

# Spark from Notebook to Cloud Native Application

Rebecca Simmonds  
Senior Software Engineer

[rsimmond@redhat.com](mailto:rsimmond@redhat.com)

@becky\_simmonds

# Aim

**To empower others  
with the tools and tips to go  
from prototype to production  
using Apache Spark**

# Prototype

# Requirements

1. Use case
2. Problem domain
3. Data set
4. Tools and techniques

# Use Case



+

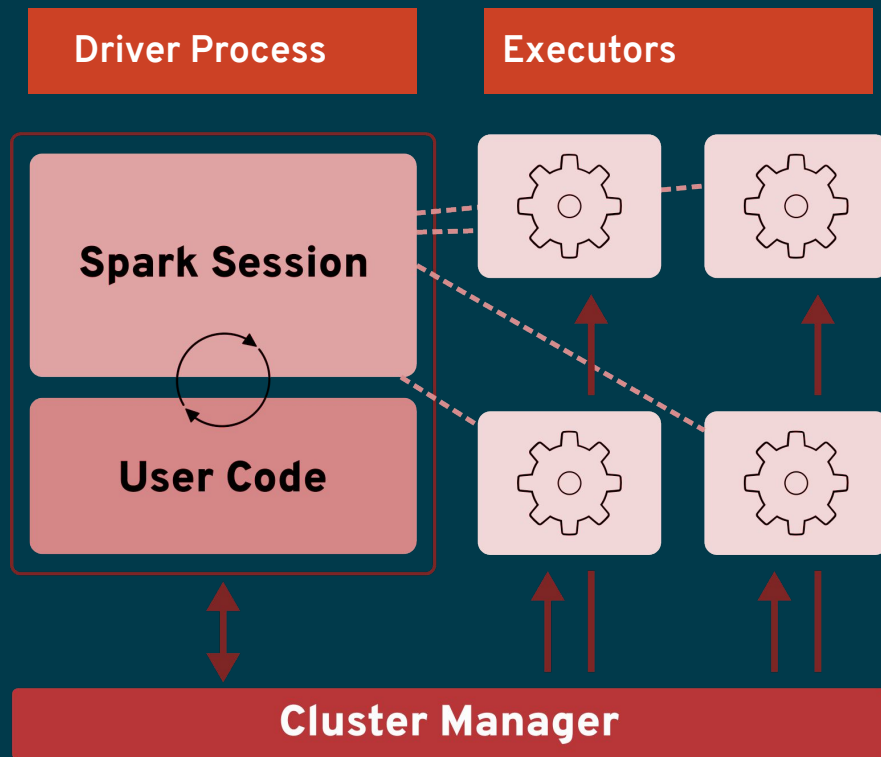


Variety	Country	Points	Region
Tinta de Toro	Spain	98	Toro
Cabernet Sauvignon	US	70	Napa Valley
Macauley	US	50	Knights Valley

# Jupyter Notebook

- Open-source web application
- Create and share live code examples
- Python code
- It empowers users with visualisation tools

# Spark





# Demo

# Conclusions

- Easy to setup and get going
- Lots of visualisations to practise with
- Great method for proof of concept

# Production

# Next Steps

1. Cloud based for scale and portability
2. Tooling and techniques
3. Database/more robust store
4. Testing

# Cloud Native Applications

Applications that are:

1. designed to run in the cloud
2. scalable
3. modular
4. and resilient

# Containers

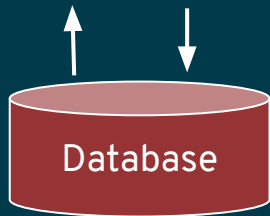
- Allow you to package and isolate a runtime environment
- Easily portable to different environments
- Scalable
- Quick and easy to deploy

# Monolithic Architecture

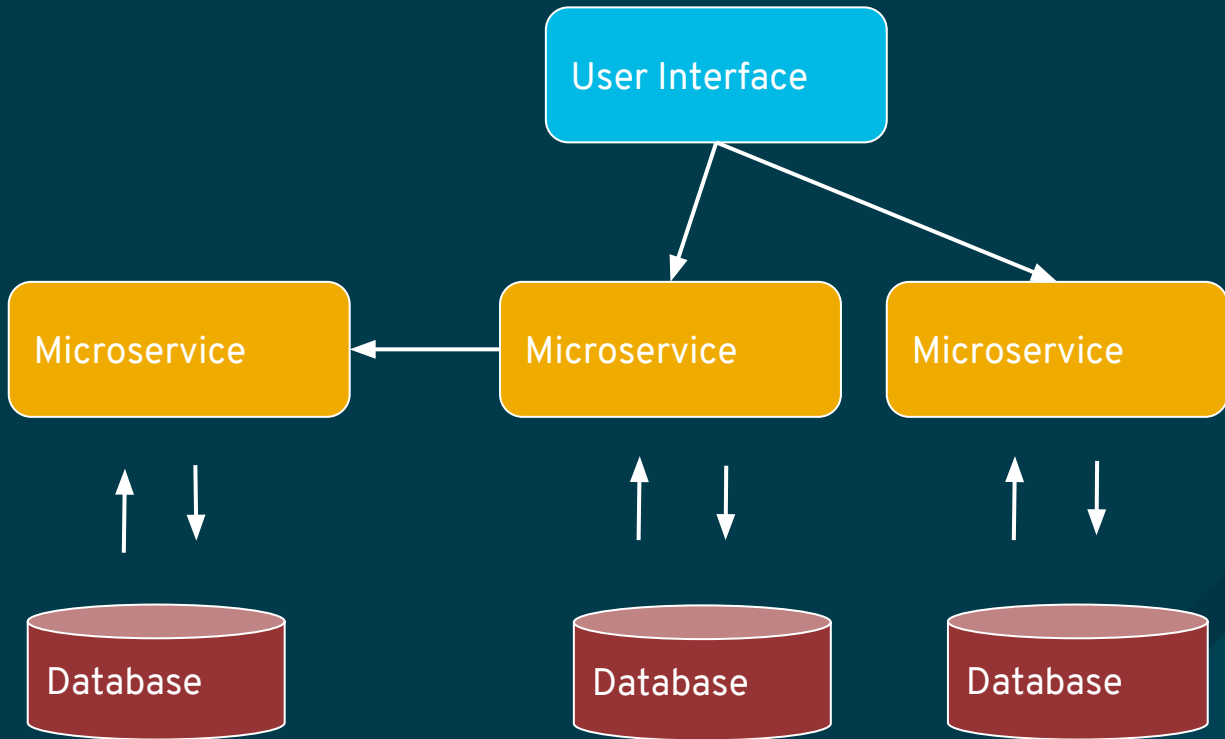
User Interface

Business Logic

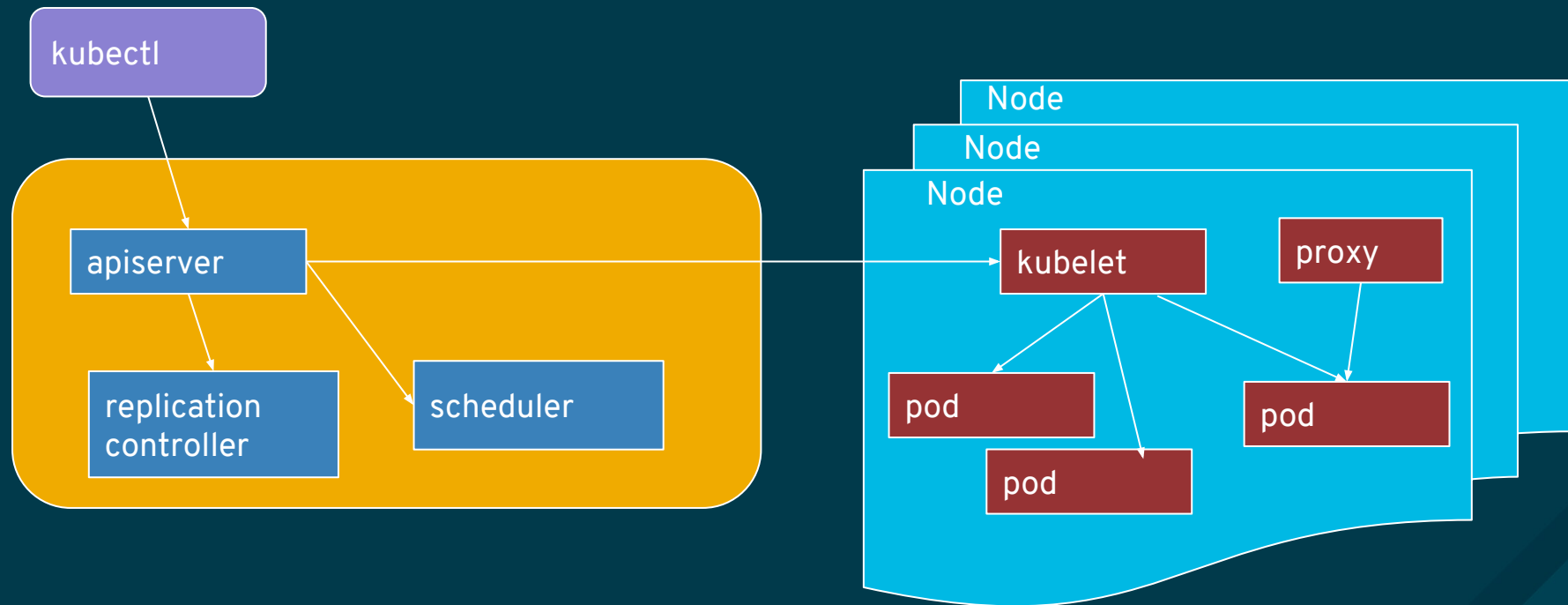
Data Access Layer



# Microservices Architecture



# Kubernetes





# Radanalytics.io

An open source community working to empower  
intelligent applications on kubernetes

Projects and tutorials to empower developers with  
machine learning techniques

# Oshinko

## Your Application



**Apache Spark**

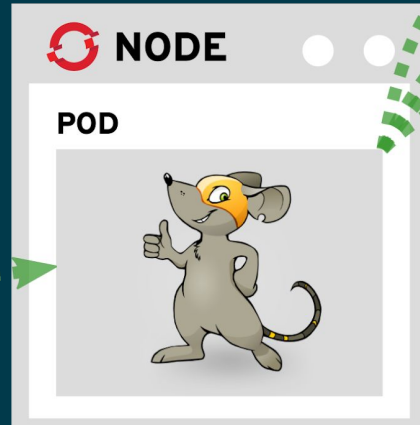


**radanalytics.io**

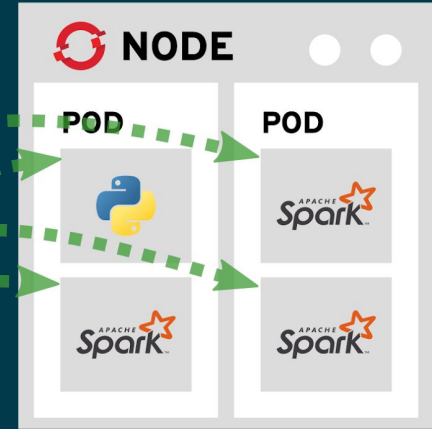


**OpenShift**

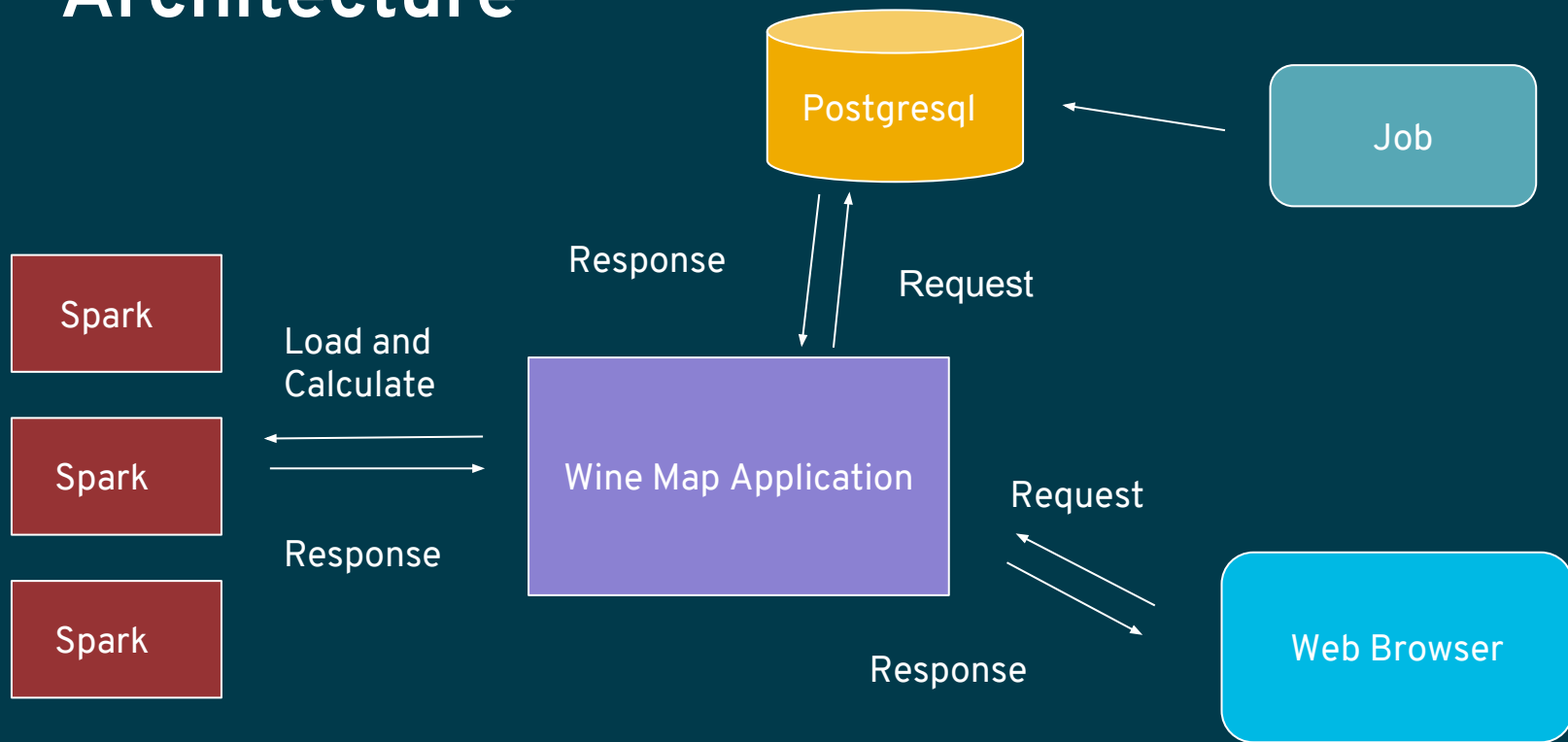
# Oshinko Deployment



Oshinko source to image



# Architecture



# Demo

**# test command**

os::cmd::try\_until\_text

**# what to test**

'oc new-app --template=oshinko-python-spark-build-dc

-p APPLICATION\_NAME=winemap

-p GIT\_URI=https://github.com/radanalyticsio/winemap.git

**# expected result**

'Success'

# Conclusion

- Jupyter notebook for prototyping
- VISIT [radanalytics.io](https://radanalytics.io)
- Deploy your own cloud native applications

@becky\_simmonds

[rsimmond@redhat.com](mailto:rsimmond@redhat.com)

<https://radanalytics.io/applications/wine-map>

