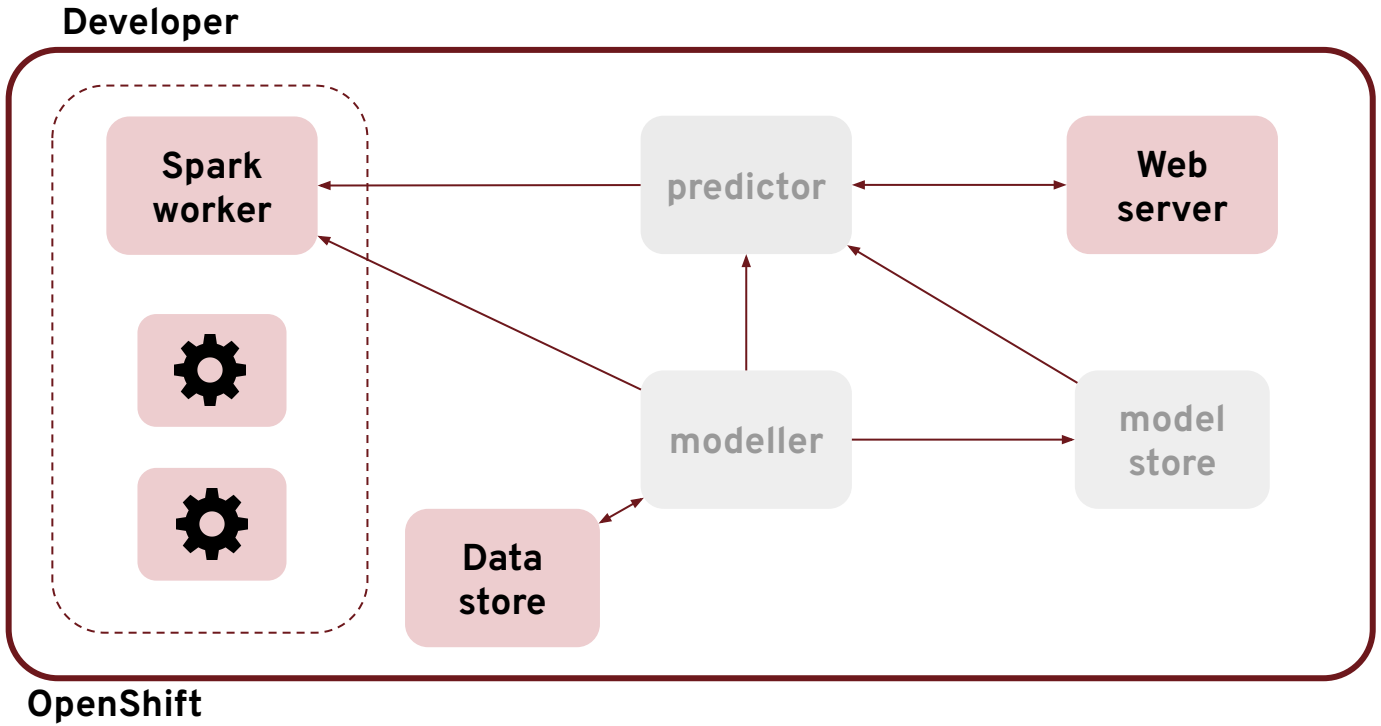
User Story

As a **data scientist**, I want a system which is flexible enough to let me focus on the recommendation algorithms. I'm also interested in reproducibility.

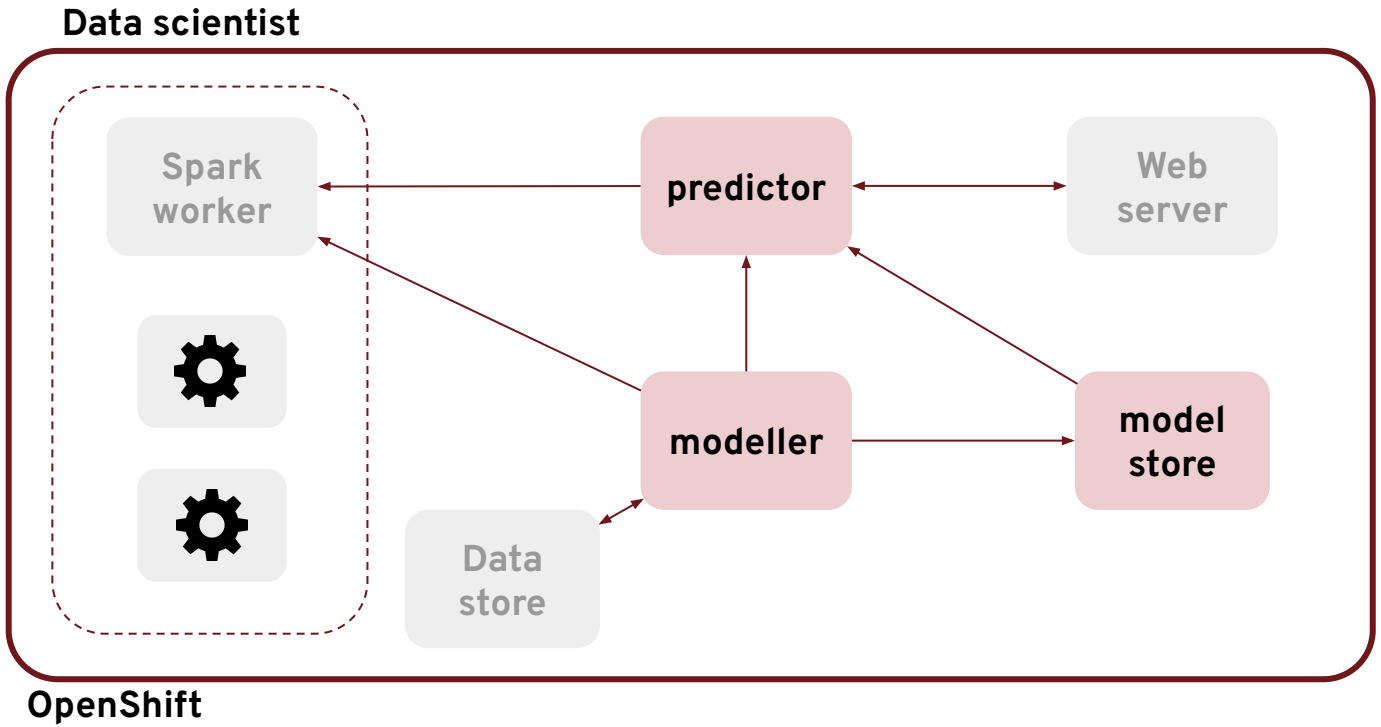
Architecture



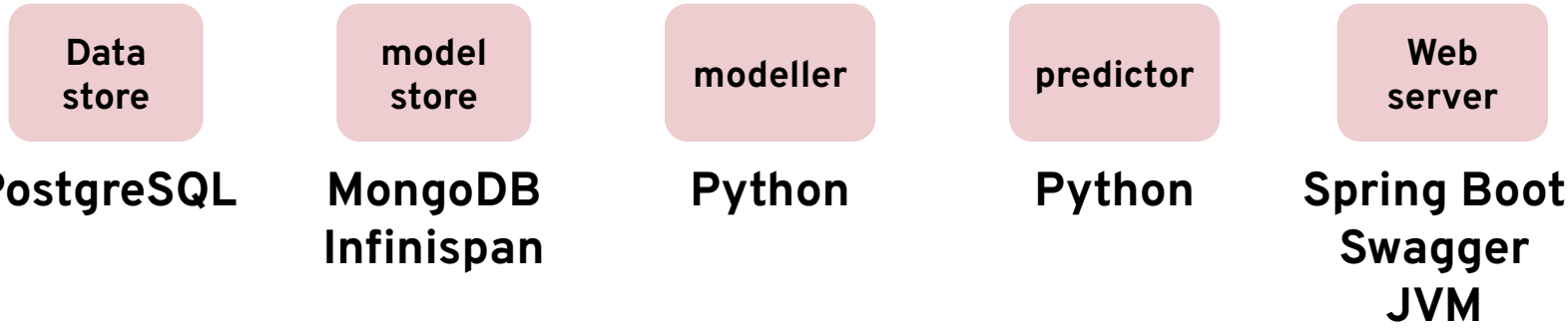
Architecture



Architecture



Open Source Technologies



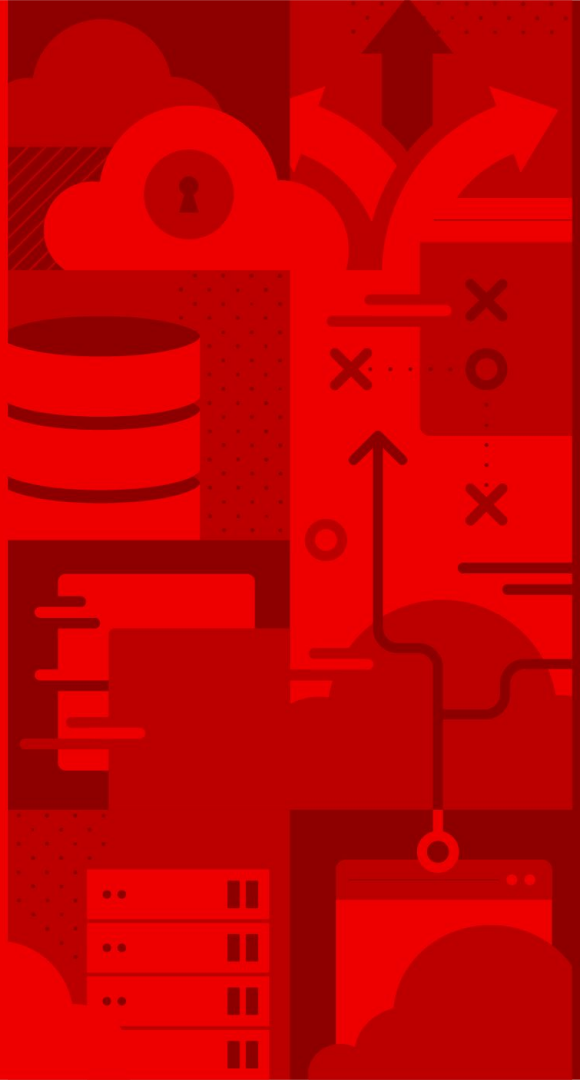
Engagement

- Learning resources
  - Workshops, conferences
- Technology showcases
- Basis for customised solutions

Demo

Why radanalytics?

Red Hat



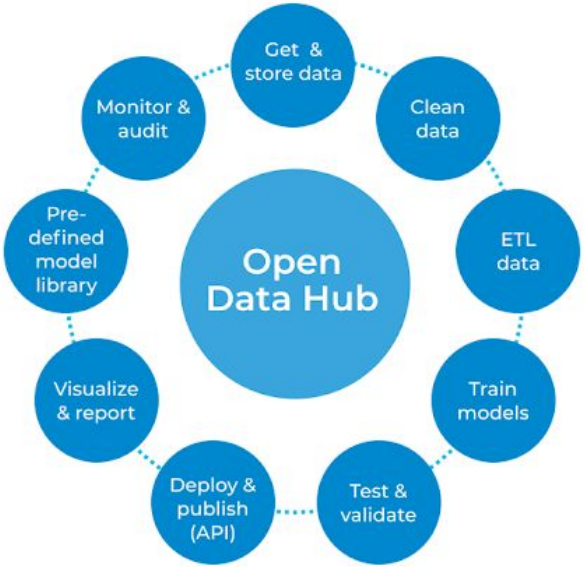
OpenDataHub

Red Hat

OpenDataHub

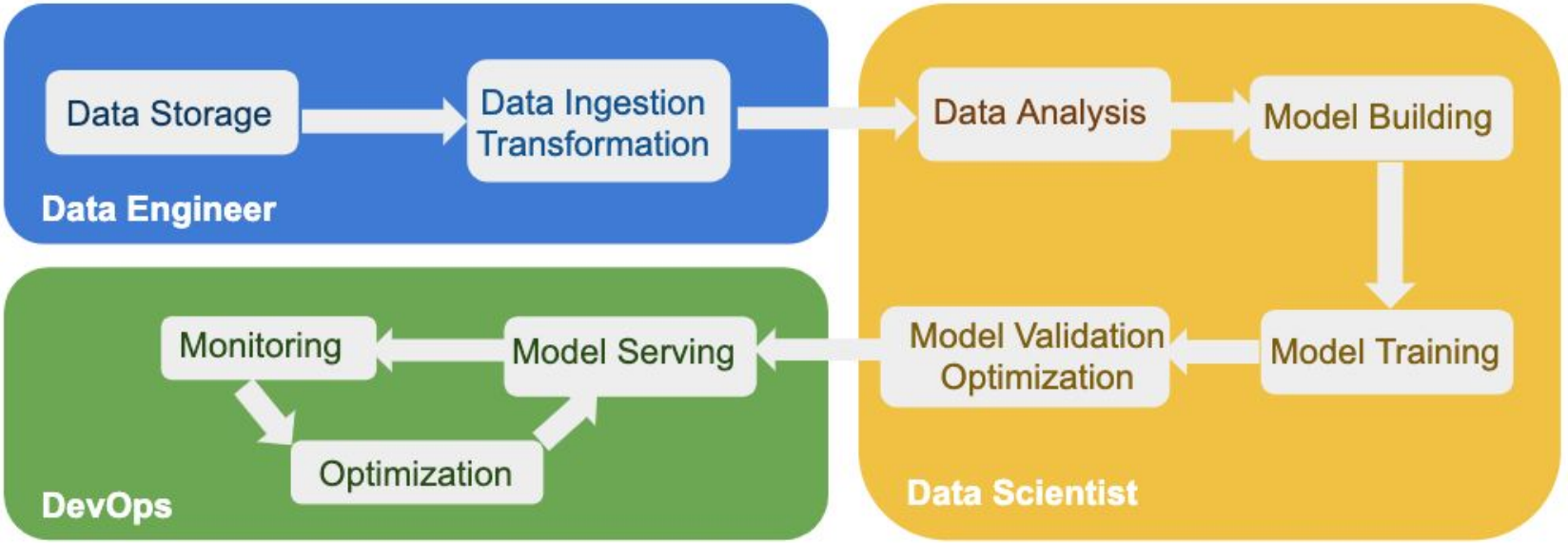
A **reference architecture** for an AI and Machine Learning as a **service platform** for OpenShift built using **open source tools**

End-to-End



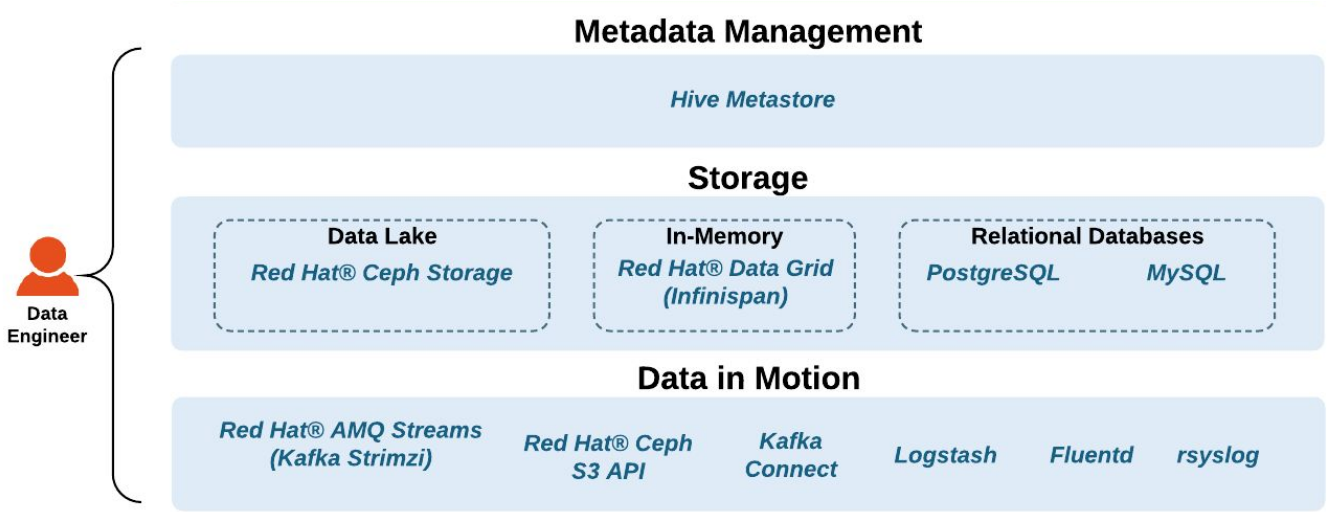
Source: https://next.redhat.com/2018/09/19/a-hub-for-open-data-at-mass-open-cloud/

Personas



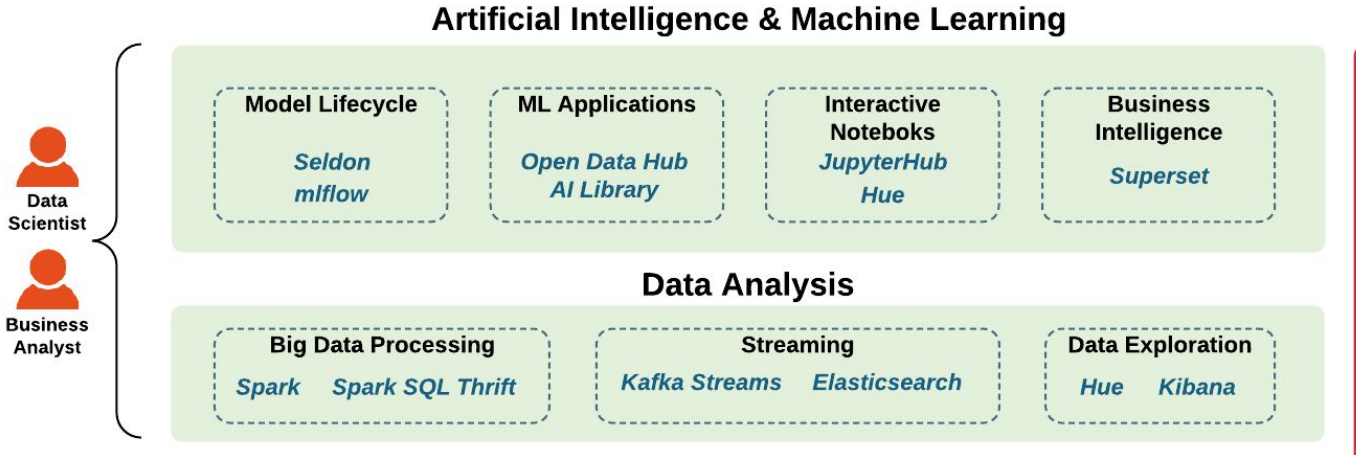
Source: https://next.redhat.com/2018/09/19/a-hub-for-open-data-at-mass-open-cloud/

Data Engineers



Source: https://opendatahub.io/arch.html

Data Scientists



Source: https://opendatahub.io/arch.html

Mass Open Cloud (MOC)

1. To create an inexpensive and efficient at-scale production cloud utility suitable for sharing and analysing massive data sets and supporting a broad set of applications.
2. To create and deploy the OCX model, enabling a healthy marketplace for industry to participate at all levels in the cloud and profit from doing so.
3. To create a testbed for research in and prototyping of cloud technology, empowering a broad community of researchers, open source developers and companies to develop new cloud computing technologies.

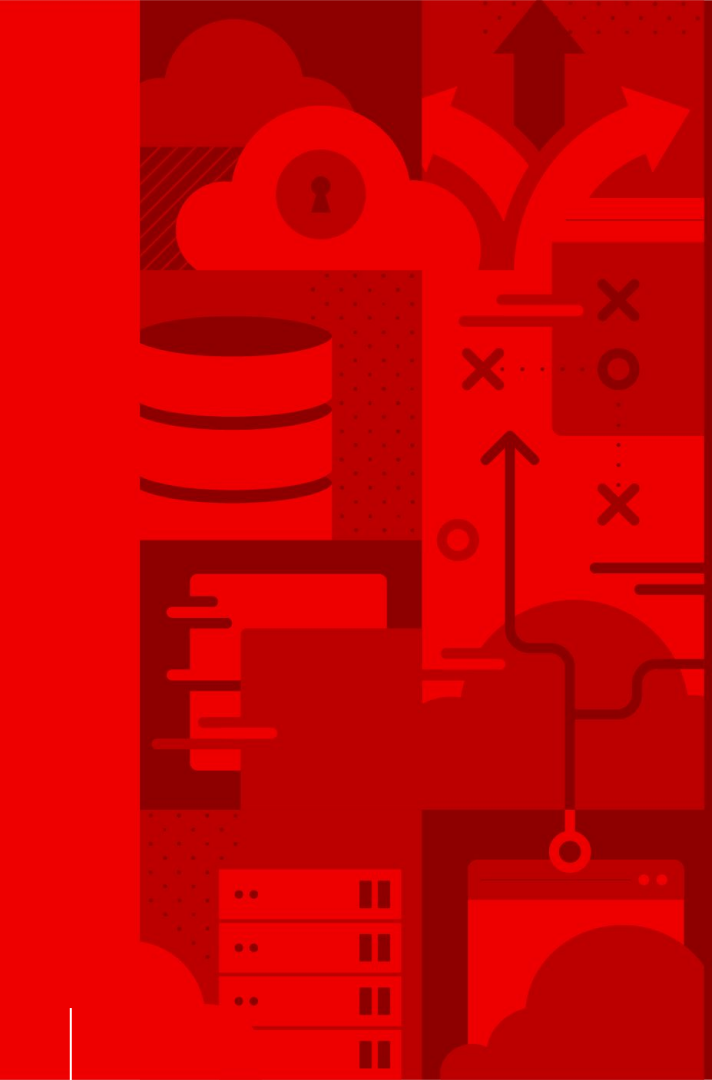
Mass Open Cloud (MOC)

Project's core partners:

- Academic (Boston University, Harvard University, Northeastern University, MIT)
- Government (Massachusetts Technology Collaborative, United States Air Force)
- Non-profit (MGHPCC)
- Industry (Cisco, Intel, NetApp, Red Hat, Two Sigma)

Challenges of Open Source

- Contribution guidelines
- Peer review
- Strategy / Focus
- Support / Documentation



Conclusions

Red Hat

Lessons learnt

- Open needs to be planned
- Communities need to be nourished to succeed

BUT

- You can have a hobby project
- Experiment and find your ideal spot

Red Hat

Conclusions

- Open is quicker and easier
- Collaboration and remote working made easier
- Relevant and customer driven application features

How you can get involved

https://radanalytics.io/
https://opendatahub.io/

Contact us:

rsimmond@redhat.com
rui@redhat.com

Red Hat