Monitoring Solr at Scale

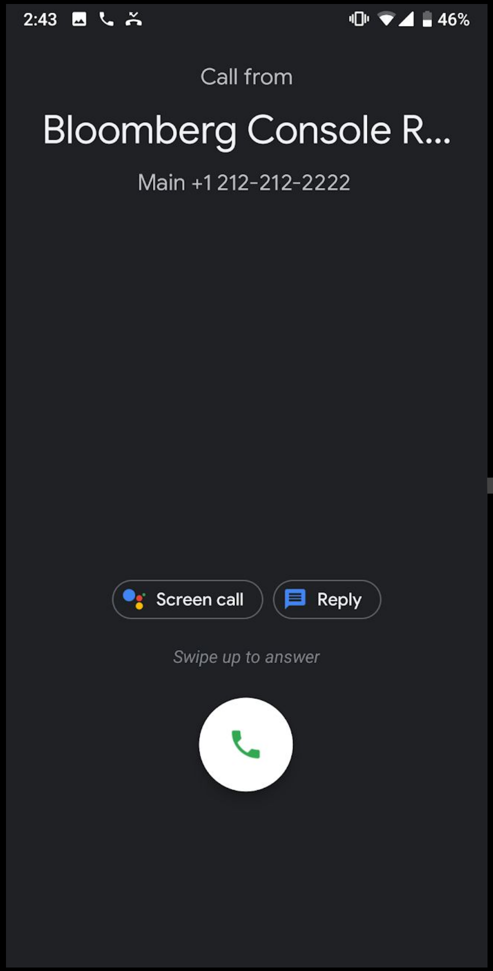
Bloomberg Engineering

Berlin Buzzwords
June 18, 2019

Ken LaPorte
Team Lead, Search Infrastructure

TechAtBloomberg.com

Has this happened to you?



2:43 46%

Call from

Bloomberg Console R...

Main +1 212-212-2222

Swipe up to answer

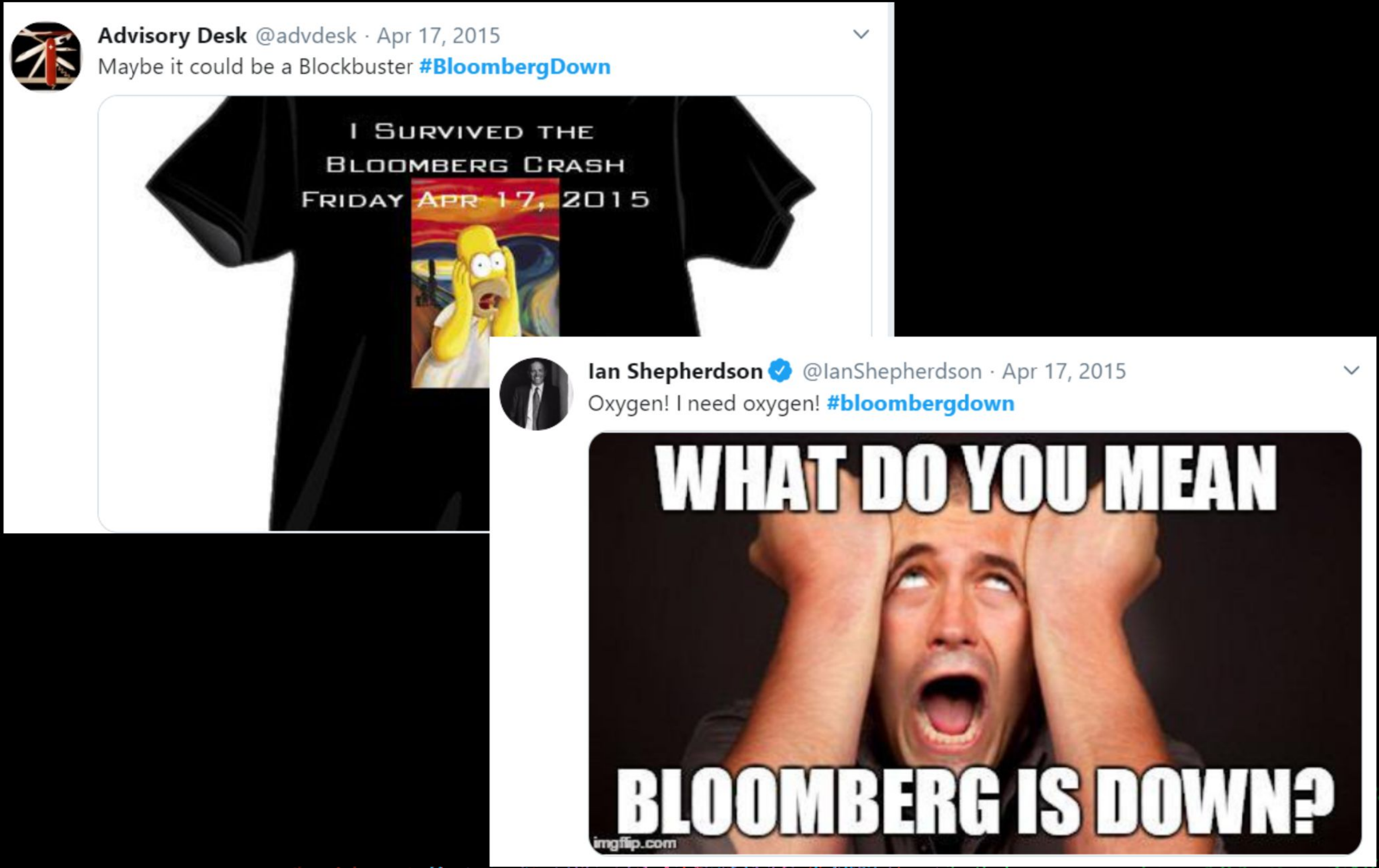
Photo by Vic, CC BY 2.0

What’s a Bloomberg?

- Largest provider of financial news and information
- Our strength is quickly and accurately delivering data, news and analytics
- Creating highly-performant and accurate information retrieval systems is core to our strength

Bloomberg

Engineering



Advisory Desk @advdesk · Apr 17, 2015

Maybe it could be a Blockbuster #BloombergDown

Ian Shepherdson @IanShepherdson · Apr 17, 2015

Oxygen! I need oxygen! #bloombergdown

Search Infrastructure at Bloomberg

- **Hundreds of search applications & ZK Ensembles**
  - Diverse use cases and scale
  - Displaced other technologies
- **10s of billions documents**
- **100s of millions new documents daily**
- **1,000s of servers**
- **10,000s of Solr instances**
- **10,000s of queries per second**
- **Critical to Bloomberg and the global financial markets**

Logos courtesy Apache Software Foundation

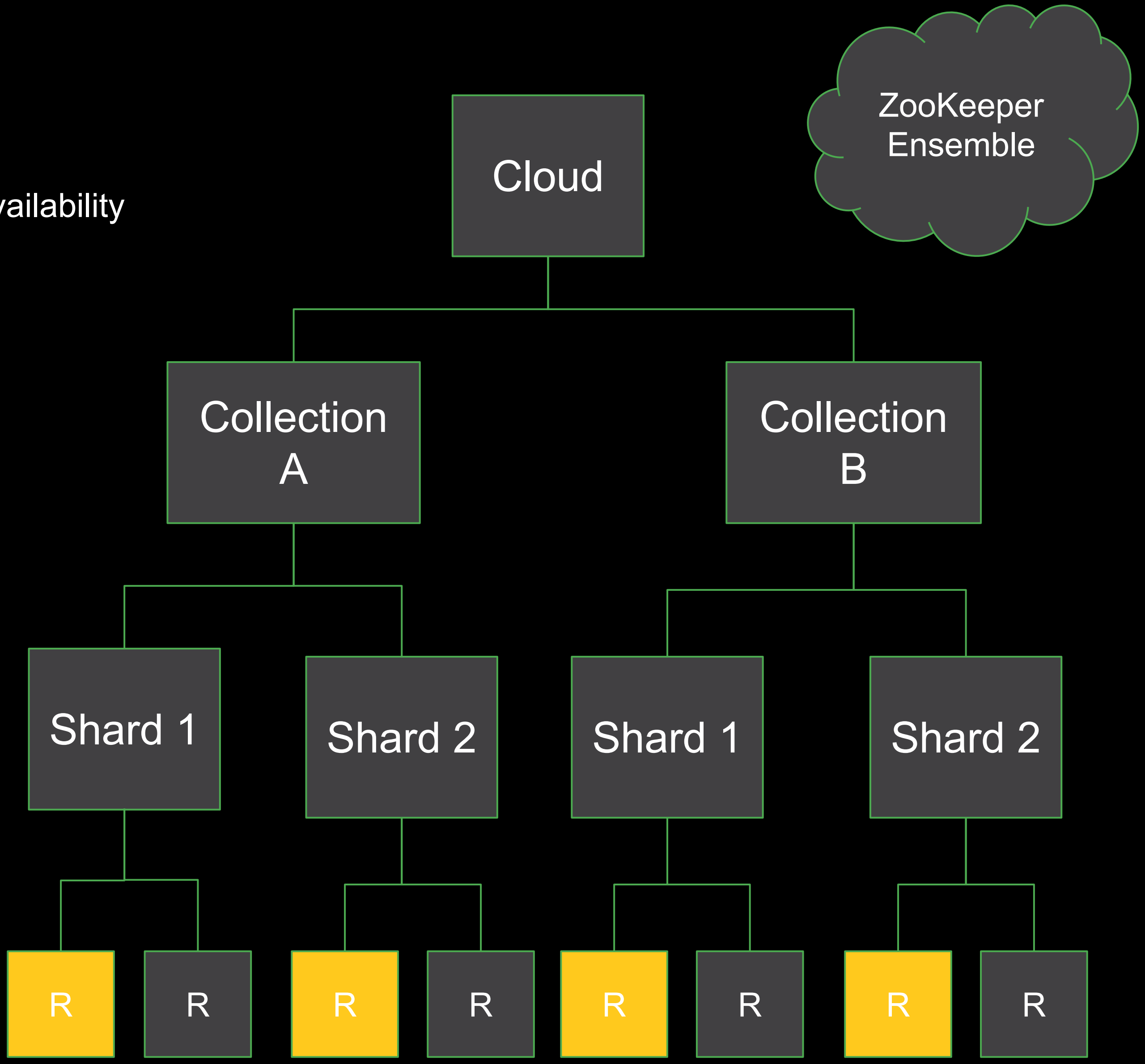
Solr Open Source Contributions at Bloomberg

- Analytics component
- Streaming expressions
- Learning to rank
- Bug fixes
- 3 committers and 2 PMC members
    - On teams throughout Bloomberg

Logos courtesy Apache Software Foundation

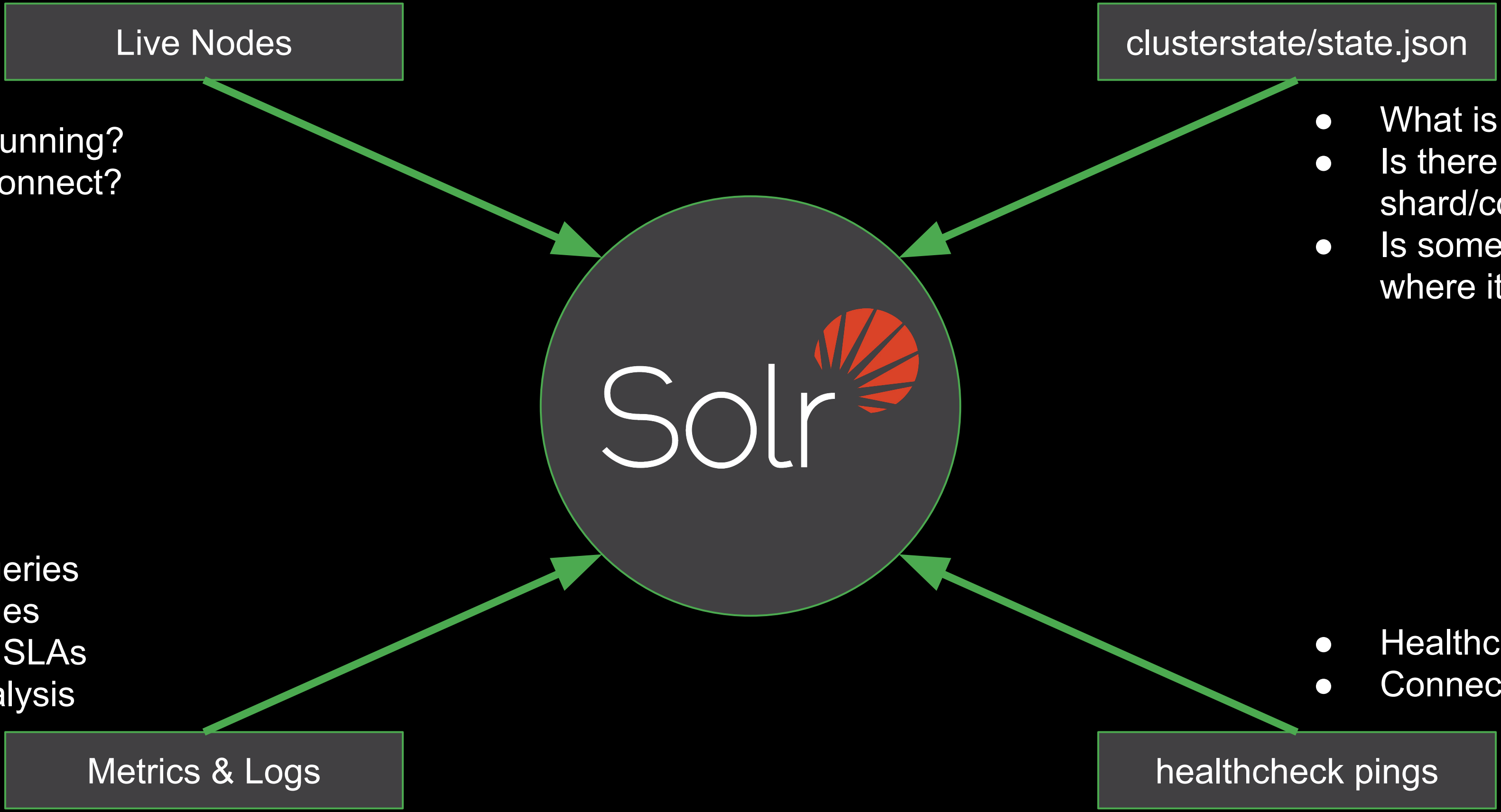
Some Terms

- Solr Cloud
  - A cluster setup that allows for fault tolerance and high availability
- Collection
  - A complete logical index
- Shard
  - A logical piece (or slice) of a collection
- Replica
  - Simply a copy of a shard
- Leader
  - The shard replica responsible for indexing
  - Represented in yellow
- ZooKeeper
  - A consensus management system used by Solr
- Service Provider
  - The service management team (Me)
- Tenant
  - The application team using a service
  - Hostile users <- Credit Anshum
- Alarming
  - Active notification of relevant events
- Monitoring
  - Next slide please



*Leader

What does it mean to monitor Solr?



Live Nodes

- Is Solr running?
- Can it connect?

clusterstate/state.json

- What is the current state?
- Is there a leader for each shard/collection?
- Is something running where it shouldn't?

Metrics & Logs

- Slow queries
- GC issues
- Tenant SLAs
- Log Analysis

Anshum Gupta's Self defence talk

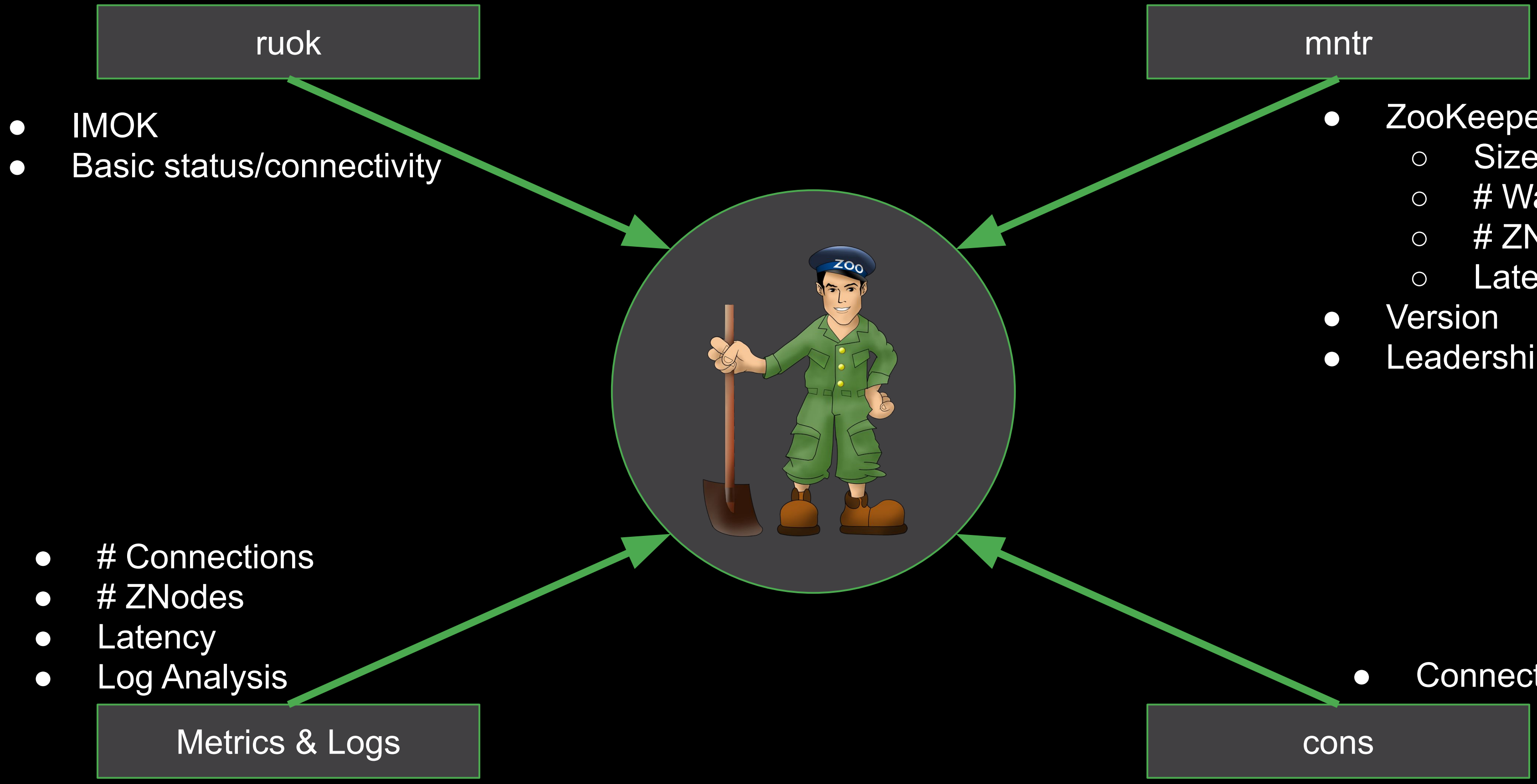
healthcheck pings

- Healthcheck pings
- Connectivity from middleware

What about ZooKeeper?

ZooKeeper 3.5 5

Dyn Reconfig.
Jetty/Rest



ruok

- IMOK
- Basic status/connectivity

mntr

- ZooKeeper stats
  - Size
  - # Watches
  - # ZNodes
  - Latency
- Version
- Leadership

- # Connections
- # ZNodes
- Latency
- Log Analysis

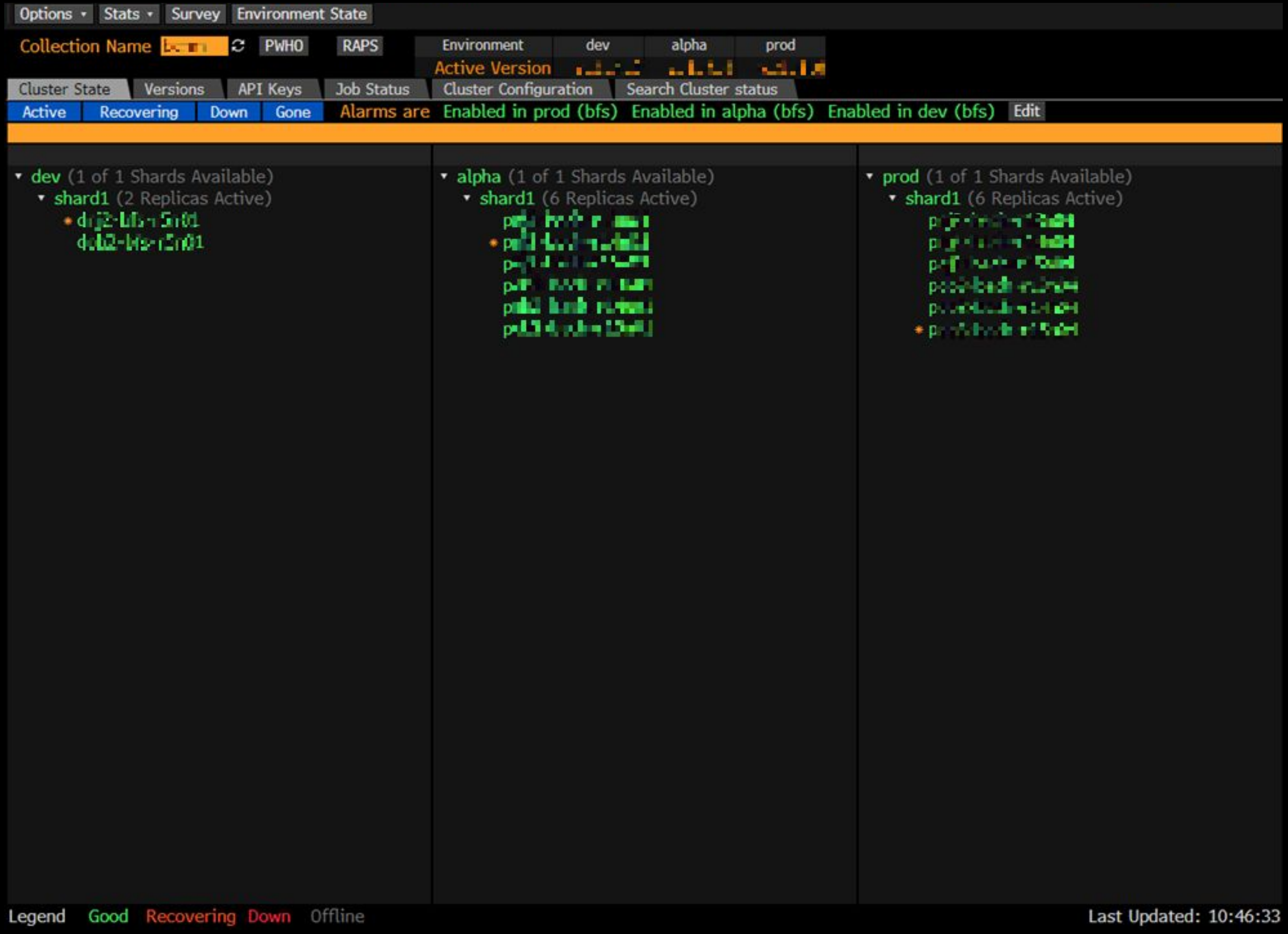
Metrics & Logs

- Connection details

cons

Can you see me now?

- Initial version
- Single Collection View
- Cross Environments
- Point in time



Collection Name [redacted]

Environment	dev	alpha	prod
Active Version	[redacted]	[redacted]	[redacted]

Alarms are Enabled in prod (bfs) Enabled in alpha (bfs) Enabled in dev (bfs)

```
▾ dev (1 of 1 Shards Available)
  ▾ shard1 (2 Replicas Active)
     • [redacted]
       [redacted]
```

```
▾ alpha (1 of 1 Shards Available)
  ▾ shard1 (6 Replicas Active)
       [redacted]
     • [redacted]
       [redacted]
       [redacted]
       [redacted]
       [redacted]
```

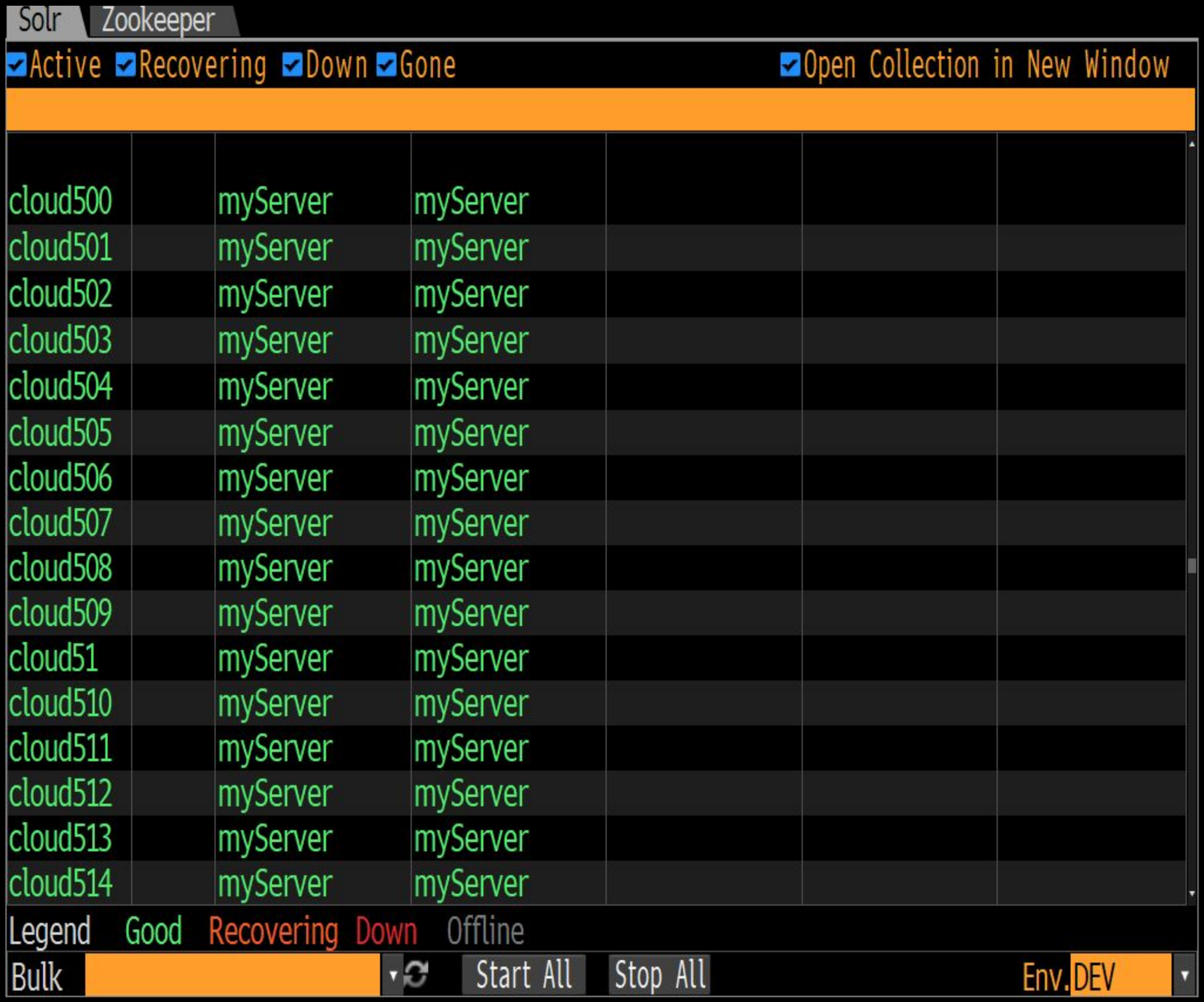
```
▾ prod (1 of 1 Shards Available)
  ▾ shard1 (6 Replicas Active)
       [redacted]
       [redacted]
       [redacted]
       [redacted]
       [redacted]
     • [redacted]
```

Legend Good Recovering Down Offline

Last Updated: 10:46:33

Service Provider View

- Display state of all clouds
- Switch between environments
- Bulk commands
  - Restart clouds by rack
- Still point in time
- Scale still a problem



☑ Active ☑ Recovering ☑ Down ☑ Gone ☑ Open Collection in New Window

cloud500		myServer	myServer			
cloud501		myServer	myServer			
cloud502		myServer	myServer			
cloud503		myServer	myServer			
cloud504		myServer	myServer			
cloud505		myServer	myServer			
cloud506		myServer	myServer			
cloud507		myServer	myServer			
cloud508		myServer	myServer			
cloud509		myServer	myServer			
cloud51		myServer	myServer			
cloud510		myServer	myServer			
cloud511		myServer	myServer			
cloud512		myServer	myServer			
cloud513		myServer	myServer			
cloud514		myServer	myServer			

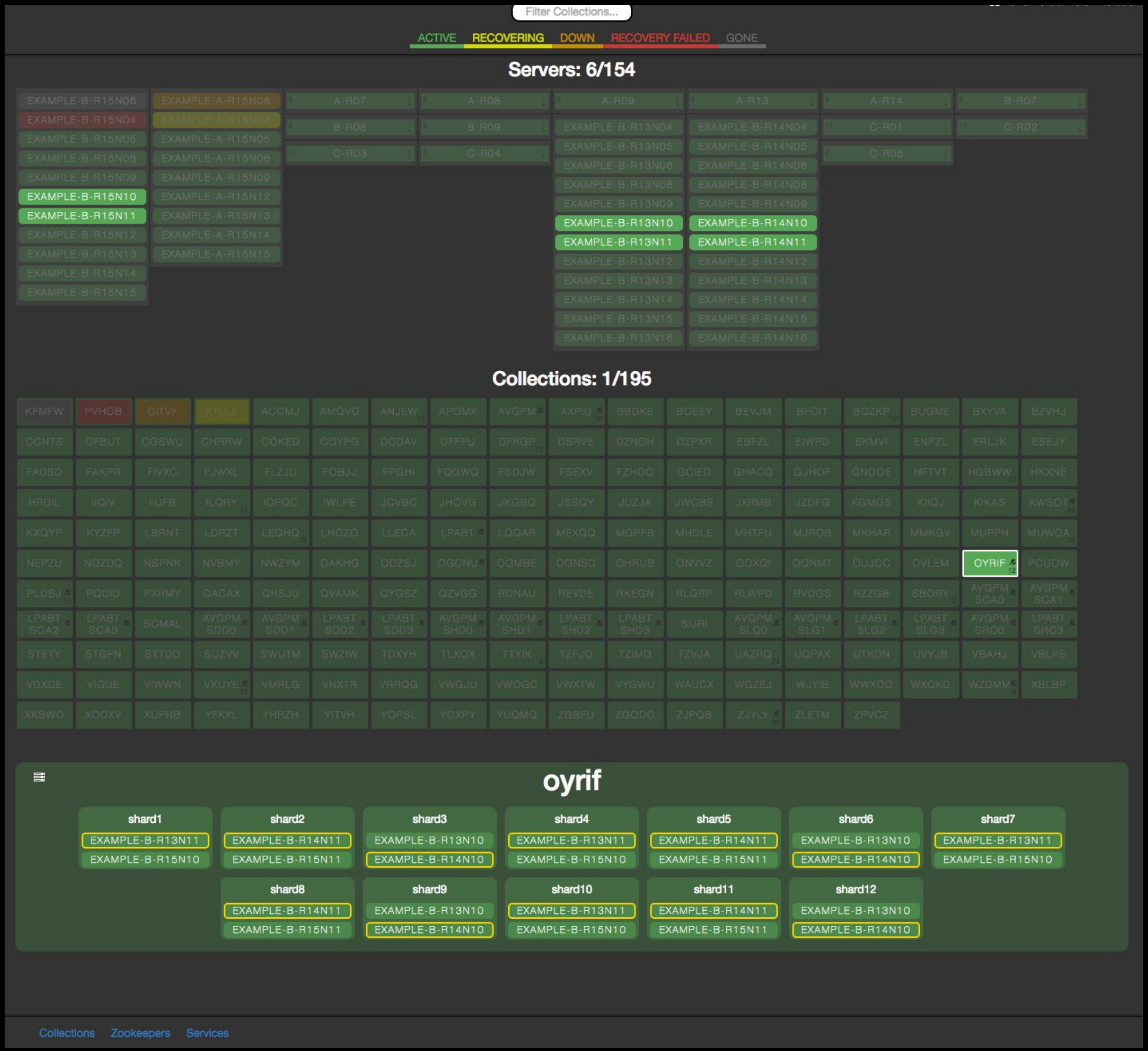
Legend Good Recovering Down Offline

Env. DEV

Names changed for display purposes

Hello Niteowl!

- Active monitor of state changes
- Backend
  - Intelligent alarms
- Frontend
  - More scalable
  - Service provider only



Filter Collections...

ACTIVE RECOVERING DOWN RECOVERY FAILED GONE

Servers: 6/154

EXAMPLE-B-R15N06
EXAMPLE-B-R15N04
EXAMPLE-B-R15N05
EXAMPLE-B-R15N08
EXAMPLE-B-R15N09
EXAMPLE-B-R15N10
EXAMPLE-B-R15N11
EXAMPLE-B-R15N12
EXAMPLE-B-R15N13
EXAMPLE-B-R15N14
EXAMPLE-B-R15N15

EXAMPLE-A-R15N06
EXAMPLE-A-R15N04
EXAMPLE-A-R15N05
EXAMPLE-A-R15N08
EXAMPLE-A-R15N09
EXAMPLE-A-R15N12
EXAMPLE-A-R15N13
EXAMPLE-A-R15N14
EXAMPLE-A-R15N15

A-R07
B-R08
C-R03

A-R08
B-R09
C-R04

A-R09
EXAMPLE-B-R13N04
EXAMPLE-B-R13N05
EXAMPLE-B-R13N06
EXAMPLE-B-R13N08
EXAMPLE-B-R13N09
EXAMPLE-B-R13N10
EXAMPLE-B-R13N11
EXAMPLE-B-R13N12
EXAMPLE-B-R13N13
EXAMPLE-B-R13N14
EXAMPLE-B-R13N15
EXAMPLE-B-R13N16

A-R13
EXAMPLE-B-R14N04
EXAMPLE-B-R14N05
EXAMPLE-B-R14N06
EXAMPLE-B-R14N08
EXAMPLE-B-R14N09
EXAMPLE-B-R14N10
EXAMPLE-B-R14N11
EXAMPLE-B-R14N12
EXAMPLE-B-R14N13
EXAMPLE-B-R14N14
EXAMPLE-B-R14N15
EXAMPLE-B-R14N16

A-R14
C-R01
C-R05

B-R07
C-R02

Collections: 1/195

KFMFW PVHOB OITVF KTLLE ACCMJ AMQVO ANJEW APOMX AVGPM AXPIU BBDKE BCEEY BEVJM BFOIT BQZKP BUGME BXYVA BZVHJ
CCNTS CFBUT CGSWU CHRRW COKED CQYPG DCDAV DFFPU DFRGP DSRVE DZNOH DZPXR EBFZL EIWPD EKMVI ENPZL ERLJK ESEJY
FADSD FAKPR FIVXC FJWXL FLZJU FOBJJ FPGHI FQGWQ FSDJW FSEXV FZHOC GCIED GHACG GJHOF GNOOE HFTVT HGBWW HKXNE
HRDIL IIQIV IIUFB ILQRY IOPQC IWLPE JCVBC JHCVG JKGBQ JSSQY JUZJA JWCBS JXRMB JZDFG KGMGS KIIQJ KIKAS KWSOT
KXQYP KYZFP LBRNT LDRZT LEQHQ LHOZO LLECA LPABT LQQAR MEXQQ MGPFB MHDLE MHTFU MJROB MKHAR MMKGV MUPPH MUWOA
NEPZU NOZDQ NSPNK NVBMY NWZYM OAKHG ODZSJ OGCNU OGMBE OGNSD OHRUB ONVVZ OOXQI OQNMT OUJCC OVLEM OYRIF PCUOW
PLDSJ PQDID PXRMY QACAX QHSJU QVAMK QYGSZ QZVGG RDNAU REVDE RKEGN RLQRP RLWPD RVDGS RZZGB SBORY AVGPM SCA0 AVGPM SCA1
LPABT SCA2 LPABT SCA3 SCMAL AVGPM SDD0 AVGPM SDD1 LPABT SDD2 LPABT SDD3 AVGPM SHD0 AVGPM SHD1 LPABT SHD2 LPABT SHD3 SIJRI AVGPM SLG0 AVGPM SLG1 LPABT SLG2 LPABT SLG3 AVGPM SRC0 LPABT SRC3
STETY STGFN STTDD SUZVV SWUTM SWZIW TDXYH TLXOX TTYIK TZFJO TZIMO TZVJA UAZRQ UQPAX UTKDN UVYJB VBAHJ VBLPS
VDXDE VIGUE VIWWN VKUYE VMRLG VNXTR VRRQG VWGJU VWOGC VWXTW VYGWU WAUCX WGZEJ WJYIB WWXOO WXQKD WZDMM XBLBP
XKSWO XODXV XUPNB YFKYL YHRZH YITVH YOPSL YOXPY YUQMQ ZGBFU ZGQDC ZJPQB ZJYLY ZLETM ZPVCZ

oyrif

shard1
EXAMPLE-B-R13N11
EXAMPLE-B-R15N10

shard2
EXAMPLE-B-R14N11
EXAMPLE-B-R15N11

shard3
EXAMPLE-B-R13N10
EXAMPLE-B-R14N10

shard4
EXAMPLE-B-R13N11
EXAMPLE-B-R15N10

shard5
EXAMPLE-B-R14N11
EXAMPLE-B-R15N11

shard6
EXAMPLE-B-R13N10
EXAMPLE-B-R14N10

shard7
EXAMPLE-B-R13N11
EXAMPLE-B-R15N10

shard8
EXAMPLE-B-R14N11
EXAMPLE-B-R15N11

shard9
EXAMPLE-B-R13N10
EXAMPLE-B-R14N10

shard10
EXAMPLE-B-R13N11
EXAMPLE-B-R15N10

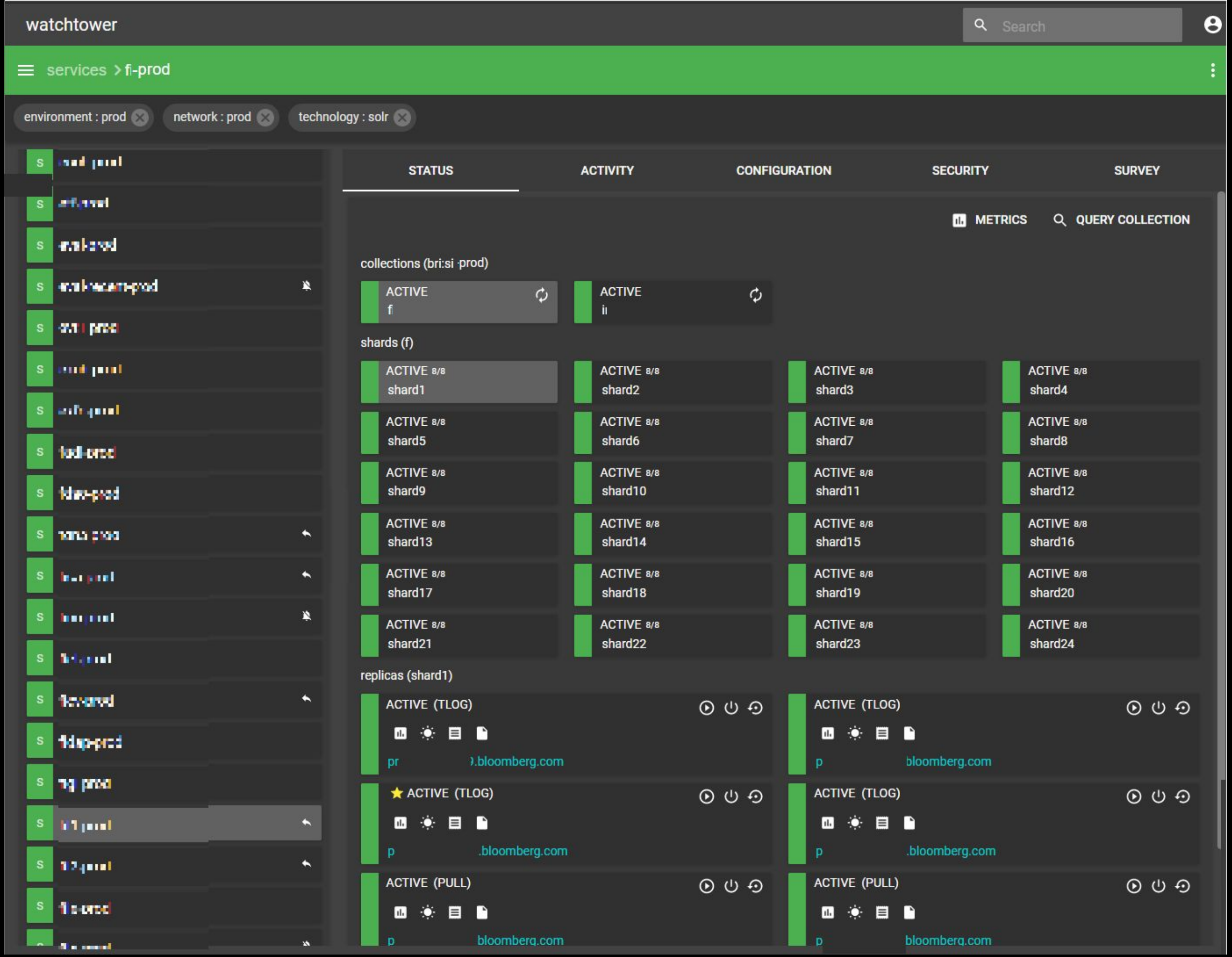
shard11
EXAMPLE-B-R14N11
EXAMPLE-B-R15N11

shard12
EXAMPLE-B-R13N10
EXAMPLE-B-R14N10

Data for display purposes only.

All along the WatchTower….

- Allows Svc provider & tenants to share a UI
- Leverages same Niteowl back-end
- Day 2 actions
  - Process management (start, stop, etc.)
  - Schema changes
  - Query interface
  - Metrics
  - Security



watchtower

Search

services > fi-prod

environment : prod ⊗ network : prod ⊗ technology : solr ⊗

S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]
S [redacted]

STATUS | ACTIVITY | CONFIGURATION | SECURITY | SURVEY

METRICS | QUERY COLLECTION

collections (bri:si ·prod)

ACTIVE f | ACTIVE ii

shards (f)

ACTIVE 8/8 shard1 | ACTIVE 8/8 shard2 | ACTIVE 8/8 shard3 | ACTIVE 8/8 shard4
ACTIVE 8/8 shard5 | ACTIVE 8/8 shard6 | ACTIVE 8/8 shard7 | ACTIVE 8/8 shard8
ACTIVE 8/8 shard9 | ACTIVE 8/8 shard10 | ACTIVE 8/8 shard11 | ACTIVE 8/8 shard12
ACTIVE 8/8 shard13 | ACTIVE 8/8 shard14 | ACTIVE 8/8 shard15 | ACTIVE 8/8 shard16
ACTIVE 8/8 shard17 | ACTIVE 8/8 shard18 | ACTIVE 8/8 shard19 | ACTIVE 8/8 shard20
ACTIVE 8/8 shard21 | ACTIVE 8/8 shard22 | ACTIVE 8/8 shard23 | ACTIVE 8/8 shard24

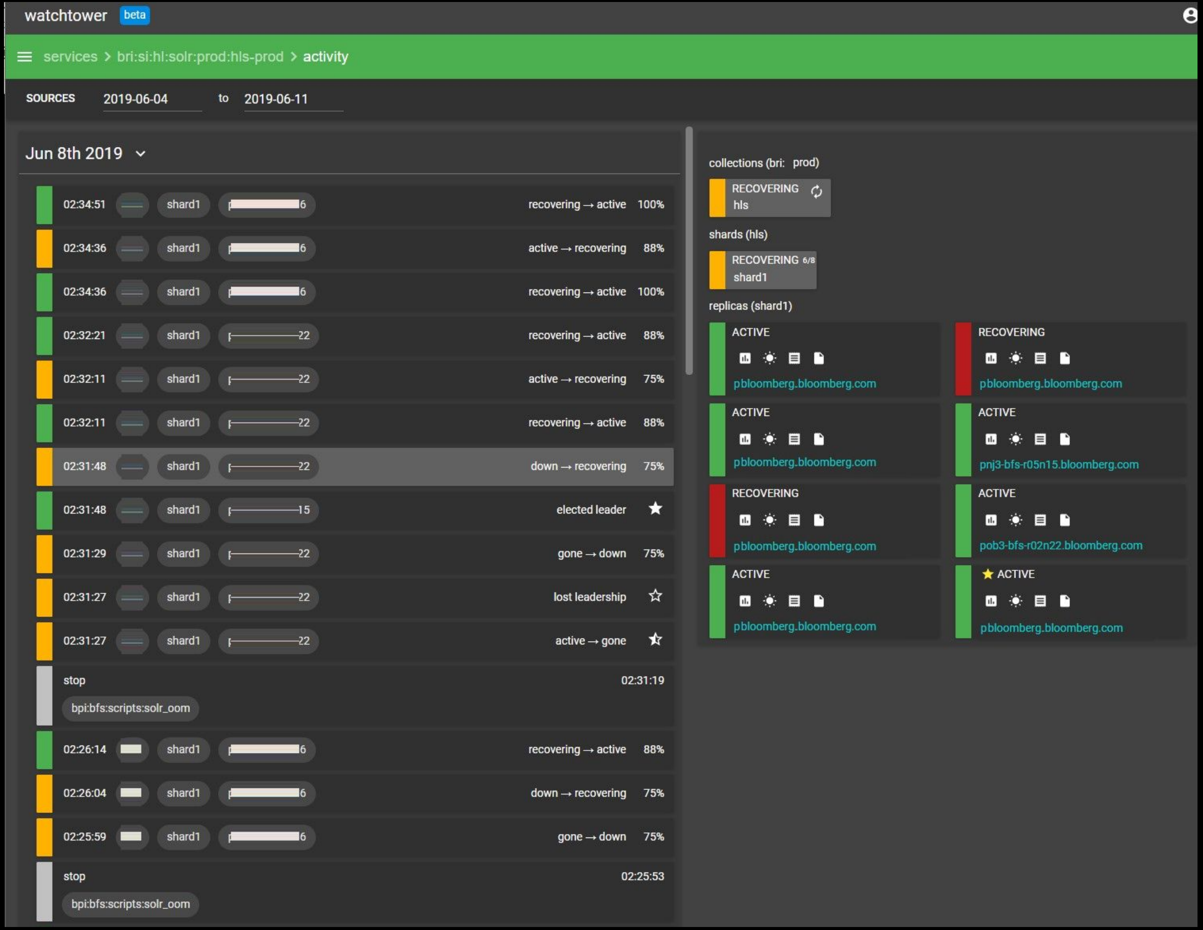
replicas (shard1)

ACTIVE (TLOG) pr).bloomberg.com | ACTIVE (TLOG) p bloomberg.com
★ ACTIVE (TLOG) p .bloomberg.com | ACTIVE (TLOG) p .bloomberg.com
ACTIVE (PULL) p bloomberg.com | ACTIVE (PULL) p bloomberg.com

Credit Jimi Hendrix & the Noun Project
Names blurred/changed for IP

Alarming, Activity & Analysis

- Alarming
  - Cloud aware
  - Consolidation across signals
  - Delay Alarms to prevent flutter
- Activity
  - Auditability is critical
  - Every state change is recorded
  - Each user action is logged
  - Additional sources (OOMs)
- Analysis
  - Events indexed into Solr
  - Analytics helped address multitude of issues



2019-06-04 to 2019-06-11

Jun 8th 2019

```
02:34:51  shard1  6   recovering → active    100%
02:34:36  shard1  6   active → recovering     88%
02:34:36  shard1  6   recovering → active    100%
02:32:21  shard1  22  recovering → active     88%
02:32:11  shard1  22  active → recovering     75%
02:32:11  shard1  22  recovering → active     88%
02:31:48  shard1  22  down → recovering       75%
02:31:48  shard1  15  elected leader           ★
02:31:29  shard1  22  gone → down             75%
02:31:27  shard1  22  lost leadership          ☆
02:31:27  shard1  22  active → gone            ☆
stop                                    02:31:19
bpi:bfs:scripts:solr_oom
02:26:14  shard1  6   recovering → active     88%
02:26:04  shard1  6   down → recovering       75%
02:25:59  shard1  6   gone → down             75%
stop                                    02:25:53
bpi:bfs:scripts:solr_oom
```

collections (bri: prod)

RECOVERING hls

shards (hls)

RECOVERING 6/8 shard1

replicas (shard1)

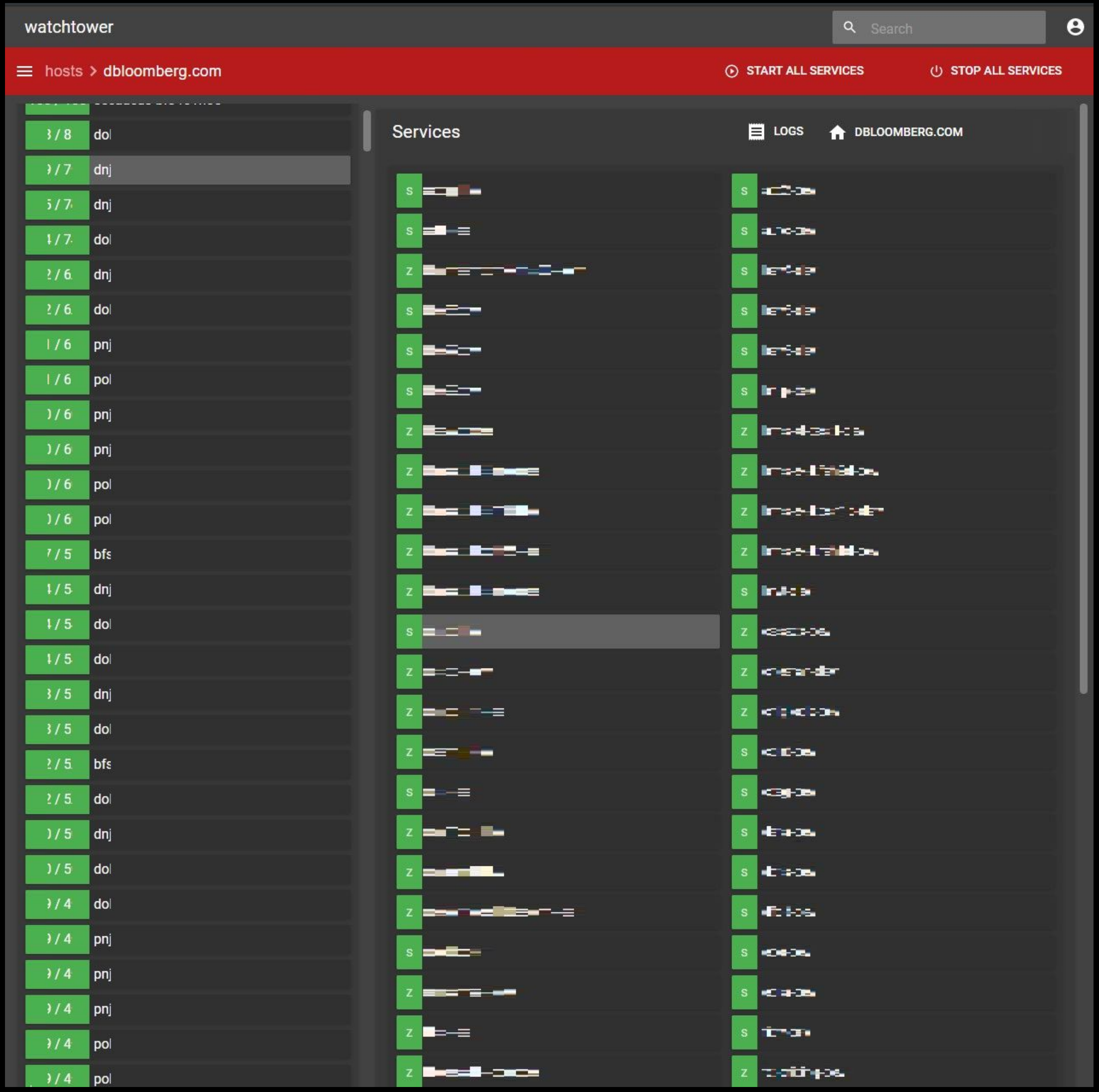
ACTIVE pbloomberg.bloomberg.com	RECOVERING pbloomberg.bloomberg.com
ACTIVE pbloomberg.bloomberg.com	ACTIVE pnj3-bfs-r05n15.bloomberg.com
RECOVERING pbloomberg.bloomberg.com	ACTIVE pob3-bfs-r02n22.bloomberg.com
ACTIVE pbloomberg.bloomberg.com	★ ACTIVE pbloomberg.bloomberg.com

Names blurred/changed for

Slice & Dice

- Slice by:
  - Host list
  - State
  - and more

- Dice by:
  - Technology
  - Tenancy
  - Network Segment
  - and more

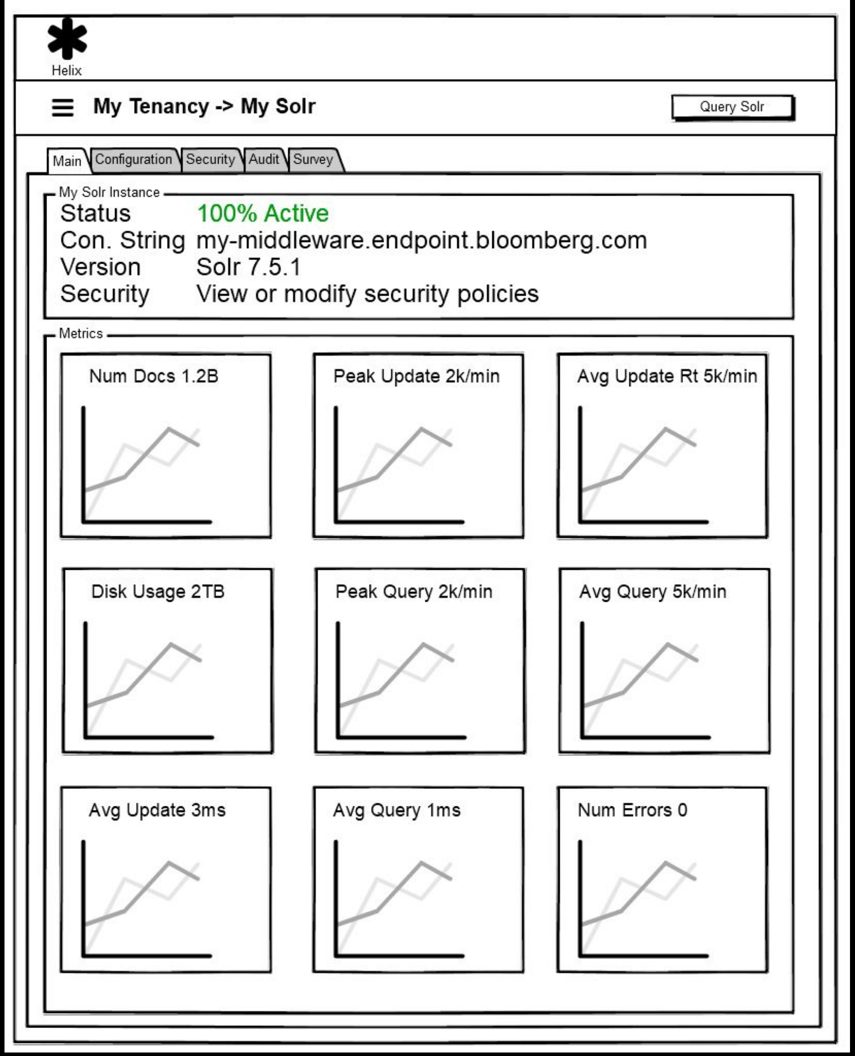


3 / 8 dol
) / 7 dnj
5 / 7 dnj
1 / 7 dol
2 / 6 dnj
2 / 6 dol
l / 6 pnj
l / 6 pol
) / 6 pnj
) / 6 pnj
) / 6 pol
) / 6 pol
7 / 5 bfs
1 / 5 dnj
1 / 5 dol
1 / 5 dol
3 / 5 dnj
3 / 5 dol
2 / 5 bfs
2 / 5 dol
) / 5 dnj
) / 5 dol
) / 4 dol
) / 4 pnj
) / 4 pnj
) / 4 pnj
) / 4 pol
) / 4 pol

Names blurred/changed for IP

nth time's a charm

- Enter Helix
- New UI
  - Shared among service providers
  - Think AWS console
- New Purpose
  - Client centric
  - Administration through automation
- Obviously still in wireframe/planning stages
- Our final UI 😁



My Solr Instance

Status 100% Active
Con. String my-middleware.endpoint.bloomberg.com
Version Solr 7.5.1
Security View or modify security policies

Metrics

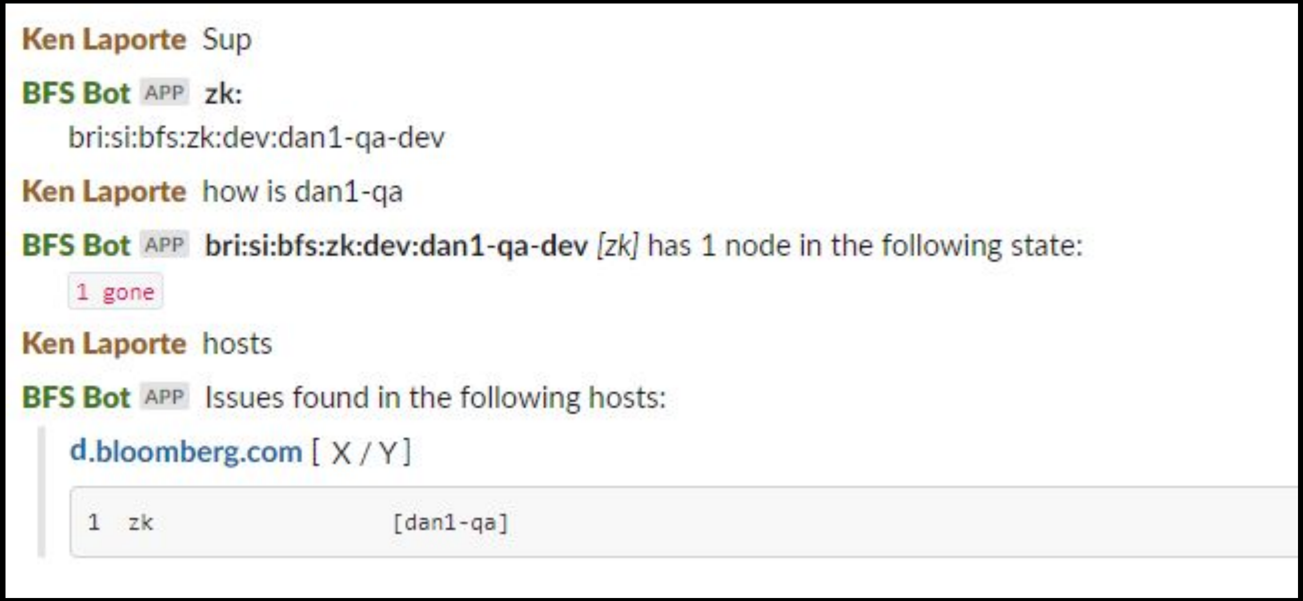
Num Docs 1.2B
Peak Update 2k/min
Avg Update Rt 5k/min
Disk Usage 2TB
Peak Query 2k/min
Avg Query 5k/min
Avg Update 3ms
Avg Query 1ms
Num Errors 0

+ = A Good Night's Sleep

- Chat Ops Goals:
  - Laptops off
  - Only wake up when necessary
  - Read-only

- Commands
  - What’s wrong (sup)?
  - Detailed views (how is *foo*)
  - Host by host breakdown (hosts)
  - More to come

- Publication of events



Ken Laporte Sup

BFS Bot APP **zk:**
bri:si:bfs:zk:dev:dan1-qa-dev

Ken Laporte how is dan1-qa

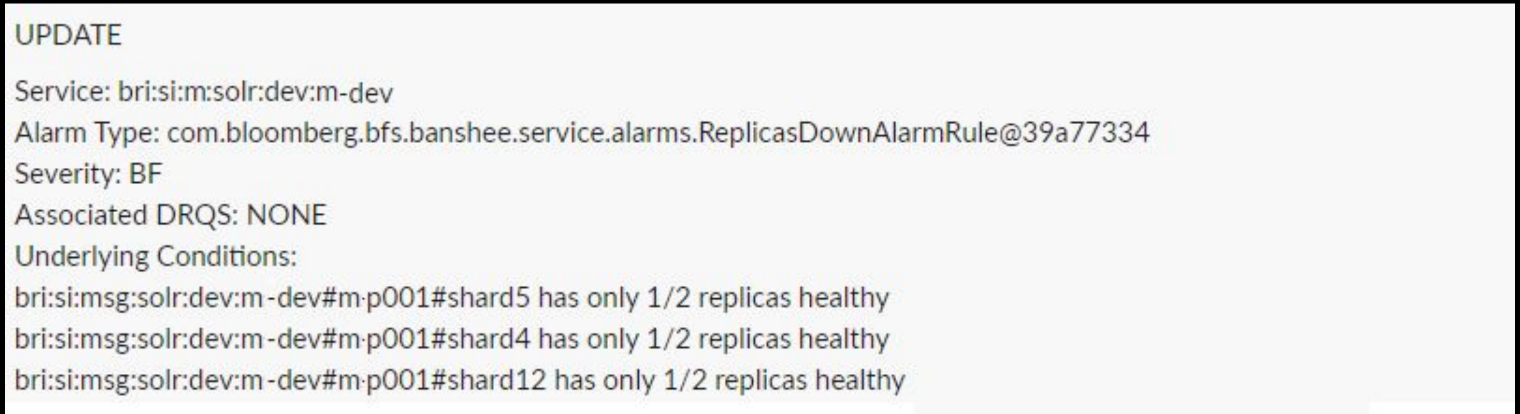
BFS Bot APP **bri:si:bfs:zk:dev:dan1-qa-dev** *[zk]* has 1 node in the following state:
`1 gone`

Ken Laporte hosts

BFS Bot APP Issues found in the following hosts:

d.bloomberg.com [X / Y]

```
1  zk                    [dan1-qa]
```

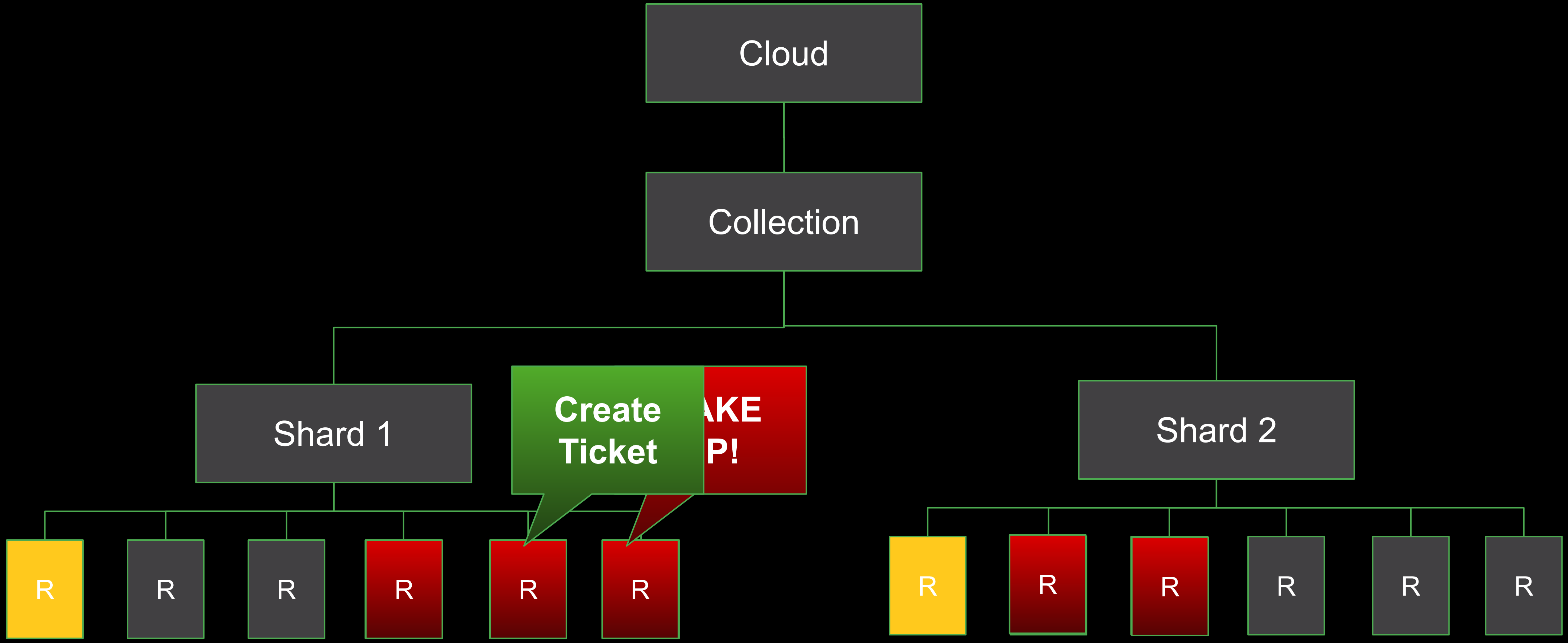


UPDATE

Service: bri:si:m:solr:dev:m-dev
Alarm Type: com.bloomberg.bfs.banshee.service.alarms.ReplicasDownAlarmRule@39a77334
Severity: BF
Associated DRQS: NONE
Underlying Conditions:
bri:si:msg:solr:dev:m-dev#m-p001#shard5 has only 1/2 replicas healthy
bri:si:msg:solr:dev:m-dev#m-p001#shard4 has only 1/2 replicas healthy
bri:si:msg:solr:dev:m-dev#m-p001#shard12 has only 1/2 replicas healthy

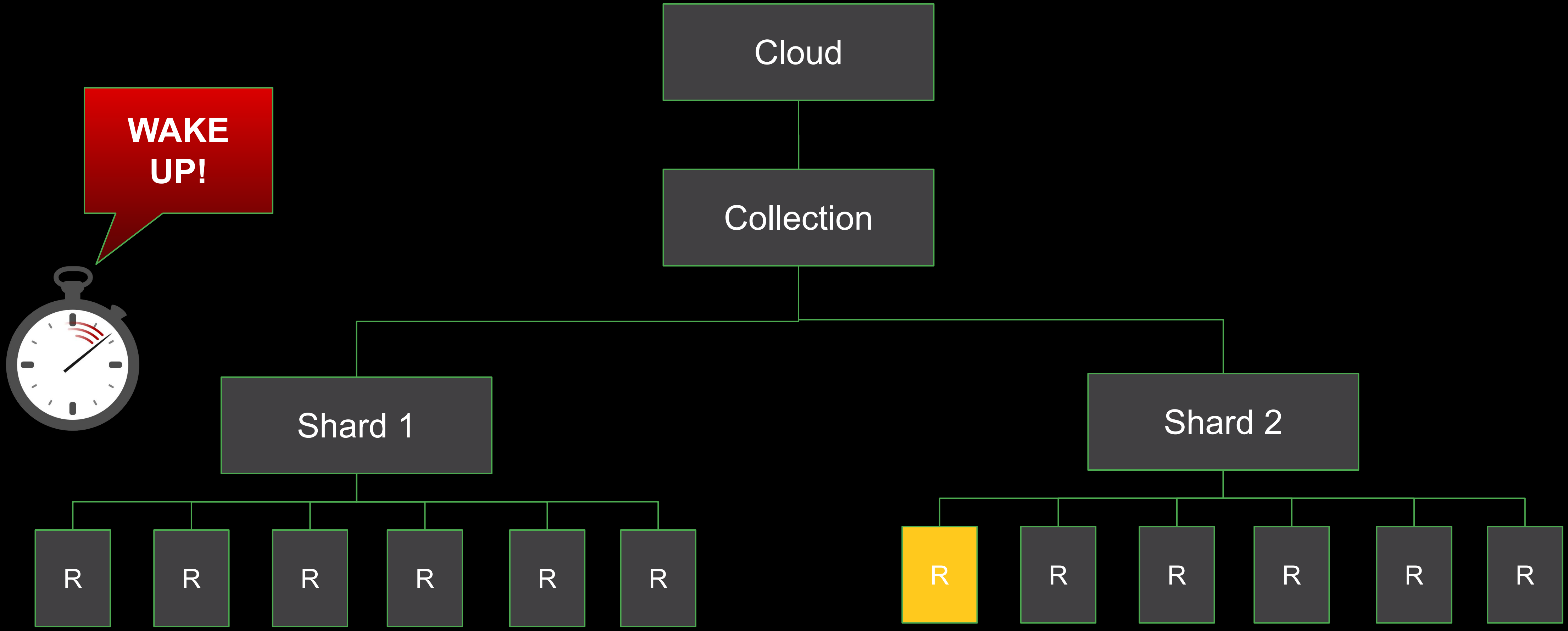
Identifiers removed for IP

Niteowl - Here today, gone tomorrow



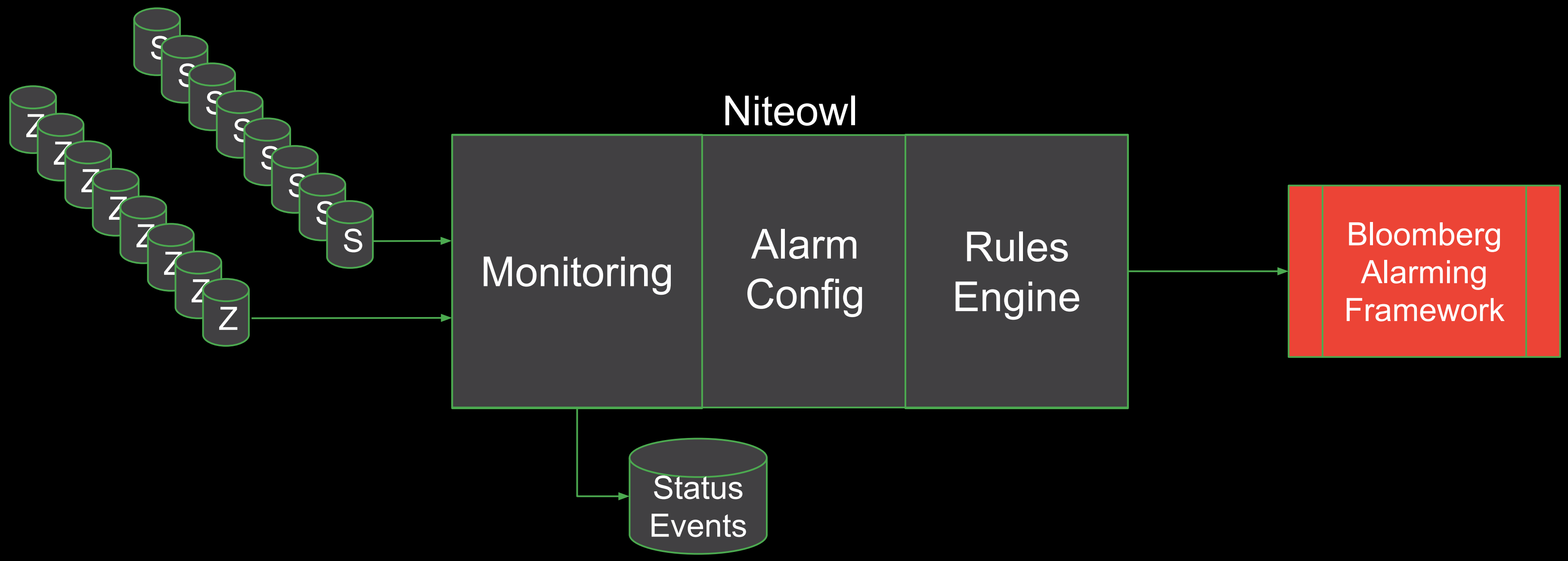
*Leader

Niteowl - Take me to your leader!

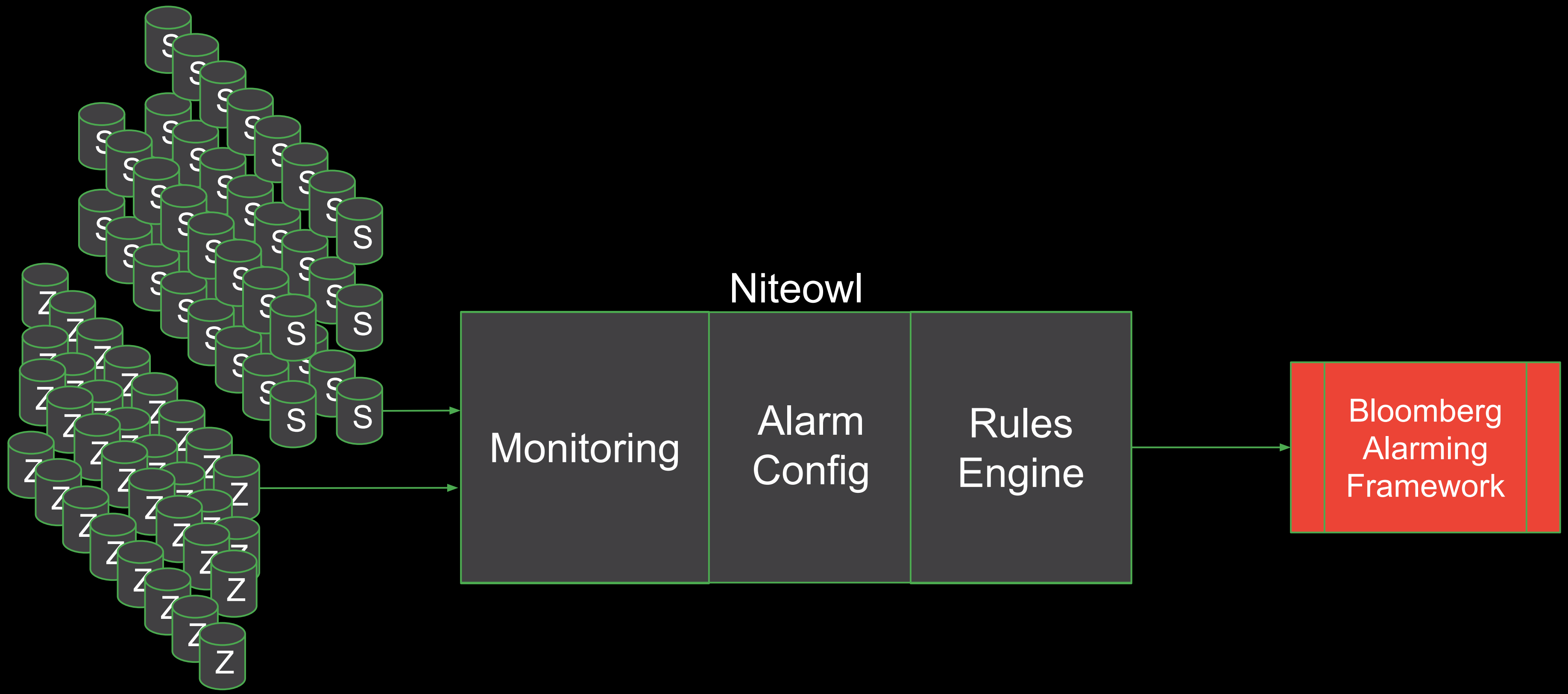


*Leader

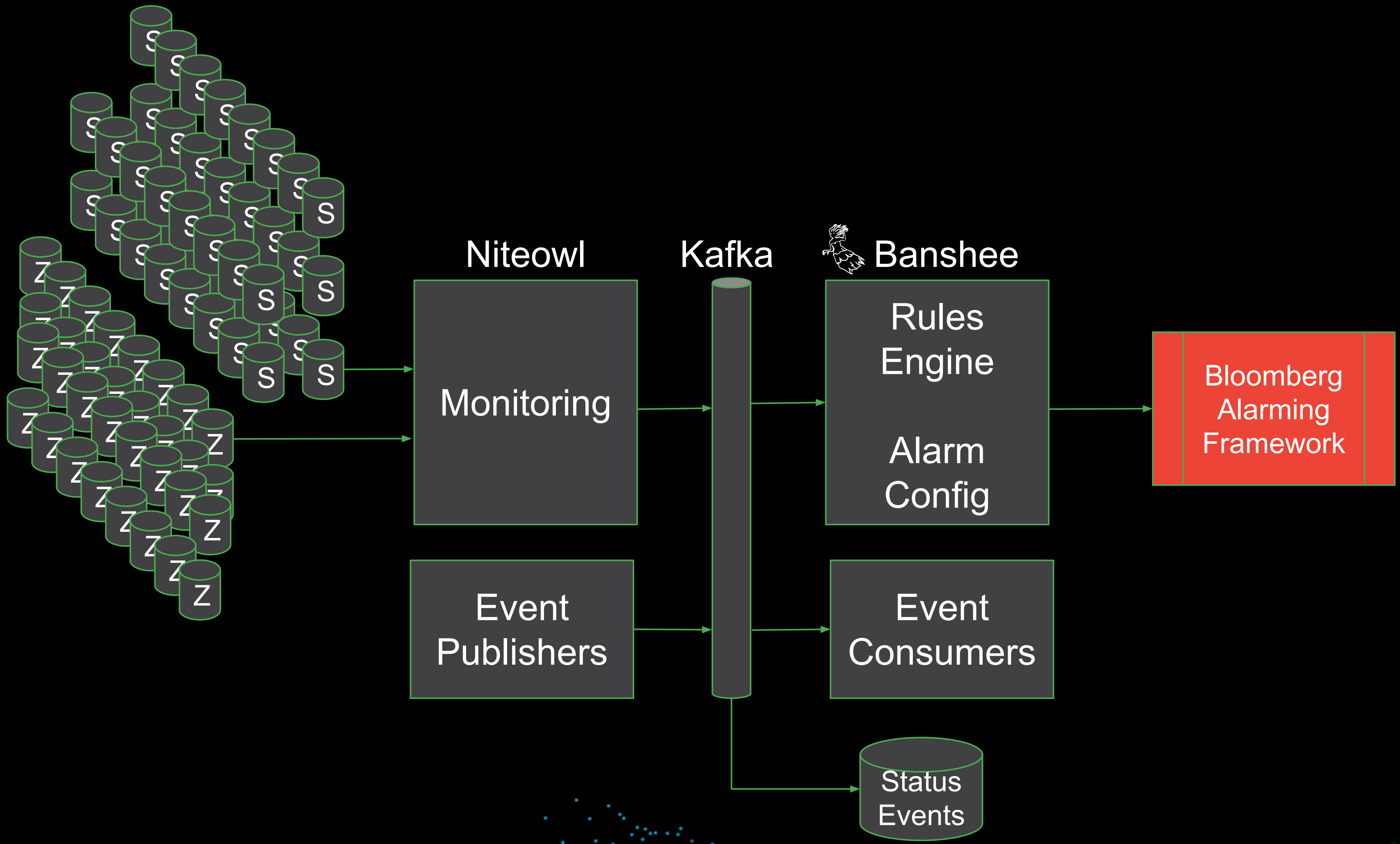
So what does all this look like?



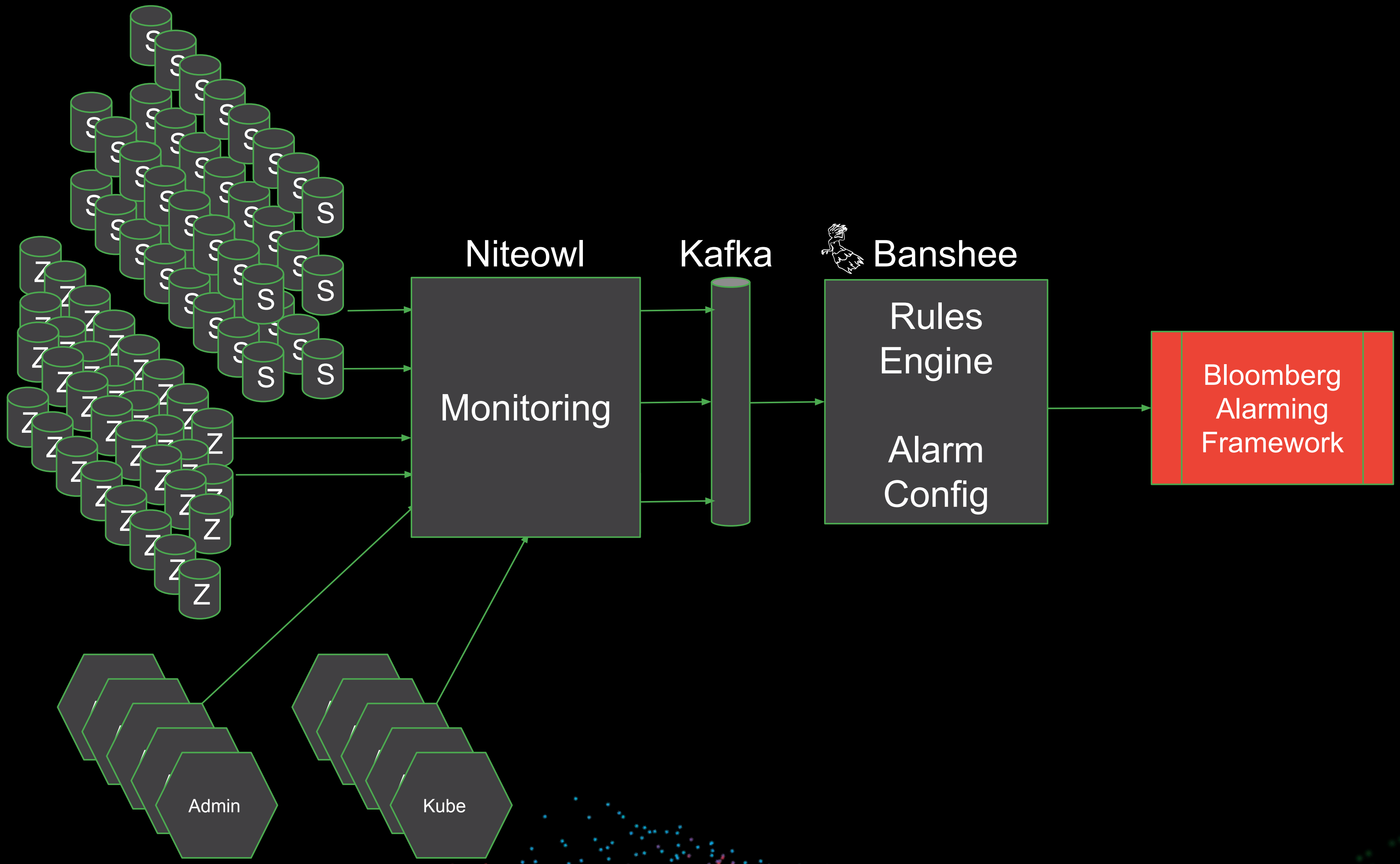
So what does all this look like?



Enter Kafka & Banshee



Run Niteowl via Kubernetes and spread the load



What’s Next?

- 2019 Goal: Rules Engine
  - Leverage Kafka Streams to create robust alarms engine
- 2019 Goal: Open Sourcing Niteowl
  - Reality: Tightly integrated into the Bloomberg Infrastructure
- 2020 Goal: Detangle Niteowl and hopefully publish as open source

Thank You, Sleep Well!

https://www.bloomberg.com/careers

Questions?

Ken LaPorte
klaporte@bloomberg.net

TechAtBloomberg.com