Adapting Kubernetes For Machine Learning Workflows

Bloomberg

Engineering

Open FinTech Forum 2018
October 10, 2018

Ania Musial, Senior Software Engineer

Keith Laban, Senior Infrastructure Engineer

TechAtBloomberg.com

About Us

- Ania Musial
  - Senior Software Engineer
  - Machine Learning Platform, AI Group

- Keith Laban
  - Senior Infrastructure Engineer
  - Data & Analytics Infrastructure

- https://www.techatbloomberg.com/post-topic/data-science/

Outline

- Machine Learning at Bloomberg
- Toward a Model Development Life Cycle
- Machine Learning the Kubernetes Way
- Concluding Remarks

TechAtBloomberg.com

Bloomberg

Engineering

Open FinTech Forum 2018

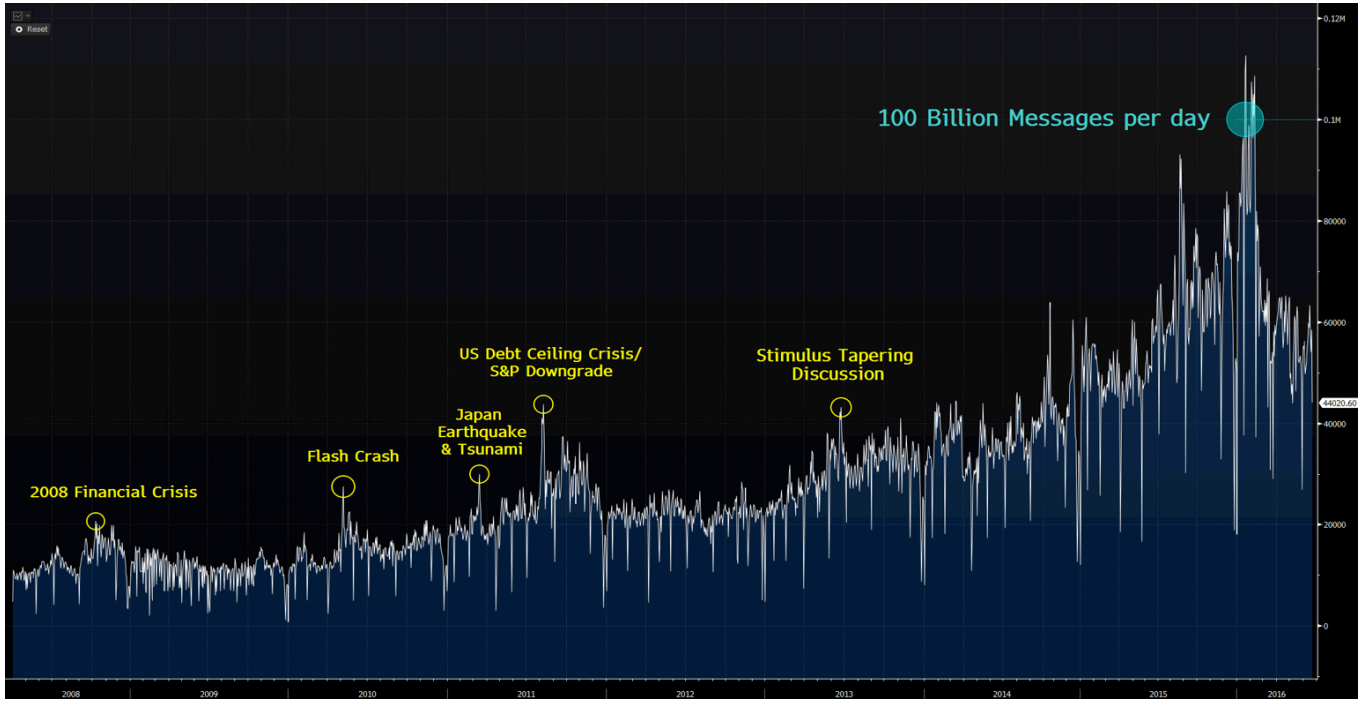
Machine Learning at Bloomberg

TechAtBloomberg.com

Bloomberg

Engineering

Scale and Performance: Market Data



TechAtBloomberg.com

Bloomberg

Engineering

Natural Language Processing & Events

TechAtBloomberg.com

Bloomberg

Engineering

Over $136 Billion Wiped Out in Minutes

TechAtBloomberg.com

Bloomberg

Engineering

Applications Contingent on ML

TechAtBloomberg.com

Bloomberg

Engineering

Open FinTech Forum 2018

Toward a Model Development Life Cycle

TechAtBloomberg.com

Bloomberg

Engineering

Different Roles, Different Priorities

Customers

- Smarter Functionality
- Privacy
- Correctness
- Uptime/Stability

Engineers

- Less operational burden
- Support for multi-tenancy
- Easy data access but also secure and properly privileged
- Avoid duplicated code
- Reproducibility of results

ML Practitioners

- Keep up with latest trends in academia and forefront of Deep Learning innovations
- Access to specialized hardware
- Rapid iteration over research ideas
- Focus on business problem, not building infrastructure

Embarking on a New Problem: What?

- What **data** is relevant to my business problem?
- What does the data **look** like?
- What **features** should I create from my data?
- What metric should I use to **measure** my model performance?
- What **algorithms** or **parameters** should I test out?
- How is my model doing on live data in **production**?

TechAtBloomberg.com

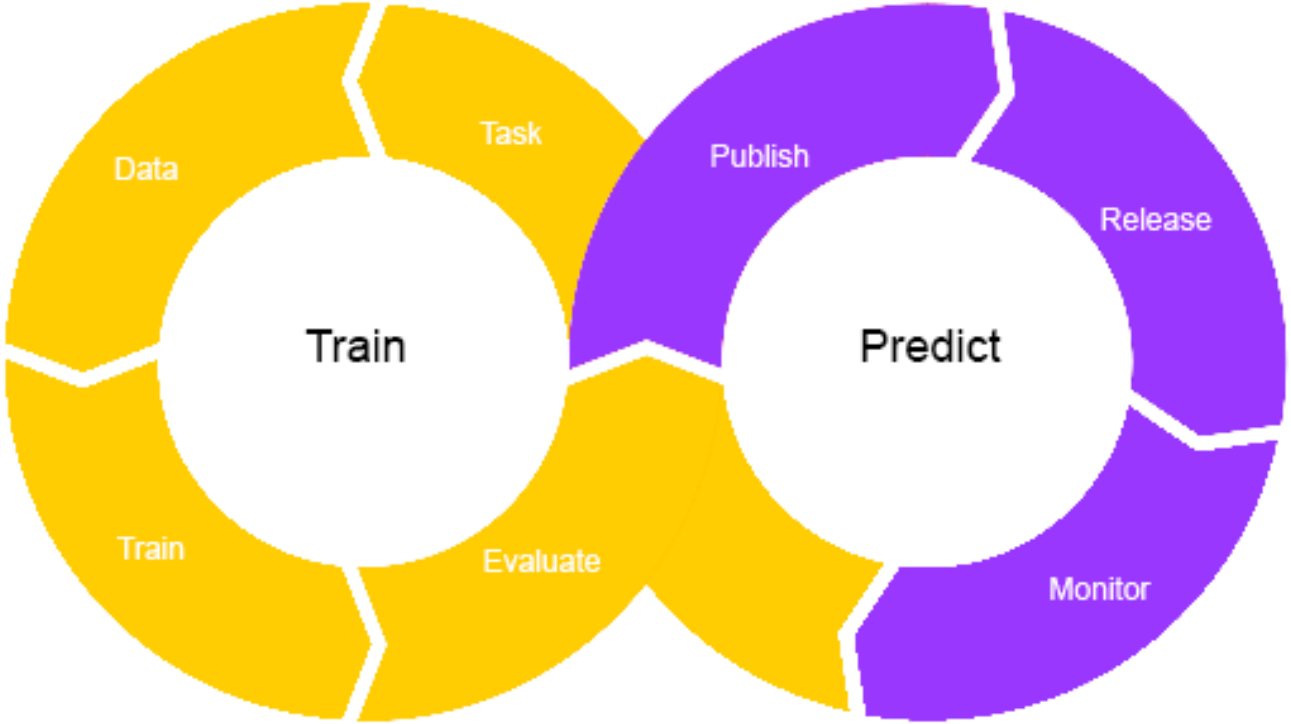
Bloomberg

Engineering

Embarking on a New Problem: How?

- Where is my data **located**?
- How do I get **permissioned** for my data?
- How do I load this data in a native **visualization** tool?
- How do I get this **toolkit** installed and operational?
- How do I **schedule** my long-running job?
- How much **compute power** is available for me to exploit?
- How do I **deploy** my model?

The Infinite Model Development Loop



TechAtBloomberg.com

Bloomberg

Engineering

Achieving Nirvana

Foundational Building Blocks

- Specialized Hardware
- Standardized Runtimes
- Batch Scheduling

Data Liquidity

- Data Access
- Security

Developer Experience

- Ease of Use
- Rapid Iteration
- Native Tooling

ML Specifics

- Experiment Management
- Hyperparameter Tuning
- Result Reproducibility

Open FinTech Forum 2018

Machine Learning the Kubernetes Way

TechAtBloomberg.com

Bloomberg

Engineering

What is Kubernetes?

“Kubernetes is an open-source system for automating deployment, scaling, and management of containerized applications.”

kubernetes

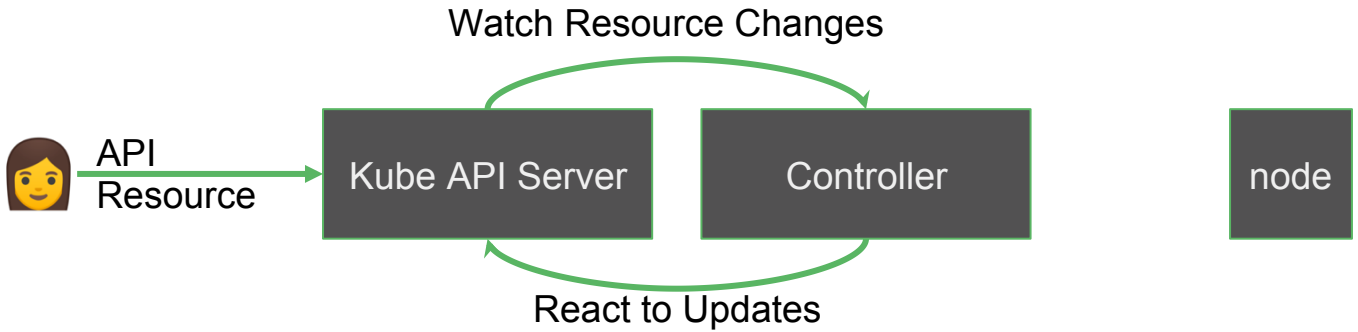
CLOUD NATIVE
COMPUTING FOUNDATION

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Life Cycle

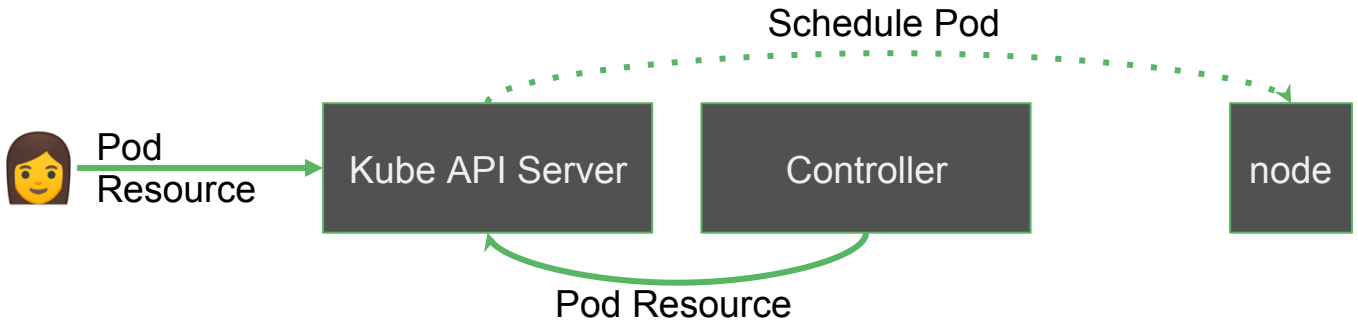


TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Life Cycle II



TechAtBloomberg.com

Bloomberg

Engineering

Deployment Resource (YAML)

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-deployment
spec:
  replicas: 3
  template:
    spec:
      containers:
      - name: nginx
        image: nginx:1.7.9
        ports:
        - containerPort: 80
```

https://kubernetes.io/docs/concepts/workloads/controllers/deployment/

Deployment Resource: Kind

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-deployment
spec:
  replicas: 3
  template:
    spec:
      containers:
      - name: nginx
        image: nginx:1.7.9
        ports:
        - containerPort: 80
```

https://kubernetes.io/docs/concepts/workloads/controllers/deployment/

TechAtBloomberg.com

Bloomberg

Engineering

Deployment Resource: Spec (Replicas)

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-deployment
spec:
  replicas: 3
  template:
    spec:
      containers:
      - name: nginx
        image: nginx:1.7.9
        ports:
        - containerPort: 80
```

https://kubernetes.io/docs/concepts/workloads/controllers/deployment/

Bloomberg

Engineering

Deployment Resource: Spec (Pod Template)

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-deployment
spec:
  replicas: 3
  template:
    spec:
      containers:
      - name: nginx
        image: nginx:1.7.9
        ports:
        - containerPort: 80
```

https://kubernetes.io/docs/concepts/workloads/controllers/deployment/

TechAtBloomberg.com

Bloomberg

Engineering

Benefits of Kubernetes

- Declarative deployment
- Hardware indifference
- Monitoring
- Replication
- Autoscaling
- Security Extensions
  - Ingress + SSL termination
  - Istio service mesh

Custom Resources

- **Resource** - a set of API objects in Kubernetes like Deployments or Pods.

- **Custom Resource** - an extension of the Kubernetes API defined by a user that declares new types of API Objects.

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Kubernetes Meets ML: TensorFlowJob

```
apiVersion: ds.bloomberg.com/v1
kind: TensorFlowJob
metadata:
  name: tf-test
  annotations:
    ai.bloomberg.com/project: foo
    ai.bloomberg.com/experiment: abcde
spec:
  framework: tensorflow-1.7-python-3
  identities:
    - hadoop:
        id: keithlaban
  pipPackages: [ai.bloomberg.com.myteam.gpu_tftraining, numpy]
  module: gpu_tftraining
  size: GpuLarge
  args:
    --data-dir hdfs://CLUSTER/projects/news/news_index
    --output-dir hdfs://CLUSTER/users/klaban1/abcde/1
```

TechAtBloomberg.com

Bloomberg

Engineering

Achieving Nirvana (Reprise)

Foundational Building Blocks

- Specialized Hardware
- Standardized Runtimes
- Batch Scheduling

Data Liquidity

- Data Access
- Security

Developer Experience

- Ease of Use
- Rapid Iteration
- Native Tooling

ML Specifics

- Experiment Management
- Hyperparameter Tuning
- Result Reproducibility

A *Model* for the ML Ecosystem

- Other Types of Custom Resources
  - Spark
  - Python
  - JVM
  - Jupyter
  - Hyperparameter Tuning
- Foundation for reproducibility and automation

Bloomberg

Engineering

Open FinTech Forum 2018

Concluding Remarks

TechAtBloomberg.com

Bloomberg

Engineering

In Conclusion

- Opportunity for Machine Learning
- Allow ML Practitioners to focus on ML
- Kubernetes is strong in the community!

TechAtBloomberg.com

Bloomberg

Engineering

Thanks!

Panel Discussion: Real-World Kubernetes Use Cases in Financial Services: Lessons Learned from Capital One, BlackRock and Bloomberg - Moderated by Ron Miller, TechCrunch

Thursday @ 4:25 PM

http://sched.co/G4Mx

TechAtBloomberg.com

Bloomberg

Engineering