

Machine Learning @ Bloomberg

Building on Kubernetes

GPU Technology Conference (GTC) 2019 [S9810]
March 19, 2019

Ian Hummel, Data Architect, Office of the CTO
David Eis, Senior Software Engineer, AI Group

[TechAtBloomberg.com](https://techatbloomberg.com)

© 2019 Bloomberg Finance L.P. All rights reserved.

Engineering

Bloomberg

Agenda

- Finance and technology
- Machine Learning at Bloomberg
- Data Science Platform infrastructure
- Results and future work

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering



Bloomberg

[TechAtBloomberg.com](https://www.techatbloomberg.com)

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

About Bloomberg

- Founded in 1981
- 325,000 Terminal subscribers
- Customers in 170 countries
- Over 19,000 employees in 192 locations
- Over 5,000 software engineers/programmers
- More reporters than The New York Times + Washington Post + Chicago Tribune

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.





Finance and technology

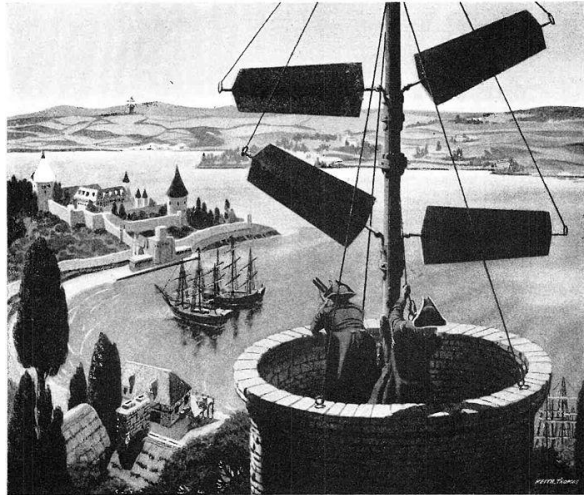
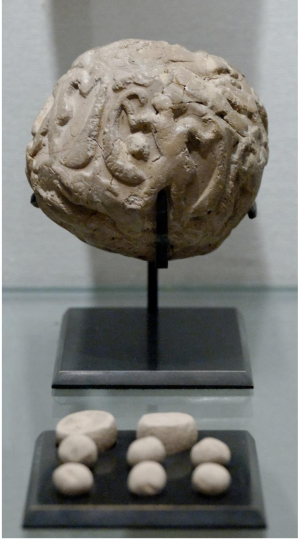
TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

Finance and technology



[https://en.wikipedia.org/wiki/Bulla_\(seal\)](https://en.wikipedia.org/wiki/Bulla_(seal))

https://en.wikipedia.org/wiki/Semaphore_telegraph

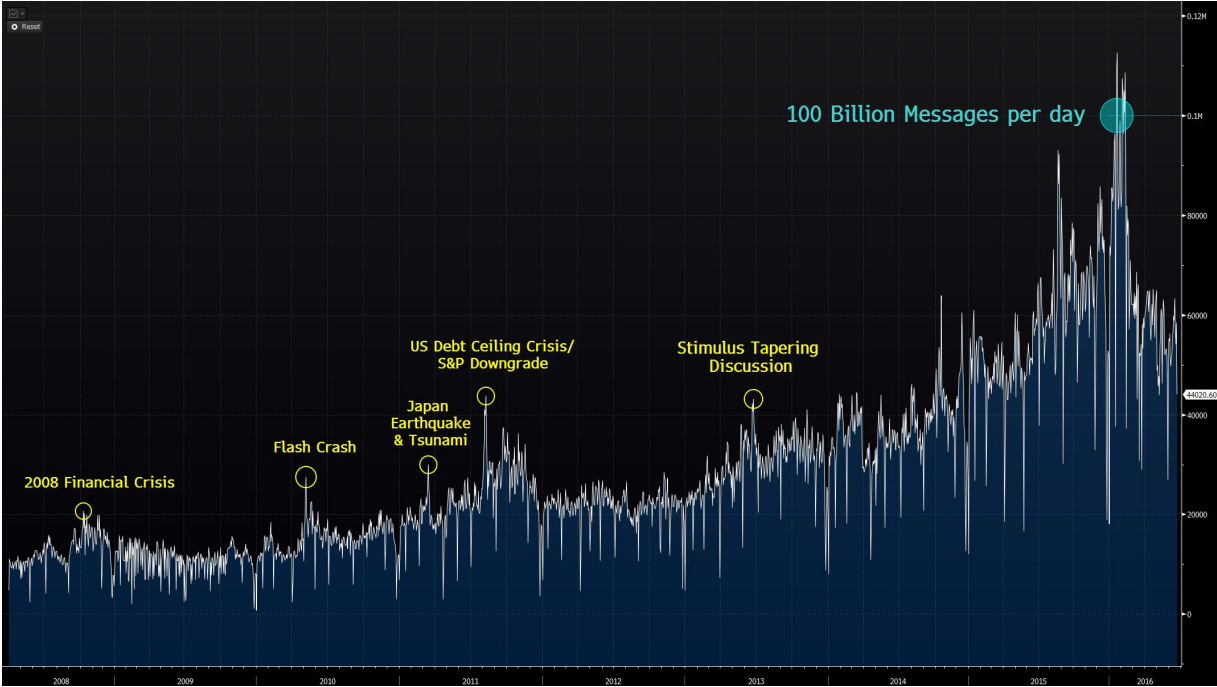
TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

Finance and technology today



TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

Understanding what's going on out there



The screenshot shows a Bloomberg News search results page for the keyword 'Coffee'. The page is divided into two main sections: 'Top Ranked News' and 'Time Ordered News'. The 'Top Ranked News' section lists three items, and the 'Time Ordered News' section lists 19 items. Each item includes a rank, a title, a source code, and a time.

Rank	News Title	Source	Time
1)	Engelhart Hires Former Sucden's Donovan to Bolster Coffee Team	BN	12:10
2)	SOFTS: Cocoa Advances on El Nino Concerns; Coffee, Sugar Climb	BFW	10/25
3)	Tata Coffee Earnings Teleconference TCO IN	BT	05:19
Time Ordered News			
4)	*SEC Short Sale Rule 201 is in Effect : YGYI (NASDAQ)	BSS	12:10
5)	Engelhart Hires Former Sucden's Donovan to Bolster Coffee Team	BN	12:10
6)	Mena Report: Tender For Sale Of Dried Coffee Berries.	NS8	11:36
7)	Mena Report: Supply Of Coffee.	NS8	09:45
8)	Mena Report: Coffee, Tea And Accompanying Products.	NS8	09:28
9)	The Columbian: Food & Drink: Kafiex celebrates global coffee village	NS1	09:09
10)	Mena Report: Food, Beverages, Tobacco And Related Products (coffee Beverage, ...	NS8	09:05
11)	Stewart & Mackertich Research_WEEKLY Dossier 26-10-2018	SMK	08:46
12)	Tuoi Tre: Tea-based drinks outplay coffee in caffeine-loving Vietnam: report	NS6	08:18
13)	Intl Coffee Org: Visit from Ambassador of Angola	IN9	06:51
14)	AGRICULTURE WEEKLY AGENDA: Sugar Ethanol; Bali; Soy Crush; Bunge	BFW	05:45
15)	Asian News: GreenBrrew all set to dominate Indian market with flavored green c...	RE6	05:28
16)	Tata Coffee Earnings Teleconference TCO IN	BT	05:19
17)	Commodity Online: India Coffee exports may fall as floods dent production	NS1	04:19
18)	*SEC Short Sale Rule 201 Not In Effect : YGYI (NASDAQ)	BSS	03:58
19)	Tata Coffee Earnings Q2 2019 Earnings Call Teleconference TCO IN	BT	02:38

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

Natural Language Processing & Events



TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

The risks of being wrong



<https://www.bloomberg.com/news/articles/2013-04-23/a-fake-ap-tweet-sinks-the-dow-for-an-instant>

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

The future is ripe for ML

All Companies | **Set Alert** | **Set View as Default** | **News Trends**

Trend Period: 8 Hours

News Reader Activity | News Sentiment | News Volume | Twitter Volume

Most Negative | Most Positive

Security	Sentiment	GN	Δ Price	Δ AVAT	Representative Headline
1) Forward Pharma A/S	-	-0.99	-3.30%	+67.00%	Forward Pharma Falls as Analysts Downg...
2) Acorda Therapeutics Inc	-	-0.95	-8.57%	-46.20%	Acorda on the Ropes After Amovra Paten...
3) Arista Networks Inc	-	-	-	-	-
4) Imagination Technologies ...	-	-	-	-	-
5) AutoZone Inc	-	-	-	-	-
6) Saudi Electricity Co	-	-	-	-	-
7) Interactive Brokers Group ...	-	-	-	-	-
8) Unitech Ltd	-	-	-	-	-
9) Nordic American Offshore L...	-	-	-	-	-
10) O'Reilly Automotive Inc	-	-	-	-	-
11) Under Armour Inc	-	-	-	-	-
12) Orient Overseas Internatio...	-	-	-	-	-
13) Biogen Inc	-	-	-	-	-
14) ReliWalk Robotics Ltd	-	-	-	-	-
15) Pricaline Group Inc/The	-	-	-	-	-
16) KazMunaiGas Exploration P...	-	-	-	-	-
17) Sartorius Stedim Biotech	-	-	-	-	-
18) Travel Expert Asia Enterpri...	-	-	-	-	-
19) Cinemark Holdings Inc	-	-	-	-	-
20) Caterpillar Inc	-	-	-	-	-
21) Kohl's Corp	-	-	-	-	-

Australia 61 2 9777 0600 Brazil 5511
Japan 81 3 3201 8900 Singapore

Twitter Publication Count - Realtime (1.1) 44.00
Last Price (K2) 38.20

CN News ☆

BAC US Equity \$↑24.23 -0.35 (-1.42%) N 24.20 / 24.22 K High 24.35 Open 24.12

Bank of America Corp At 16:15 d Volume 4

Search News Sources Display & Edit Custom

Bank of America Corp <Narrow More>

Match Bank of America Corp

1) Wall Street Journal: How Donald Trump r
2) Wall Street Journal: Warren Buffett, long
3) BI Company Primer: Bank of America
4) ABC 4: Judge won't dismiss Counts 1 & 1
5) Phil van dyke: \$BAC And if you get in, H
6) *MARKET ON CLOSE IMBALANCE BUY 2260!
7) Inside Mtg [Reg]: Banks' Mortgage-Bank
8) Deseret News: Judge declines to toss ra
9) Options Clearing: PRO Listings -BANK OF
10) Seeking Alpha: New: The Market Might B
11) RT @WSJ: Warren Buffett, long a critic of
12) Bank of America Corp.'s 6 7/8% Capital s
13) Salt Lake Trib: Swallow trial: Judge den
14) Phil van dyke: \$BAC We know that DT wi

SoFi Raises \$500 Million Led by Silver Lake for Global Expansion

- Funds will be used for new products and international growth
- The investment is said to value SoFi at \$4.3 billion

By Selina Wang

(Bloomberg) -- **Social Finance Inc.** said it raised \$500 million in a financing round led by private equity firm **Silver Lake Partners**. It plans to use the funds for international expansion of its online lending business and development of new financial products.

Other investors in the round include **SoftBank Group Corp.** and **GPI Capital**, and the funding brings SoFi's total investment to \$1.9 billion, the company said. Several sovereign wealth funds from countries in Asia and Europe invested as well, said **Mike Cagney**, SoFi's chief executive officer and co-founder. The international group will purchase SoFi's loans in addition to taking an equity stake, he said.

"Our challenge has always been our reliance on wholesale capital markets," Cagney said. "Silver Lake is not a balance sheet partner, but their investor base is hugely important to us, and there's a big overlap between their limited partners and our loan buying universe."

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering



Machine Learning at Bloomberg

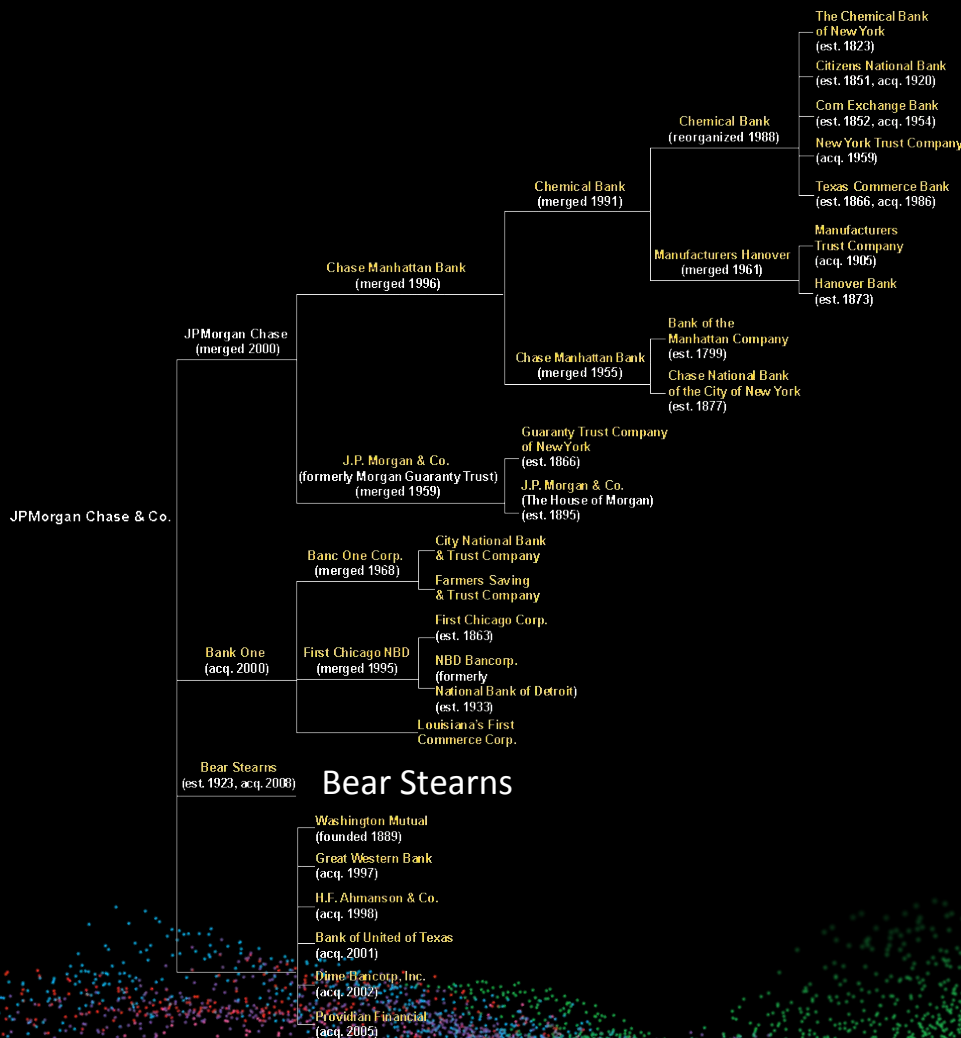
TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

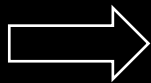
Engineering

Financial named entity extraction



TechAtBloomberg.com

Explaining Knowledge Graph relationships



“In 1971, Siegl, Bowker and Baldwin established the Starbucks Coffee Company.”

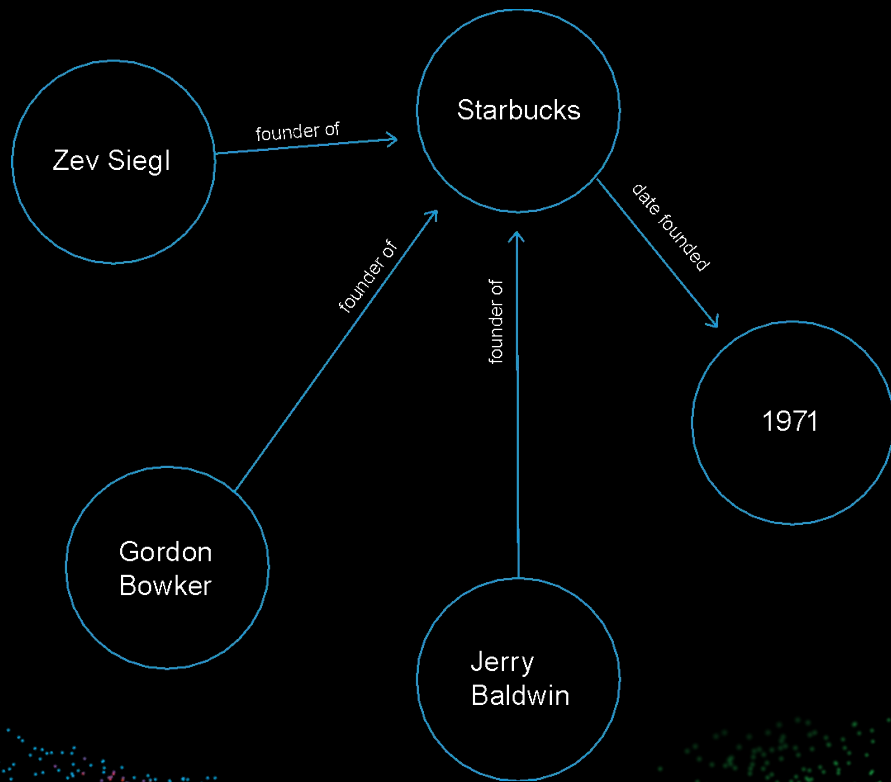


Table Extraction

NVIDIA CORPORATION AND SUBSIDIARIES
CONSOLIDATED STATEMENTS OF INCOME
(In millions, except per share data)

	Year Ended		
	January 27, 2019	January 28, 2018	January 29, 2017
Revenue	\$ 11,716	\$ 9,714	\$ 6,910
Cost of revenue	4,545	3,892	2,847
Gross profit	7,171	5,822	4,063
Operating expenses			
Research and development	2,376	1,797	1,463
Sales, general and administrative	991	815	663
Restructuring and other charges	—	—	3
Total operating expenses	3,367	2,612	2,129
Income from operations	3,804	3,210	1,934
Interest income	136	69	54
Interest expense	(58)	(61)	(58)
Other, net	14	(22)	(25)
Total other income (expense)	92	(14)	(29)
Income before income tax	3,896	3,196	1,905
Income tax expense (benefit)	(245)	149	239
Net income	\$ 4,141	\$ 3,047	\$ 1,666
Net income per share:			
Basic	\$ 6.81	\$ 5.09	\$ 3.08
Diluted	\$ 6.63	\$ 4.82	\$ 2.57
Weighted average shares used in per share computation:			
Basic	608	599	541
Diluted	625	632	649

NVDA US \$ ↓ 148.25 -1.01 7148.25 / 148.29Q 1x1
At 11:12 d Vol 3,856,684 0 145.66Q H 149.33Z L 144.80Q Val 618.983

NVIDIA US Equity 90 Actions 97 Export 90 Settings Financial Analysis

NVIDIA Corp BQL Periods 10 Annuals Cur FRC (USD)

1 Key Stats 2 I/S 3 B/S 4 C/F 5 Ratios 6 Segments 7 Addl 8 ESG 9 Custom 10 Shared

11 Adj Highlights 12 GAAP Highlights 13 Earnings 14 Enterprise Value 15 Multiples 16 Per Share 17 Stock Value

In Millions of USD except Per Share

	2016 Y 01/31/2016	2017 Y 01/29/2017	2018 Y 01/28/2018	2019 Y 01/27/2019	2020 Y Est 01/31/2020	2021 Y Est+ 01/31/2021
Revenue						
Consensus Estimate	4,922.0	6,849.6	9,473.6	11,712.3	11,181.8	13,400.6
Comparable Actual	5,010.0	6,910.0	9,714.0	11,716.0		
Revenue Surprise %	1.8	0.9	2.5	0.0		
GAAP Actual	5,010.0	6,910.0	9,714.0	11,716.0		
Adjusted Actual	5,010.0	6,910.0	9,714.0	11,716.0		
Earnings Per Share						
Consensus Estimate	1.58	2.94	4.47	6.60	5.36	7.13
Comparable Actual	1.67	3.06	4.92	6.64		
EPS Surprise %	5.9	4.1	10.1	0.7		
GAAP Actual	1.08	2.57	4.82	6.63		
Adjusted Actual	1.36	2.60	4.63	6.10		
EBIT						
Consensus Estimate	1,055.6	2,157.6	3,361.6	4,406.3	3,717.2	5,028.8
Comparable Actual	1,125.0	2,221.0	3,617.0	4,407.0		
EBIT Surprise %	6.6	2.9	7.6	0.0		
GAAP Actual	747.0	1,934.0	3,210.0	3,804.0		
Adjusted Actual	968.0	1,953.0	3,224.0	3,850.0		

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

Automating the news

Starbucks Shares Fall After Chain's U.S. Sales Disappoint Again

By Leslie Patton

DEEP ECO: Starbucks Pressured, Industry Sales Slowest in 8 Yrs

By Bloomberg Automation [Feedback?](#)

(Bloomberg) -- Starbucks Corp.'s revenue may be pressured after a U.S. Census Bureau report showed limited-service eating places nominal retail sales fell 0.1 percent in November, the first decrease in eight years.

- The monthly report may provide an indication of the performance of companies in the industry Bloomberg classifies as limited service restaurants; Starbucks gets 89 percent of its revenue from this industry.
- Analysts estimate the company's revenue will rise 8.2 percent in the current fiscal quarter.
- This retail sales index was up 1 percent in the first two months of the company's current fiscal quarter.
- Starbucks' first quarter started on Oct. 2, 2017.
- In the past six years, Starbucks' revenue has had a 0.7 correlation with this index, according to data compiled by Bloomberg.

SBUX US Equity
Graphic Dashboard»

1M Imp Vol 3M
28.07



The New York Times

By Jaclyn Peiser

Feb. 5, 2019

As reporters and editors find themselves the victims of layoffs at digital publishers and traditional newspaper chains alike, journalism generated by machine is on the rise.

Roughly a third of the content published by Bloomberg News uses some form of automated technology. The system used by the company, Cyborg, is able to assist reporters in churning out thousands of articles on company earnings reports each quarter.

The program can dissect a financial report the moment it appears and spit out an immediate news story that includes the most pertinent facts and figures. And unlike business reporters, who find working on that kind of story a chore, the program's output is a relief.

Subscribe to The New York Times.

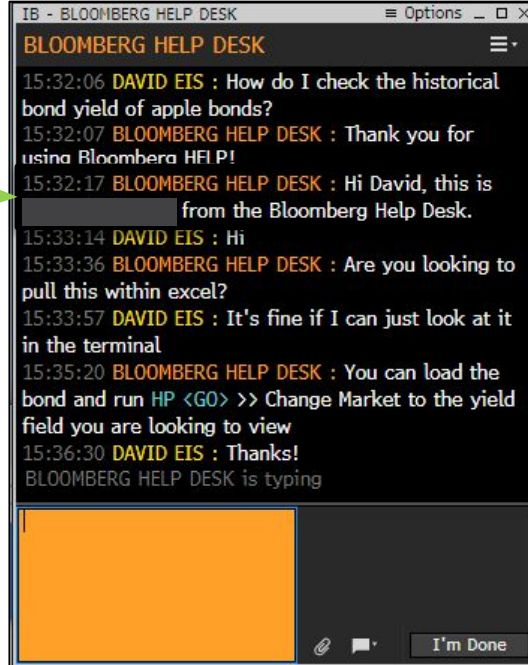
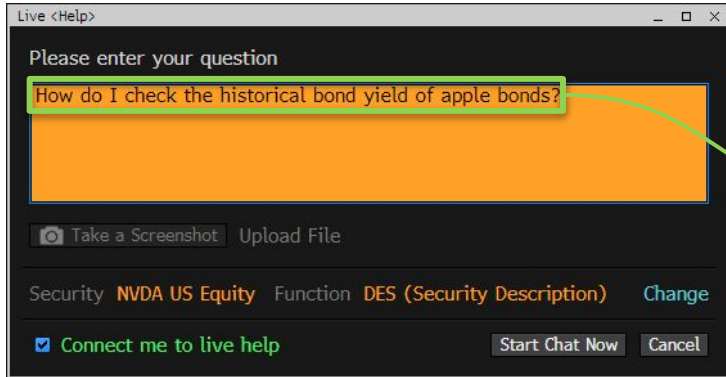
TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

HELP HELP - Query routing



News QA

The screenshot displays the Bloomberg News search interface. At the top, there are two tabs: "Custom Searches" and "News & Research Menu". Below the "Custom Searches" tab, a search bar contains the text "news on brexit from the financial times last thursday". A dropdown menu shows "News Search Options" with the query broken down into individual words: "NEWS", "ON", "BREXIT", "FROM", "THE", "FINANCIAL", "TIMES", "LAST", and "THURSDAY". To the right of this menu are three options: "Smart Match", "Match All Words", and "Exact Match".

Below the search bar, a navigation menu includes "Search News", "Actions", "Custom Searches", and "Translate". The search results are displayed in a list format. The first result is "UK Vote on EU Member..." from FTI on 02/07. The search bar also shows "FTI" and "02/07".

On the left side of the interface, there is a sidebar with various categories and sources:

- 3) READ Most Read News
- 4) NI News Categories
- 5) NH News: All Sources
- 6) Recommended Content >
- 7) FIRS First Word
- 8) BI Bloomberg Intelligence
- 9) BRIE Bloomberg Brief Newsletters
- 10) QUIC Bloomberg QuickTakes
- 11) BIZW Bloomberg Businessweek
- 12) Opinion & Analysis >
- 13) OPIN Bloomberg Opinion
- 14) BIP Bloomberg Intelligence Primers
- 15) BNEF Bloomberg New Energy Finance
- 16) ECO Economics

The search results list includes the following items:

- 1) FastFT: Theresa May pushes back against Tusk's sharply-worded comments FTI 02/07
- 2) FastFT: May and Juncker promise talks to find path through Brexit impasse FTI 02/07
- 3) Mideast investors shun commercial property FTI 02/07
- 4) May and EU agree on talks to break deadlock FTI 02/07
- 5) Plans hatched in secret to revive economy after a no-deal Brexit FTI 02/07
- 6) Brexit/property funds: cash test dummies FTI 02/07
- 7) Mideast investors shun commercial property FTI 02/07
- 8) Brexit/property funds: cash test dummies FTI 02/07
- 9) Mideast investors shun commercial property FTI 02/07
- 10) Brexit/property funds: cash test dummies FTI 02/07
- 11) Brexit/property funds: cash test dummies FTI 02/07
- 12) Brexit/property funds: cash test dummies FTI 02/07
- 13) Mideast investors shun commercial property FTI 02/07
- 14) May and EU agree to reopen talks to break Brexit impasse FTI 02/07
- 15) Jeremy Corbyn's shift on Brexit heightens Labour tensions FTI 02/07
- 16) The Brexit delusion of creating 'Singapore upon Thames' FTI 02/07
- 17) Bank of England pulls back on multiple rate rise plans FTI 02/07
- 18) FastFT: Kellogg swings to fourth-quarter loss as Brexit planning costs bite FTI 02/07
- 19) UK draws up secret plan to boost economy after no-deal Brexit FTI 02/07

Charts QA



TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering



Why build a Data Science Platform?

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

ML opportunities

What questions do teams ask before starting a new project?

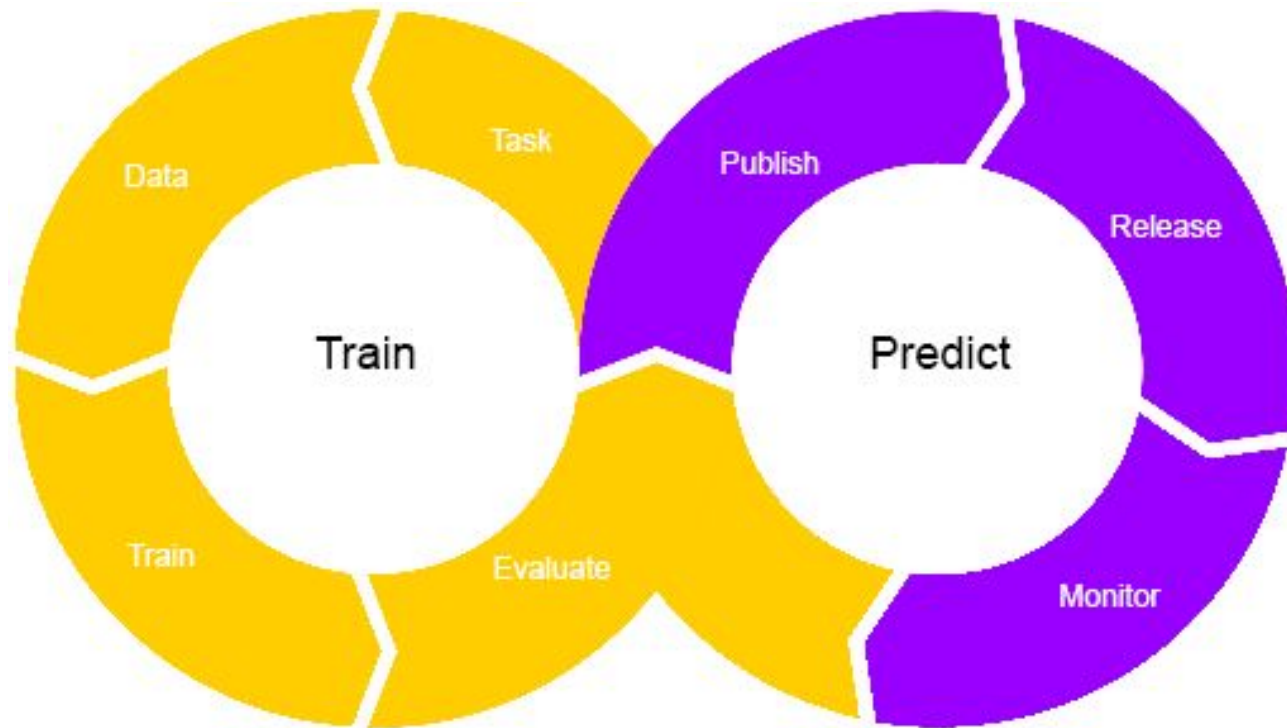
- What **data** is relevant to my business problem?
- What does the data **look** like?
- What **features** should I create from my data?
- What metric should I use to **measure** my model performance?
- What **algorithms** or **parameters** should I test out?
- How is my model doing on live data in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

Challenges for ML projects

Even in an established enterprise, ML teams face the same challenges again and again...

- 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 model development lifecycle



TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

Different roles, different priorities

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

Engineers

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

Customers

- Smarter Functionality
- Privacy
- Correctness
- Uptime/Stability

Blueprint for an ML infrastructure...

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





Kubernetes to the rescue!

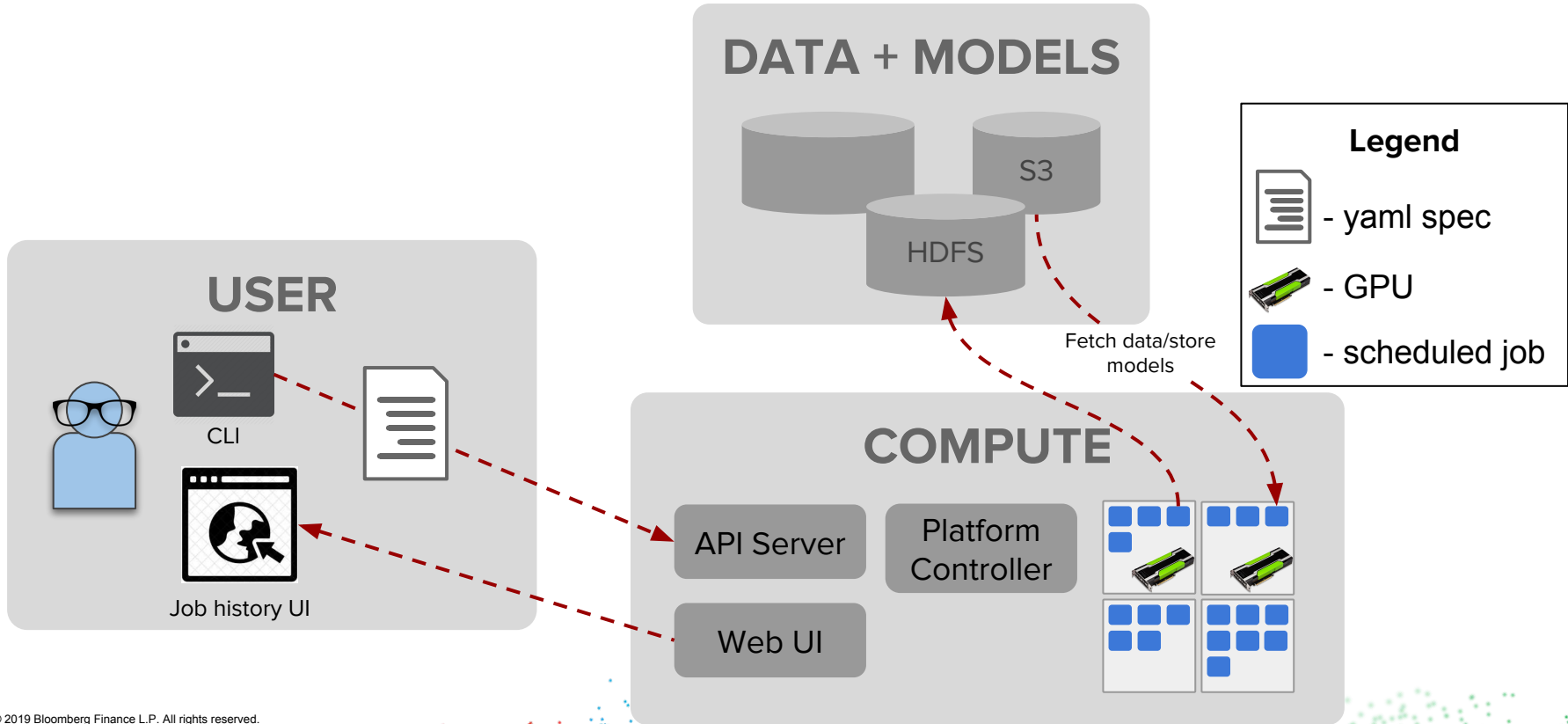
TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

DS Platform architecture

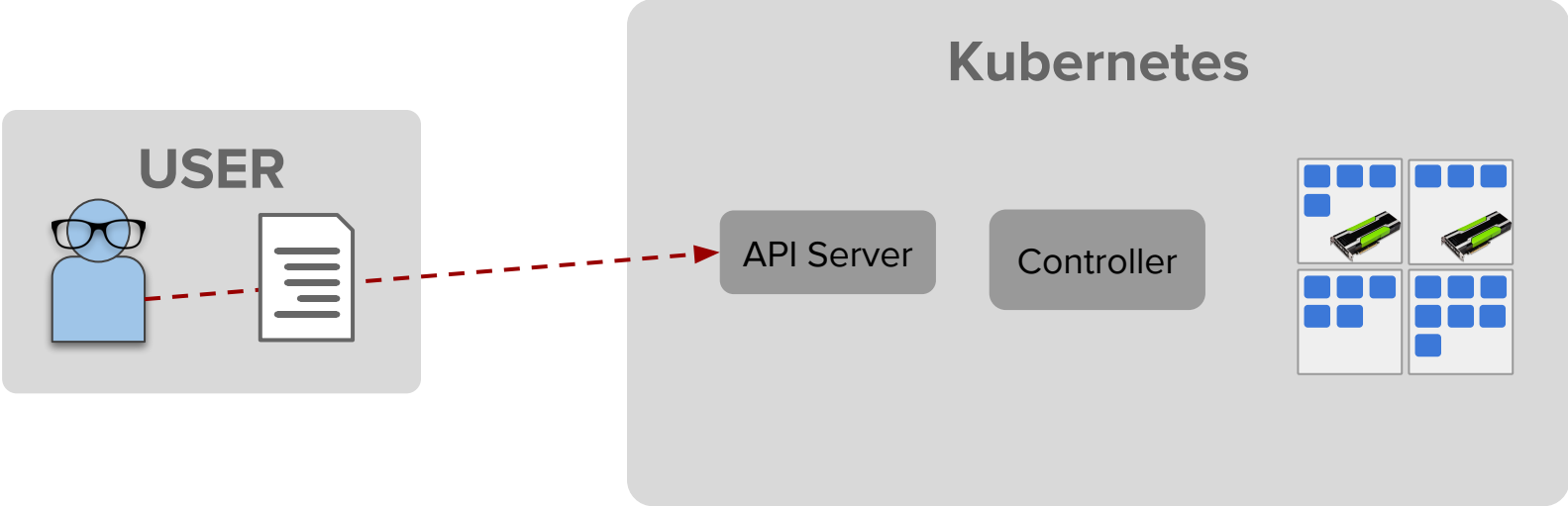


What is Kubernetes?

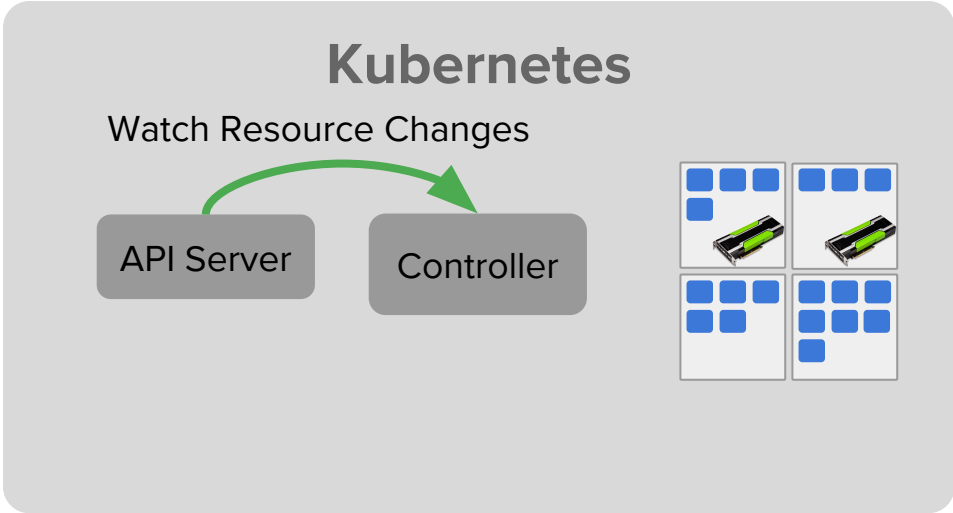
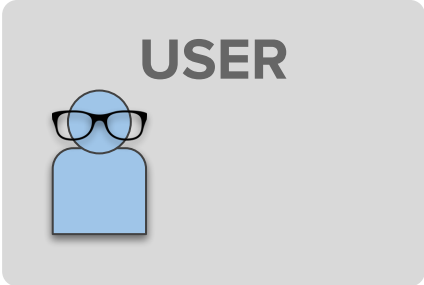
“Kubernetes is an open-source system for automating deployment, scaling, and management of containerized applications.”

- <https://kubernetes.io/>

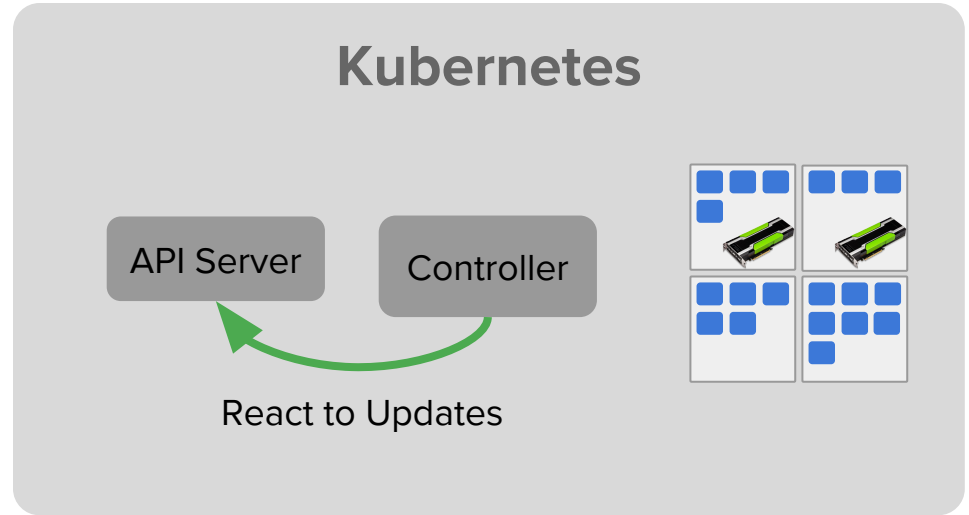
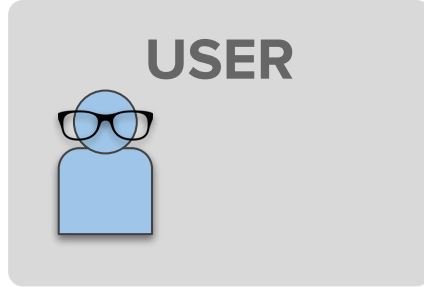
Kubernetes life cycle



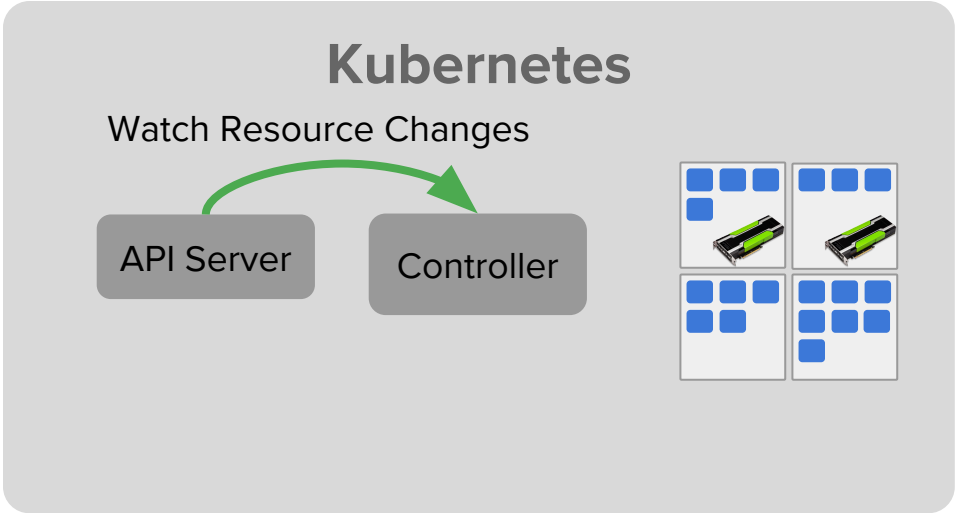
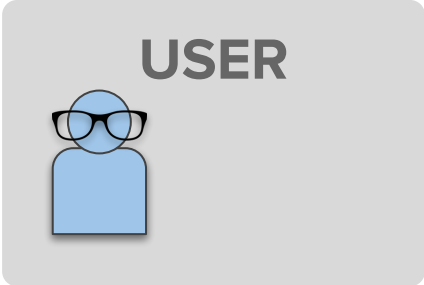
Kubernetes life cycle



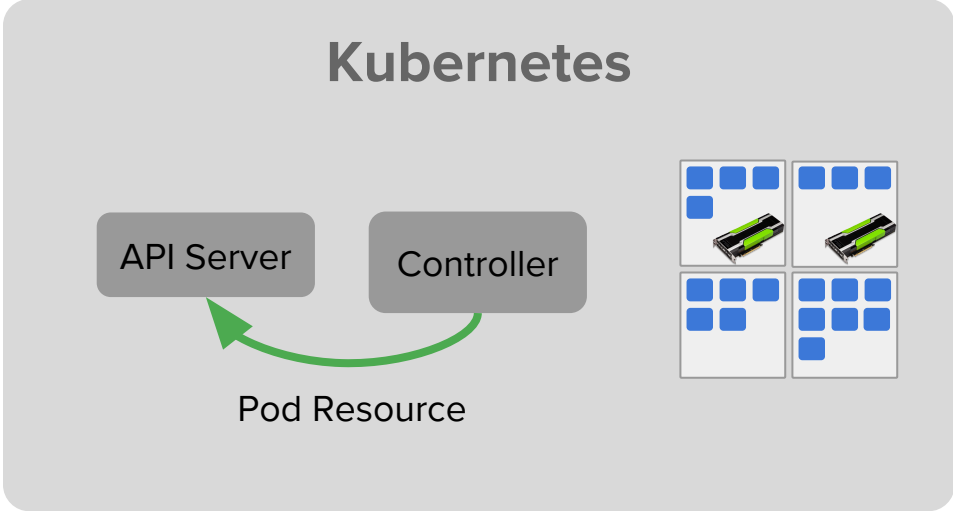
Kubernetes life cycle



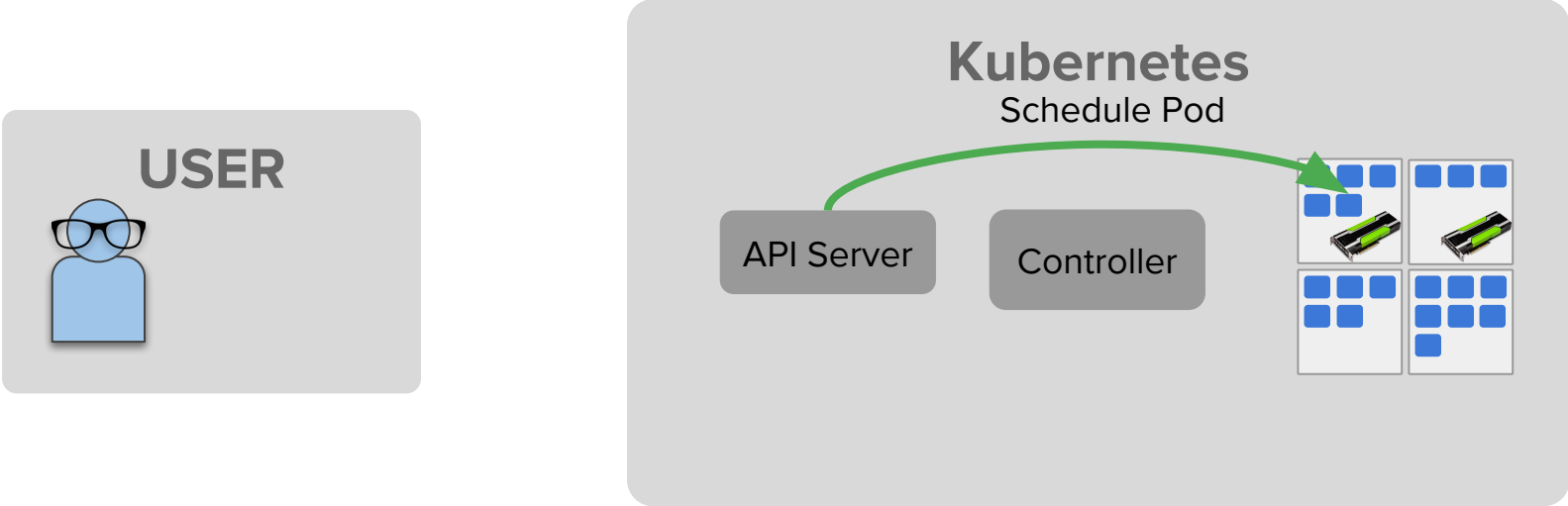
Kubernetes life cycle



Kubernetes life cycle



Kubernetes life cycle



Deployment resource (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/>

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

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/>

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/>

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

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

© 2019 Bloomberg Finance L.P. All rights reserved.

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, such as 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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/fancy_model
```

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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/abcde/foo/1
```

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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/abcde/foo/1
```

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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/abcde/foo/1
```

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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/abcde/foo/1
```

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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/abcde/foo/1
```

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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/abcde/foo/1
```

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-1.13-python-3
  identities:
    - hadoop:
      id: davideis
  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/deis/abcde/foo/1
```

Bloomberg

Engineering

A model for the ML ecosystem

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

Tooling

- Home
- Cluster Status
- Jobs
 - PythonJob
 - TensorFlowJob
 - SparkJob
 - JVMJob
 - JupyterNotebook
 - HyperTune
 - Launch Notebook
 - Documentation

Bloomberg Data Science Platform

deis ds-ob-dev01 s-mlplatform Logout



TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

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





Helping teams succeed with the DS Platform

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

Building infrastructure is not for the faint-hearted

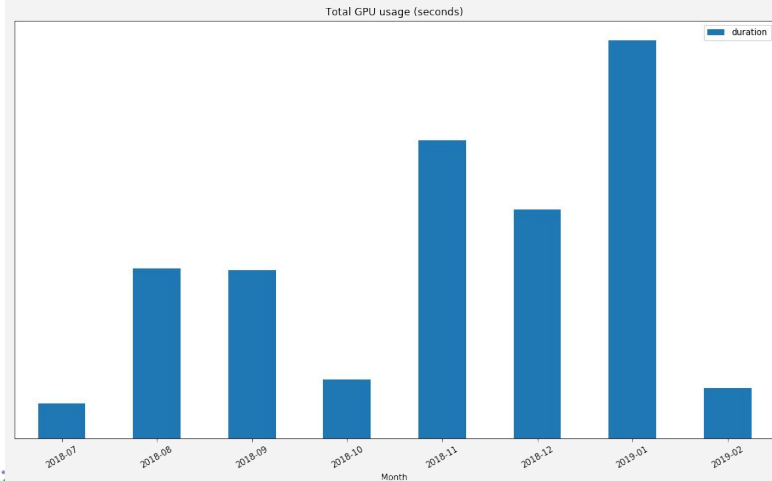
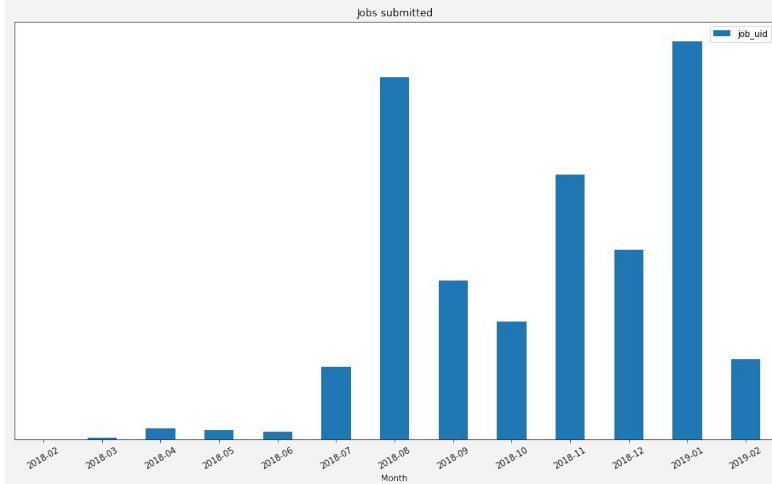
- Demanding users
- Multi-tenancy is hard
- New hardware designs hard to incorporate
- ML projects require consulting
- Effective tooling means team requires broad skill set
- Open source moving incredibly fast
- What's your value-add??

Strong emphasis on metrics

- Number of teams using the platform
- Number of users using the platform
- How often they submit jobs
- Duration of jobs
- Types of jobs submitted
- Hardware utilization
- Net Promoter Score (NPS) surveys

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.



What's next for us?

- More tooling
- Kubernetes batch/job queueing
- KNative serving
- Dataset and model management
- Additional ML services like data drift detection

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering



Thanks!

Reach out to us at

ihummel@bloomberg.net and deis@bloomberg.net

TechAtBloomberg.com

© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering