

# How a centralized Telemetry team drives value @ Bloomberg

GrafanaCon LA 2019  
February 25, 2019

Sean Hanson, Software Developer  
Stig Sorensen, Manager Production Visibility

TechAtBloomberg.com

# Sean Hanson and Stig Sorensen



TechAtBloomberg.com

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

Bloomberg  
Engineering

# Bloomberg in a Nutshell



The **Bloomberg Terminal** delivers a diverse array of information on a single platform to facilitate financial decision-making.

[TechAtBloomberg.com](http://TechAtBloomberg.com)

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

**Bloomberg**  
Engineering

# Bloomberg by the Numbers

- Founded in **1981**
- **325,000+** subscribers in **170 countries**
- Over **19,000 employees** in 192 locations
- **More journalists** than *The New York Times* + *Washington Post* + *Chicago Tribune*

TechAtBloomberg.com

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

Bloomberg  
Engineering

# Bloomberg Technology by the Numbers

- **5,000+** software engineers
- **150+** technologists and data scientists devoted to machine learning
- One of the largest private networks in the world
- **120 billion** pieces of data from the financial markets each day, with a peak of more than 10 million messages/second
- **2 million** news stories ingested / published each day (500+ news stories ingested/second)
- News content from **125K+** sources
- Over **1 billion** messages and Instant Bloomberg (IB) chats handled daily

[TechAtBloomberg.com](http://TechAtBloomberg.com)

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

**Bloomberg**  
Engineering

# People care when Bloomberg doesn't work



Lou Whiteman

Millions of traders are looking out the window for the first time in years.

#BloombergDown

2

20

30



Owen Callan

@OwenCallan



Right now there's a very sad Bloomberg algo that wants to send out a \*BLOOMBERG DOWN GLOBALLY\* flash headline, but just can't get the irony



Philip O'Sullivan

@pdosullivan

Bloomberg being down has finally made me understand how teenagers feel when Facebook crashes.



TechAtBloomberg.com

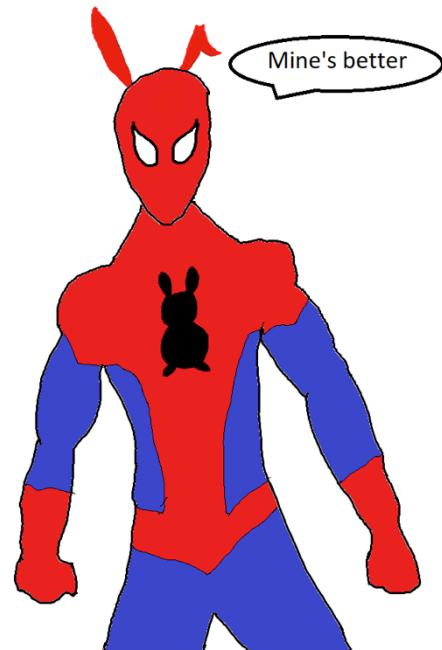
© 2019 Bloomberg Finance L.P. All rights reserved.

Bloomberg

Engineering

# Origin Story

- Common circumstances behind outages
- Multiple teams receive (seemingly) unrelated alerts
- Each team consults different telemetry systems
- Lingers until multiple tickets escalate to “Major Outage”
- Post-mortems create \*another\* data source

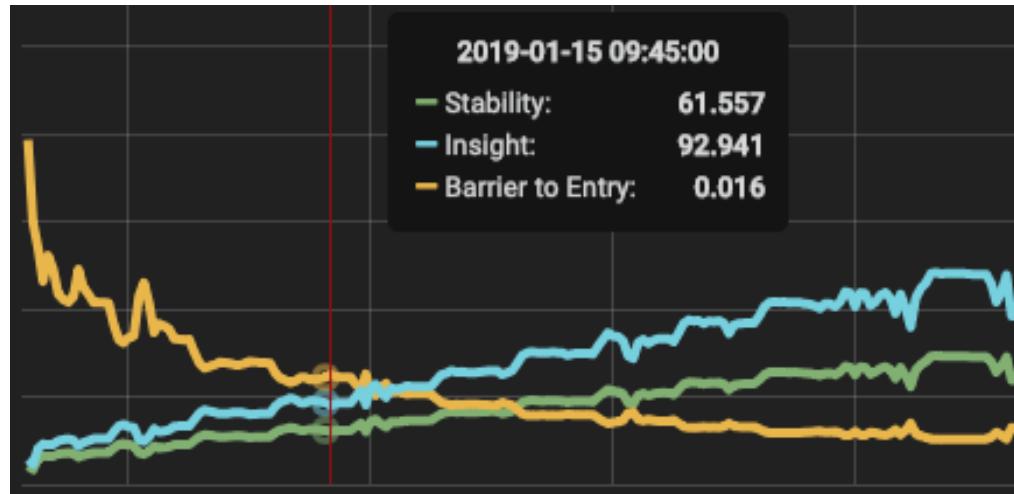


TechAtBloomberg.com

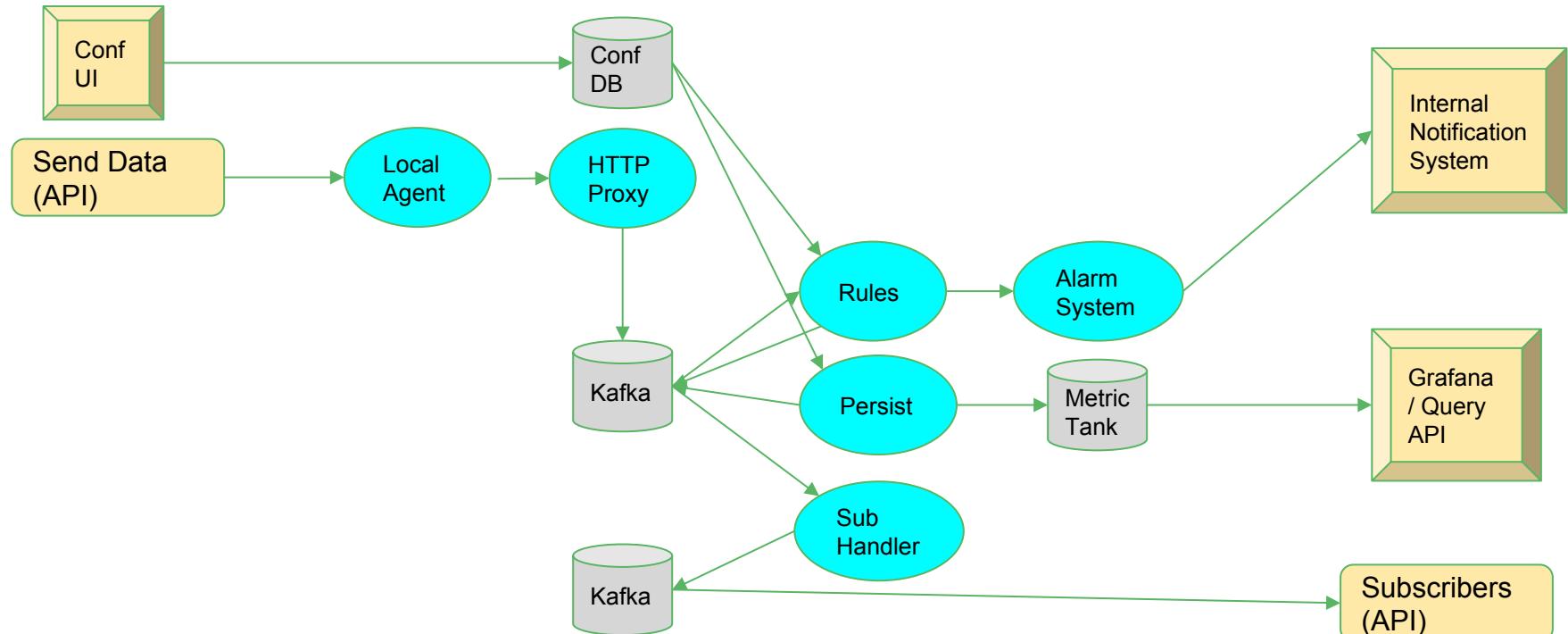
© 2019 Bloomberg Finance L.P. All rights reserved.

# Telemetry team objectives

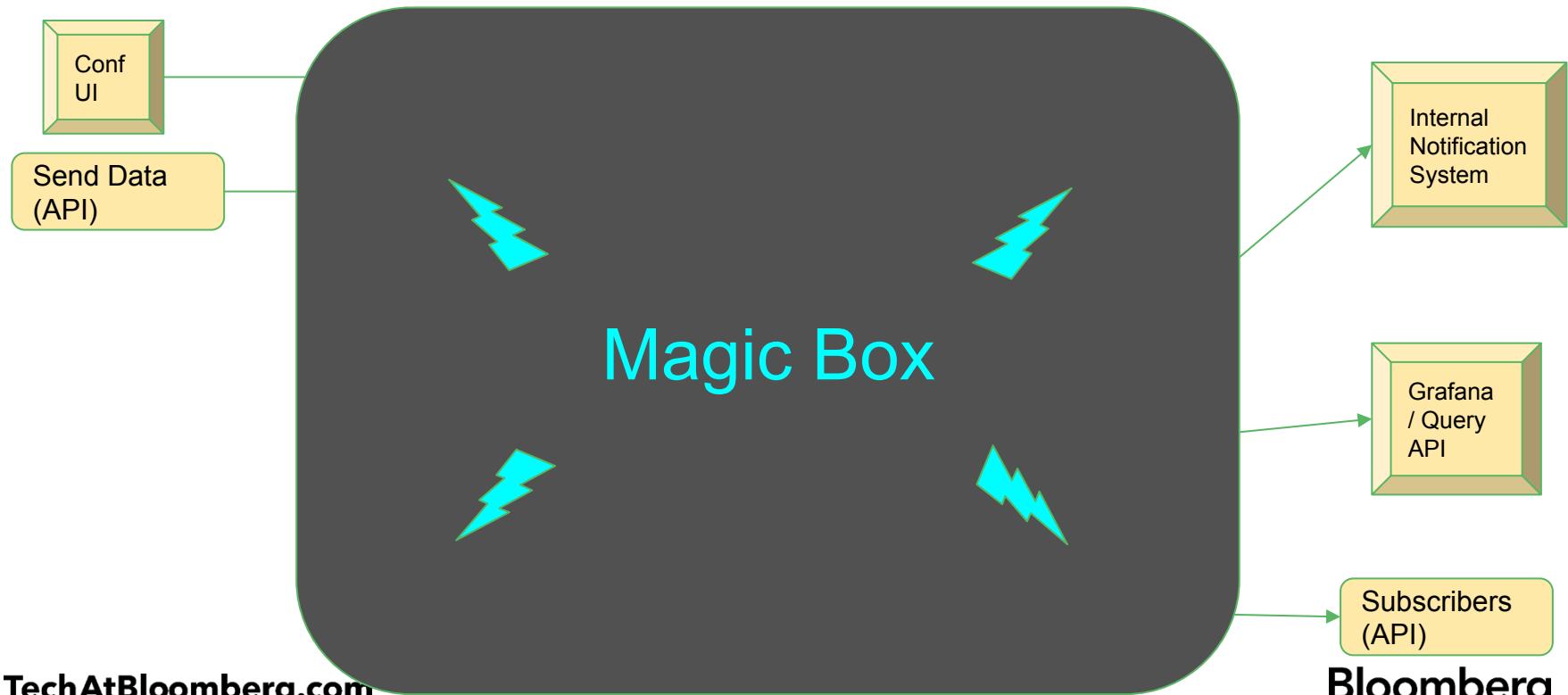
- Boost insights
- Improve stability
- Reduce barrier to entry
- Drive best practices



# Internal Architecture (simplified)



# Internal Architecture (simplified)



# Boost Insight

- Collect “free” metrics
  - OS metrics
  - Common frameworks / infrastructure
  - Entire process table
- Generic dashboards for these metrics
- High-level system health dashboards
  - Great for project owners and managers
  - Drilldowns for SREs / Devs
- Query API
  - Let users extract their own insights

TechAtBloomberg.com

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

Bloomberg  
Engineering

# Improve Stability

- Machines monitored from creation
  - Even the machine creation process can publish metrics!
- Alert on well-known metrics by default
- Common frameworks and infrastructure
  - e.g., Alert on services that \*always\* have outstanding requests
- Alerts on processes using significant resources
- Unified Alert system
  - One source for configurations
  - Easier to find related alerts
  - Faster root cause analysis

# Drive Best Practices

- APIs enforce consistency, while documentation explains
  - When and how to use tags
  - Aggregation vs. sampling
  - Data rollups
- Meet with other teams regularly
  - Really listen to what they need
  - Guide them to solutions
  - Fill in infrastructure gaps where needed

TechAtBloomberg.com

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

Bloomberg  
Engineering

# Reduce Barrier to entry

- One source for visualizations
  - Application / System / Infrastructure dashboards all live together
  - Templates for custom dashboards
  - Queries are the same as the programmatic interface
- Simple API for publishing
- Single place for configuring...everything
- Alerts have a consistent look and feel
  - Allows for “reflex” building

# Telemetry Status

- Metrics
  - 30K+ monitored machines
  - 5M data points/sec
    - Doubled in 9 months
  - 200M active time series
  - 2,500 metrics rules
- Logs
  - 7M lines/sec (2.5GB/s)
  - 21,000 Log Rules
    - 100 Million regex/sec
  - 40 TB/day persisted
    - Multiple stores
- Grafana
  - 2,000+ unique users a week
  - 3,500+ dashboards in 500+ folders
  - 75 queries/sec

TechAtBloomberg.com

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

Bloomberg  
Engineering

# The Cost of “Free”

- Growing dimensionality
- Scaling pains
- Discoverability

TechAtBloomberg.com

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

Bloomberg  
Engineering

# Growing Dimensionality

- Users want increased ability to drilldown
- Various frameworks cause “series churn”
  - Kubernetes
  - Hbase
  - Elastic scaling
- More time-series == more RAM

TechAtBloomberg.com

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

Bloomberg  
Engineering

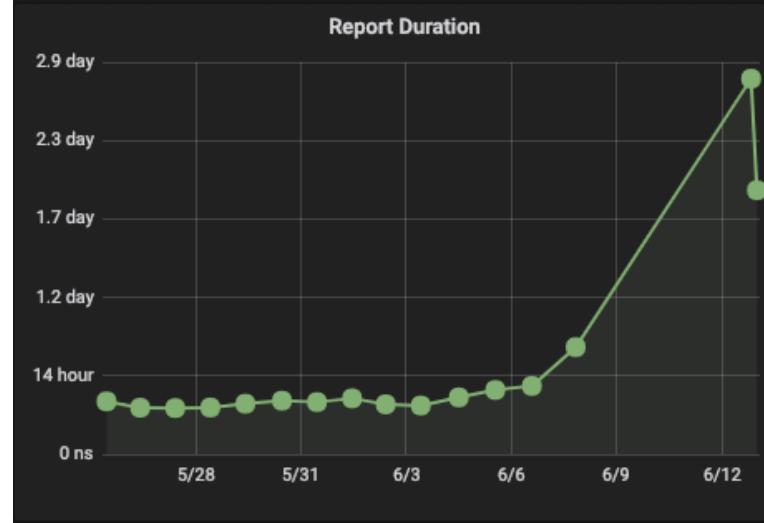
# Growing Dimensionality - Solution

- MetricTank allows pattern based retention and pruning rules
- Users pick their flavor
  - Short lived - store more time-series for shorter time
  - Long lived - store fewer time-series for longer time
  - Default - somewhere in between
- Automatically tag data with a policy
  - MetricTank does the rest

Metrics Namespace: bachHbaseRegion	
<a href="#">Create Metric Rule</a>	<a href="#">Create Monitored Process Configuration</a>
<a href="#">Edit</a>	<a href="#">Delete</a>
Description	HBase Region Metrics
Machines	705
Time Series	<a href="#">7,940,340</a>
Maximum Time Series	12,000,000
Retention Policy <a href="#">?</a>	<u>Short-lived</u>
Tags	N.A.

# Scaling Pains

- MetricTank slowness
- Autocomplete noticeably slow
- Releasing Query API really exposed it
  - Got worse with volume
  - Programs are more sensitive to latency than humans
- One report went from 9 hour runtime to 2 days
  - Oh, and it was a daily report ¯\\_(ツ)\_/¯



# Scaling Pains - Solution

- GC pauses exacerbating a “slowest member” problem
  - We run 120 shard groups with 2 replicas
  - Each render request needed to talk to 119 peers
- Implemented “Speculative Querying”
  - When a slow peer is detected, issue a duplicate query to a replica peer
  - Take the first to respond
  - “Win” rate of 65-75% under normal load
  - P90 render latencies 290ms -> 12ms
- Implemented Graphite functions natively
  - 3x-13x faster than when proxied through Graphite-web
- All users benefit from these enhancements

# Discoverability

- 1000s of dashboards from 100s of teams
- Infrastructure dashboards are frequently copied
- The more popular the dashboard, the more copies exist
- Hard to determine if a dashboard already exists
- Coarse permissioning leads to chaos
  - Users overwriting each others' dashboards
  - Almost impossible to control

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

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

Bloomberg  
Engineering

# Discoverability - Solution?

- Folders
  - Folder per team
  - Permissions per folder
- Users can still copy dashboards to their folder
  - Except now, names can be exactly the same :/
- Generic dashboard owners complaining
  - Users are finding the wrong dashboards and complaining

Q apdx\_dash|

aimstp

- APDX dashboard for moblotsv

gutsInfrastructure

- APDX\_DASH
- APDX\_DASH Copy NM
- APDX\_DASH Multi
- APDX\_DASH Safe Div

mail

- APDX\_DASH - All Requests
- APDX\_DASH with Yesterday Overlaid

tssr

- APDX\_DASH - Saru Copy for BAS connection limit exampe - Dec 2018 SRE meetup
- APDX\_DASH - TS SRE
- APDX\_DASH copy for TS specific analysis
- APDX\_DASH Multi Copy

tomspm

- APDX\_DASH Allan Testing

TechAtBloomberg.com

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

Bloomberg  
Engineering



apdx\_dash

drqsGroup986

APDX\_DASH Copy  
drqsGroup986

dlib

APDX\_DASH Copy  
dlib

APDX Top

mobcollab

APDX\_DASH Copy  
mobcollab

tiappliedtelemetry

APDX\_DASH Copy  
tiappliedtelemetry

APDX Top

wfsvc

APDX\_DASH Copy  
wfsvc

APDX Top

equityAppFramework

APDX\_DASH Copy  
equityAppFramework

APDX Top

ldrvcore

APDX\_DASH Copy  
ldrvcore

APDX Top

aimtm

APDX\_DASH Copy  
aimtm

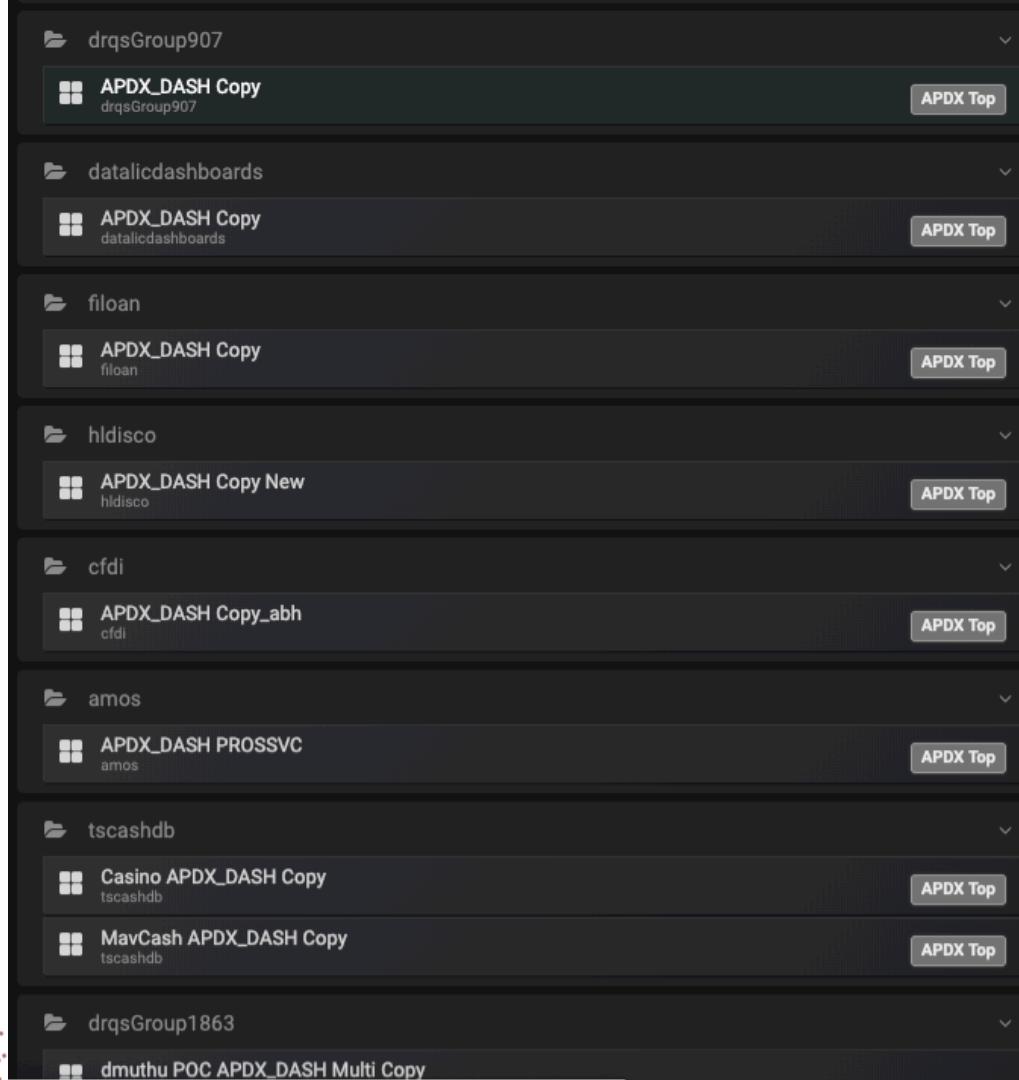
APDX Top

drqsGroup907

TechAtBloomberg.com

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

Bloomberg  
Engineering



**TechAtBloomberg.com**

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

**Bloomberg**  
Engineering

# Discoverability - Possible Solutions

- Enhanced Autocomplete
  - Search for keywords and description
- Non-copyable flags / tags
- Ability to mark a dashboard as “Official”
- Mark a dashboard as “Experimental” too

TechAtBloomberg.com

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

Bloomberg  
Engineering

# Summary

- Unifying telemetry across 5K+ engineers is hard
- APIs should guide users to best practices
- Some users love telemetry (too much!)
- Many others see telemetry as a “nice to have”
  - Make it easy to have
  - Better yet, make it free!
- Having the right tools makes things much easier
  - Grafana
  - MetricTank
- Don’t be afraid to get your hands dirty!
- Investing in telemetry pays dividends ~(ˇˇ~)

TechAtBloomberg.com

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

Bloomberg  
Engineering

# The end...

Sean Hanson - [shanson7@bloomberg.net](mailto:shanson7@bloomberg.net)

Stig Sorensen - [ssorensen3@bloomberg.net](mailto:ssorensen3@bloomberg.net)

TechAtBloomberg.com

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

Bloomberg  
Engineering