

SERVICE DISCOVERY IN THE CLOUD

CONSUL @ SCALE14x

DARRON@FROESE.ORG - @DARRON

DARRON FROESE



TWO MAIN COMPONENTS

WHAT IS SERVICE
DISCOVERY?

I PROVIDE THIS SERVICE AT THIS IP ADDRESS AND PORT.

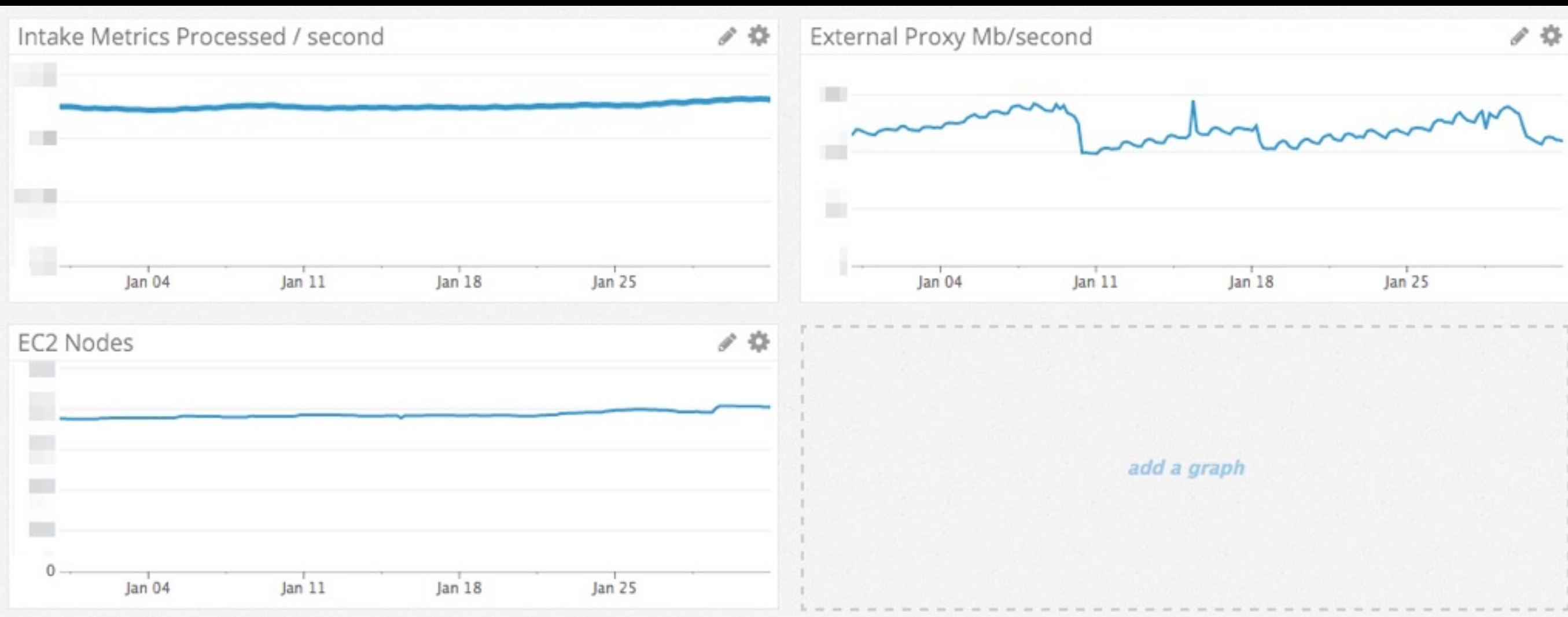
SERVICE REGISTRATION

CAN YOU TELL ME WHERE TO CONNECT TO THIS SERVICE?

SERVICE DISCOVERY

WINTER 2014 @ DATADOG

WHERE WERE WE?



LATE 2014

- 4 year old codebase.
- Cutting apart our monolith.
- Rapid growth across the board.

WE COULDN'T DO IT THE SAME WAY ANYMORE.

OLD PRACTICES WERE FRAYING

Usually, you start off with some static configuration which gets you pretty far. Things get more complicated as you start deploying more services. With a live system, service locations can change quite frequently due to auto or manual scaling, new deployments of services, as well as hosts failing or being replaced.

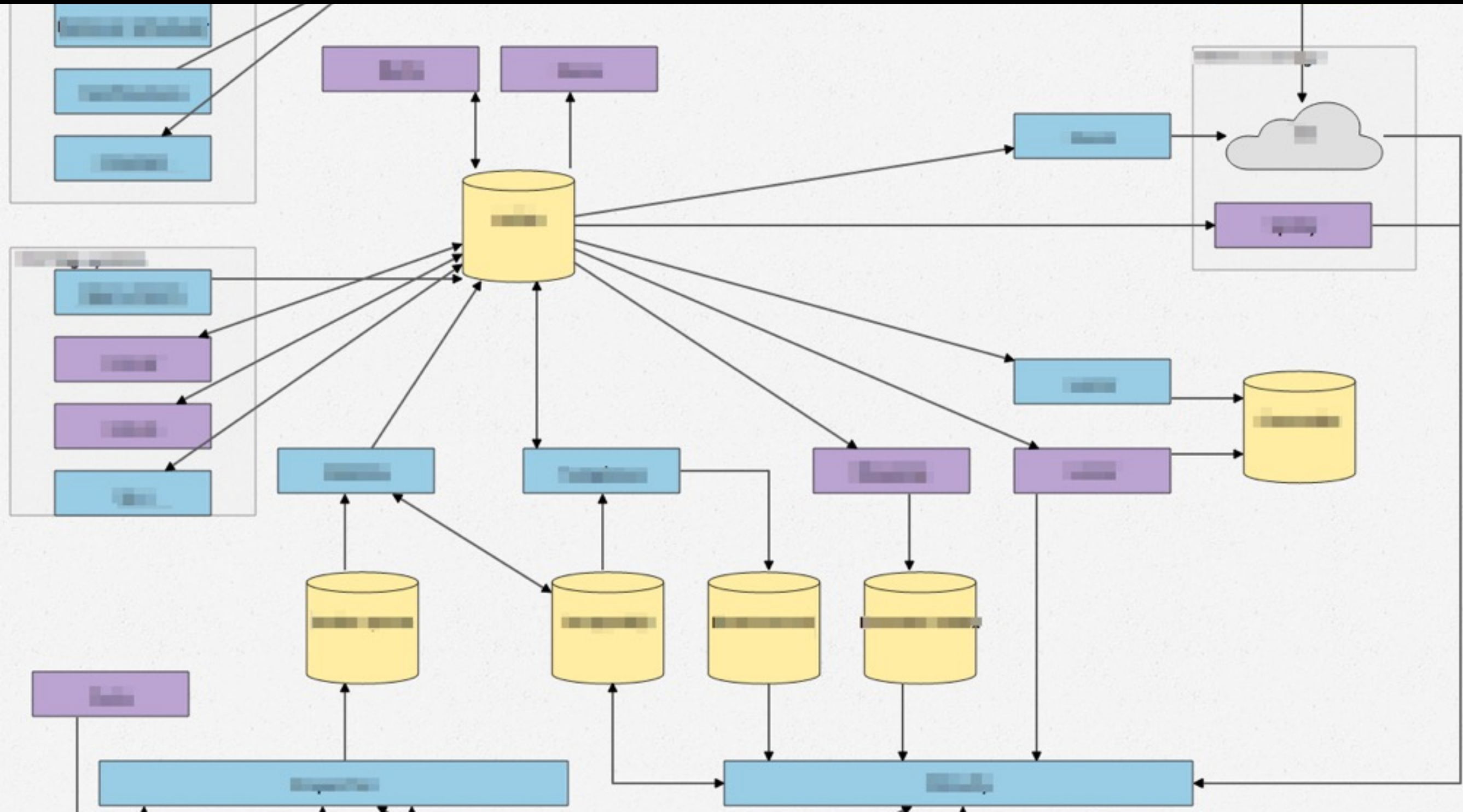
Dynamic service registration and discovery becomes much more important in these scenarios in order to avoid service interruption.

DISCOVERY WAS A HYBRID

- Chef searches. 30 minutes to update.
- Large numbers of manually managed IP addresses.
- There was nothing really wrong with it - but it was getting harder to manage.

"MO SYSTEMS. MO PROBLEMS." - THE NOTORIOUS B.I.G.

DISTRIBUTED SYSTEMS



BACKING STORE FOR DOCKER CONTAINERS

MID 2014

octohost / octohost-cookbook

Unwatch 9

Star 89

Fork 10

Code

Issues 0

Pull requests 1

Wiki

Pulse

Graphs

Settings

Removing etcd and adding consul.

Browse files

master v1.9.1 ... 1.1.2



darron committed on Jun 15, 2014

1 parent d1a819d

commit 6a00c0bd13fac70ea98e6c785041062dd9389845

Showing 6 changed files with 7 additions and 36 deletions.

Unified

Split

2 Berkfile

View

	@@ -10,5 +10,5 @@	cookbook 'redis', git: 'https://github.com/darron/redis-cookbook.git'
10	10	cookbook 'openresty', git: 'https://github.com/darron/openresty-cookbook.git'
11	11	cookbook 'gitreceive', git: 'https://github.com/darron/gitreceive-cookbook.git'
12	12	cookbook 'sysdig', git: 'https://github.com/darron/sysdig-cookbook.git'
13		-cookbook 'etcd', git: 'https://github.com/spheromak/etcd-cookbook.git'
	13	+cookbook 'consul', git: 'https://github.com/darron/consul-cookbook.git'
14	14	cookbook 'chef-sugar', git: 'https://github.com/sethvargo/chef-sugar'

4 metadata.rb

View

	@@ -4,7 +4,7 @@	

NOVEMBER 2014

OVERALL PLAN

To provide a stable system to:

1. Register and provide a catalog of services that are available in the Datadog server cluster.
2. Make a distributed Key/Value store available on any machine in the Datadog server cluster.

This will allow us to:

1. Respond to individual component outages in seconds rather than minutes by integrating the service catalog health checks and configuration generation.
2. Build configuration files across the entire cluster in seconds by interacting with the Key Value store and Service Catalog in a simple and automated fashion.
3. Help to make the Datadog server cluster more resilient by leveraging the Service Catalog and Key Value store to automate additional failover and configuration tasks.

WHAT IS CONSUL?



Service Discovery

Consul makes it simple for services to register themselves and to discover other services via a DNS or HTTP interface. Register external services such as SaaS providers as well.



Failure Detection

Pairing service discovery with health checking prevents routing requests to unhealthy hosts and enables services to easily provide circuit breakers.



Multi Datacenter

Consul scales to multiple datacenters out of the box with no complicated configuration. Look up services in other datacenters, or keep the request local.

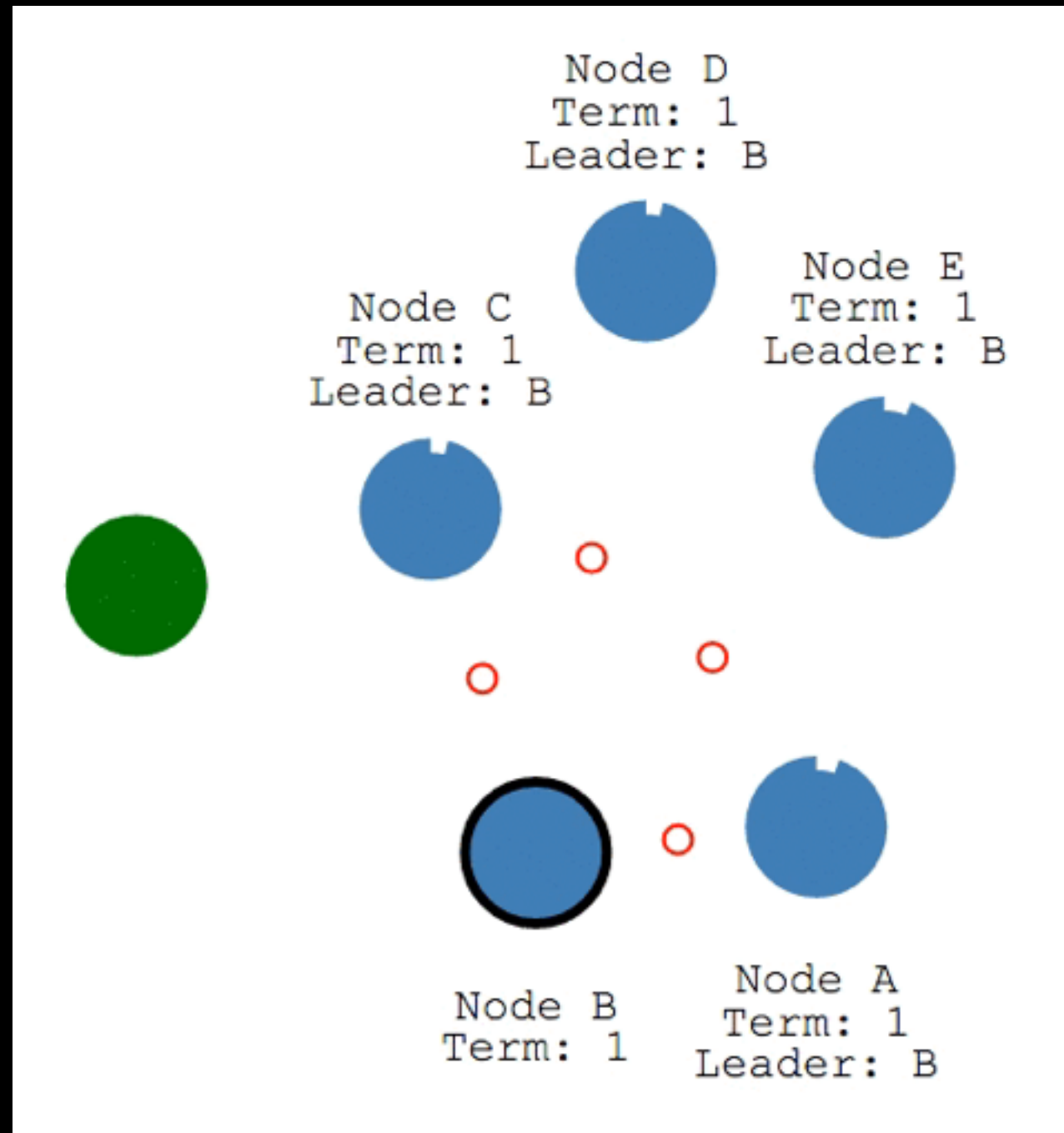


Key/Value Storage

Flexible key/value store for dynamic configuration, feature flagging, coordination, leader election and more. Long poll for near-instant notification of configuration changes.

[HTTP://THESECRETLIVESOFDATA.COM/RAFT/](http://thesecretlivesofdata.com/raft/)

RAFT CONSENSUS



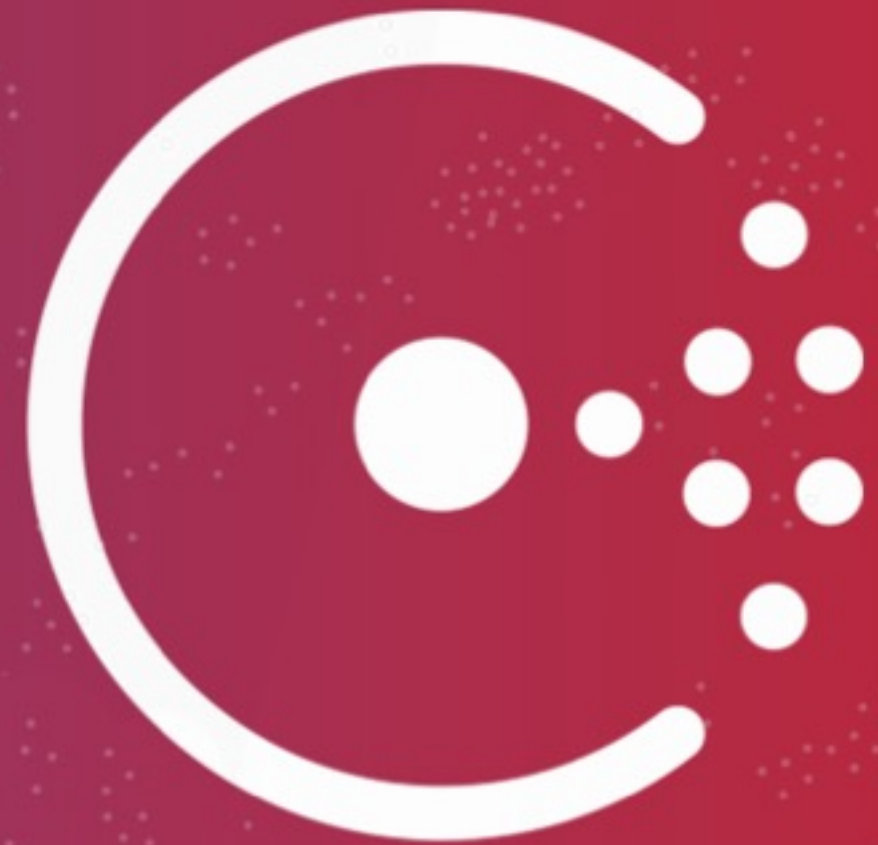
WE WEREN'T SURE.

CAN IT HELP DATADOG?

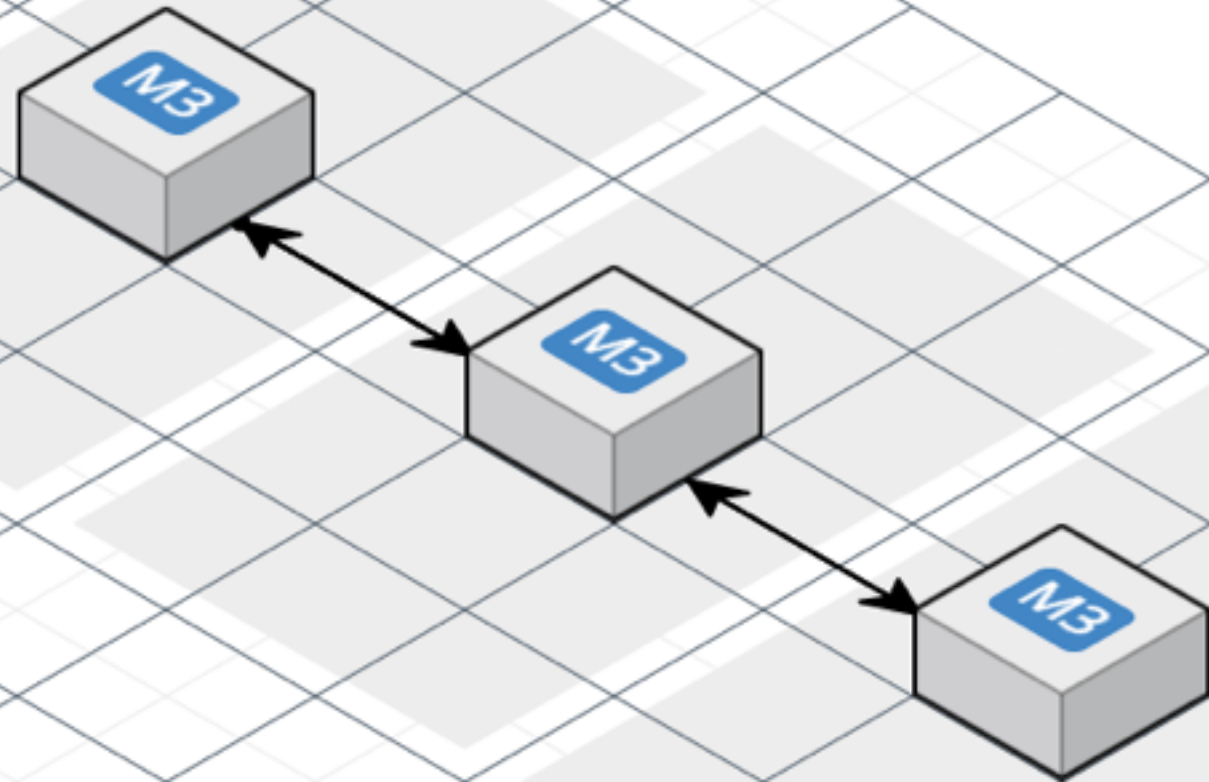


[Intro](#) [Docs](#) [Community](#) [Demo](#) [Download](#)

Service discovery and
configuration made easy.
Distributed, highly
available, and datacenter-
aware.



STAGING



- ~100 nodes in total.
- 3 x m3.medium server nodes
4GB of RAM - 3 ECU - 1 cpu core - SSD drives.

PHASE 1 PLAN

- Initial deploy
- Small amount of services.
- Minimal KV usage
- How will it act?
- Consul 0.4.1.

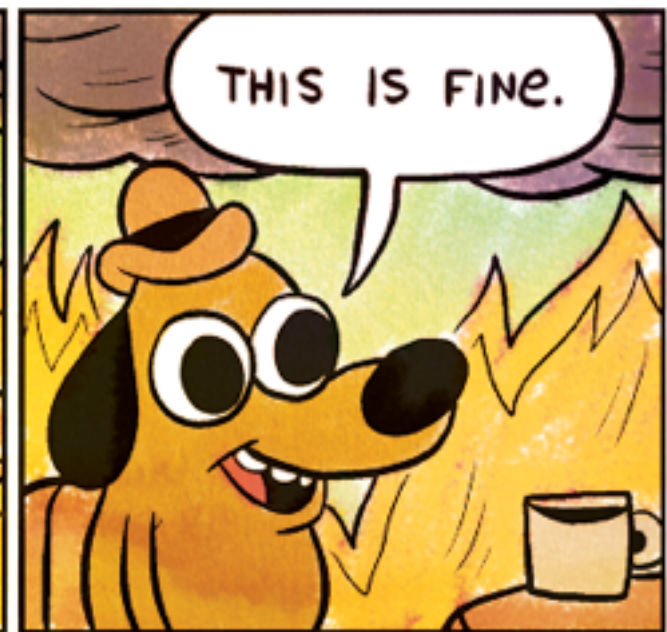


"MONITOR FIRST"

BEFORE PROD

[HTTPS://BLOG.FROESE.ORG/PRESENTATIONS/](https://blog.froese.org/presentations/)

IT'S PROBABLY FINE
SHIP IT



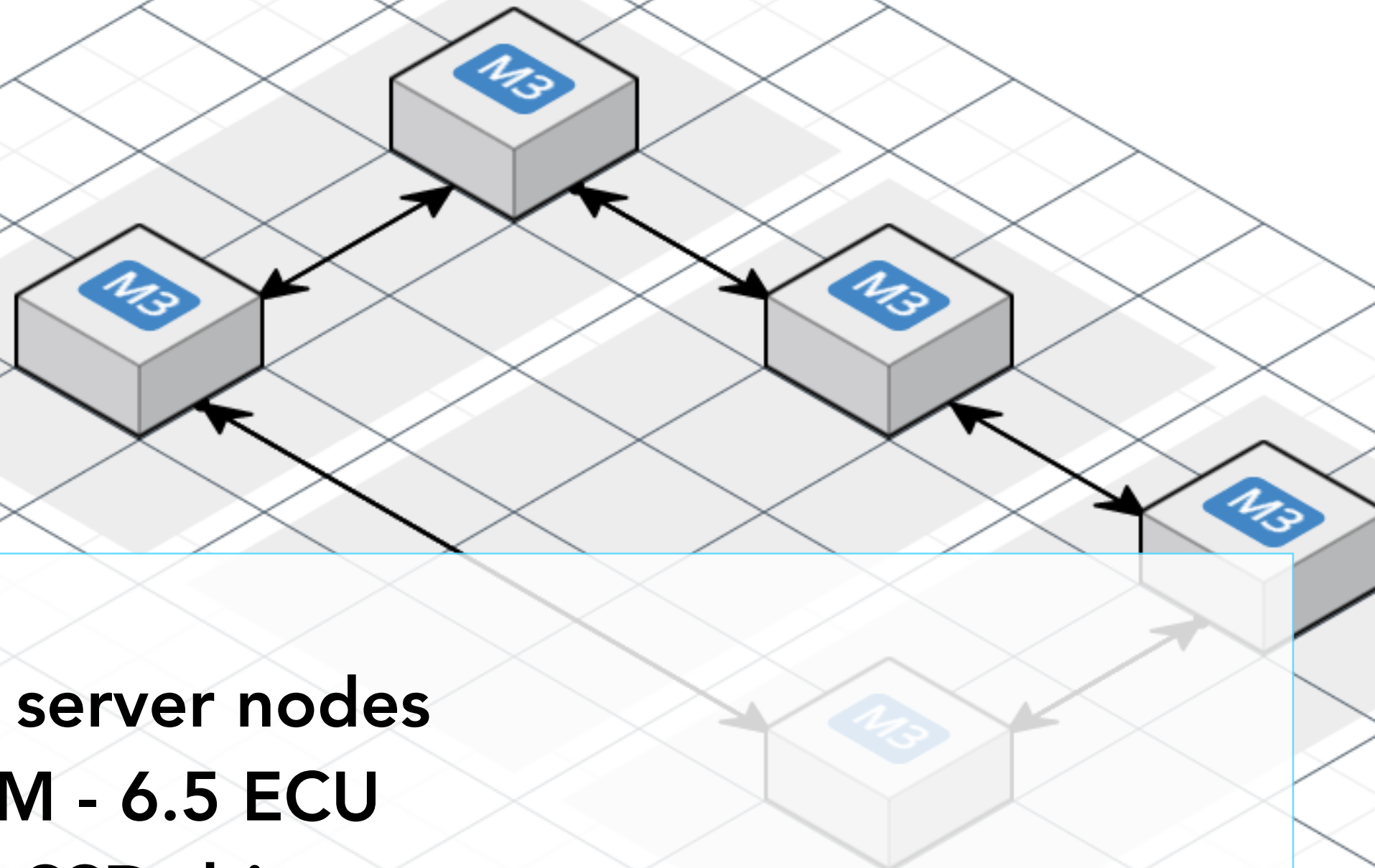
LATE DECEMBER 2014.

DEPLOYED TO PROD

```
desc 'Provision the first Consul bootstrap host'
task :consul_bootstrap, :az do |t, args|
  roles = BASE_ROLES + ['consul-bootstrap']
  size = IS_PROD ? 'm3.large' : 'm3.medium'
  az = args[:az] or pick_az
  provision(roles, 'consul-bootstrap', UBUNTU_1404_HVM_AMI, size, az, PROD_BACKEND, FIRST_EPHEMERAL)
  puts "Provisioned the first Consul bootstrap host."
end

desc 'Provision a Consul server host'
task :consul_server, :az do |t, args|
  roles = BASE_ROLES + ['consul-server']
  size = IS_PROD ? 'm3.large' : 'm3.medium'
  az = args[:az] or pick_az
  provision(roles, 'consul-server', UBUNTU_1404_HVM_AMI, size, az, PROD_BACKEND, FIRST_EPHEMERAL)
  puts "Provisioned a new Consul server host."
end
```

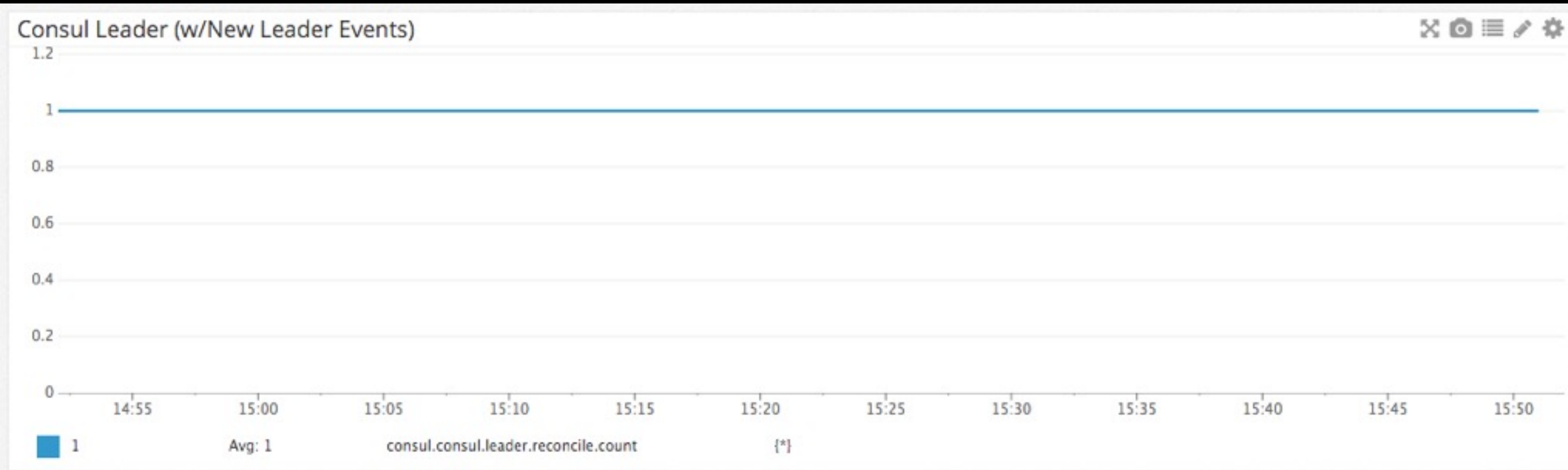
PROD



- **5 x m3.large server nodes**
7.5GB of RAM - 6.5 ECU
2 cpu cores - SSD drives.
- **Rapidly required us to spin up 2 more server nodes**
- it wasn't stable at 3.

AND ALL WAS WELL

IT STABILIZED



ONE OF THE FIRST THINGS WE ADDED.

DATADOG SERVICE

DATADOG SERVICE

```
1 {  
2   "service": {  
3     "name": "datadog",  
4     "tags": [  
5       "consul-server",  
6       "az-us-east-1c"  
7     ],  
8     "check": {  
9       "interval": "60s",  
10      "script": "/bin/true"  
11    }  
12  }  
13 }
```

service.json hosted with ❤️ by GitHub

[view raw](#)

ONE OF THE FIRST THINGS WE ADDED.

DATADOG SERVICE

```
ssh  1
[prod]darron@i-b9162b6a:~$ hostname
i-b9162b6a
[prod]darron@i-b9162b6a:~$ ssh_to_role zookeeper
Warning: Permanently added 'zookeeper.datadog.service.consul,10.239.179.108' (ECDSA) to the list of known hosts.
[prod]darron@i-f7c9f309:~$ hostname
i-f7c9f309
[prod]darron@i-f7c9f309:~$ hosts_by_role zookeeper
i-07c182fc
i-8b2f0e67
i-ae218b43
i-f0bbb80e
i-f7c9f309
[prod]darron@i-f7c9f309:~$
```

TOYED WITH AND DISABLED

CONSUL EXEC

```
ssh 981
[staging]darron@i-86402c2e:~$ consul exec -service datadog -tag consul-server w
i-ccce987b: 23:51:24 up 52 days, 18:11, 0 users, load average: 0.00, 0.01, 0.05
i-ccce987b: USER      TTY      FROM              LOGIN@   IDLE   JCPU   PCPU   WHAT
i-ccce987b:
==> i-ccce987b: finished with exit code 0
i-bc13b90c: 23:51:24 up 52 days, 18:30, 0 users, load average: 0.04, 0.05, 0.05
i-bc13b90c: USER      TTY      FROM              LOGIN@   IDLE   JCPU   PCPU   WHAT
i-bc13b90c:
==> i-bc13b90c: finished with exit code 0
i-1dc56ea3: 23:51:24 up 52 days, 18:20, 0 users, load average: 0.00, 0.01, 0.05
i-1dc56ea3: USER      TTY      FROM              LOGIN@   IDLE   JCPU   PCPU   WHAT
i-1dc56ea3:
==> i-1dc56ea3: finished with exit code 0
3 / 3 node(s) completed / acknowledged
[staging]darron@i-86402c2e:~$
```


STRONGLY CONSISTENT
KEY VALUE STORE AVAILABLE
ON LOCALHOST WITH AN HTTP QUERY.

GIT2CONSUL

[HTTPS://GITHUB.COM/CIMPRESS-MCP/GIT2CONSUL](https://github.com/cimpress-mcp/git2consul)

GIT2CONSUL

CONSUL-CONFIG

Configuration we feed into the Consul KV store with git2consul. — Edit

 3,450 commits

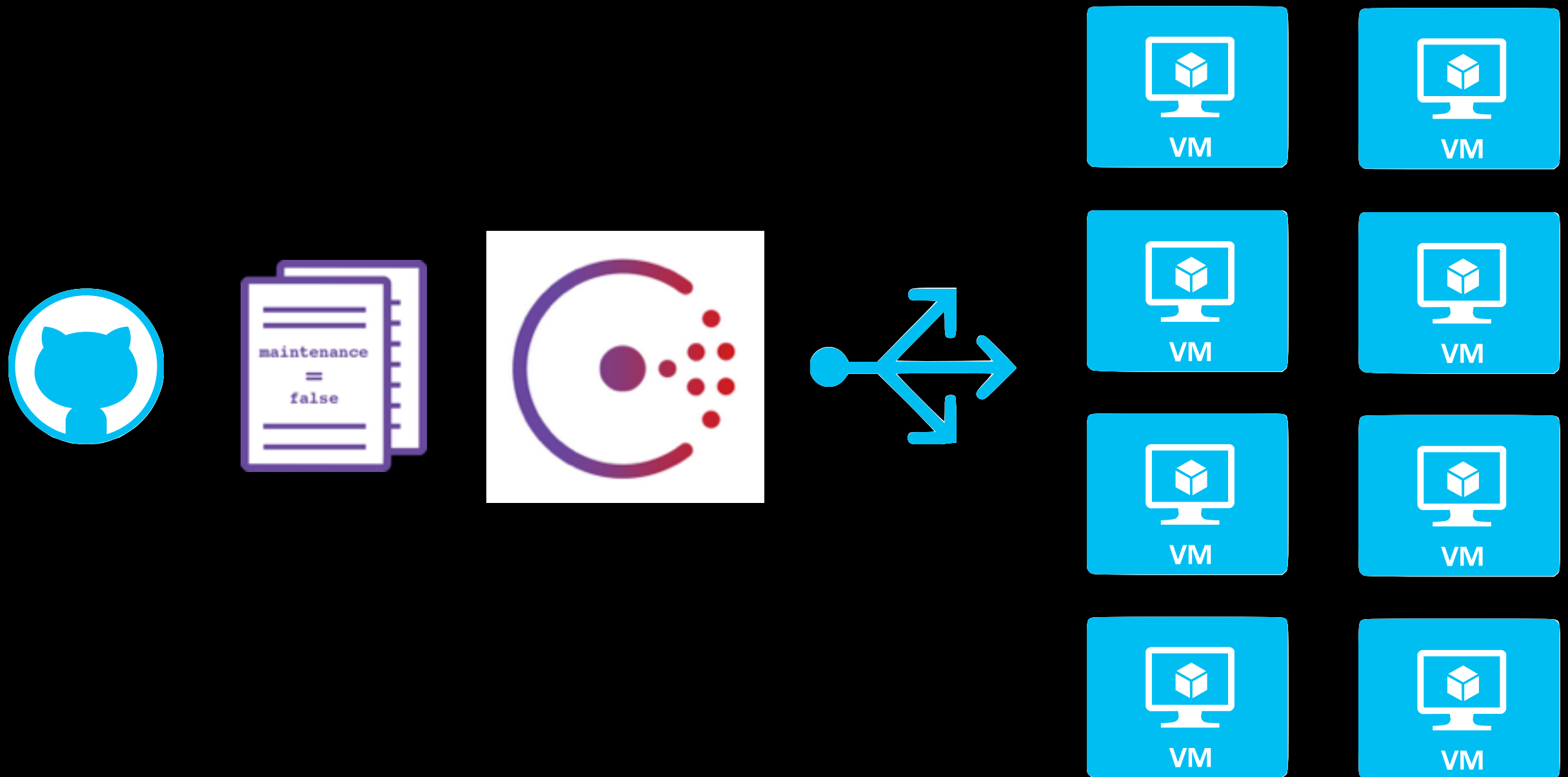
 38 branches

 0 releases

 56 contributors

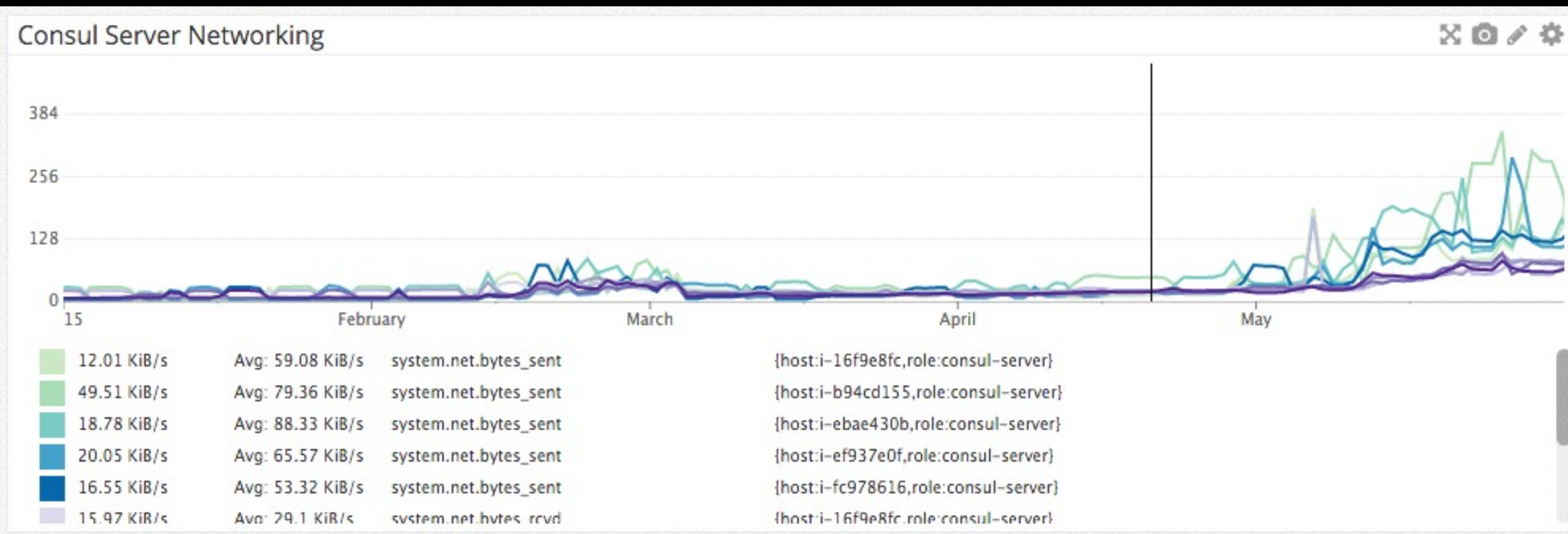
HOW IT WORKS

GIT2CONSUL + CONSUL-CONFIG



UP AND TO THE RIGHT

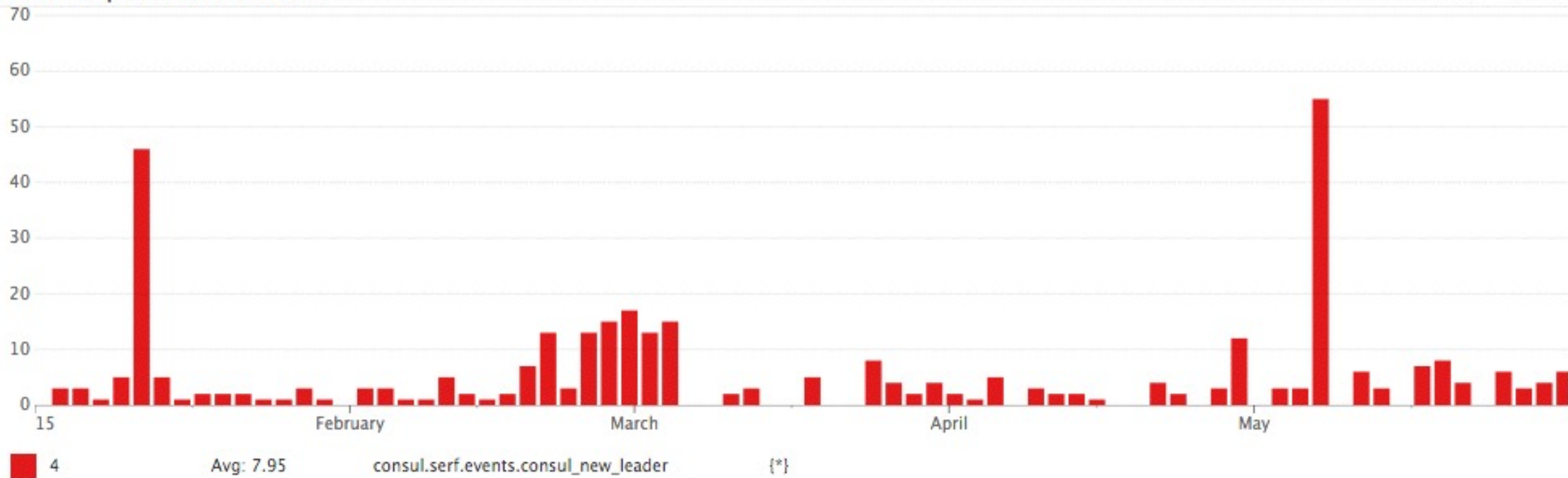
MORE AND MORE USE



PRETTY COMMON - MOSTLY HARMLESS

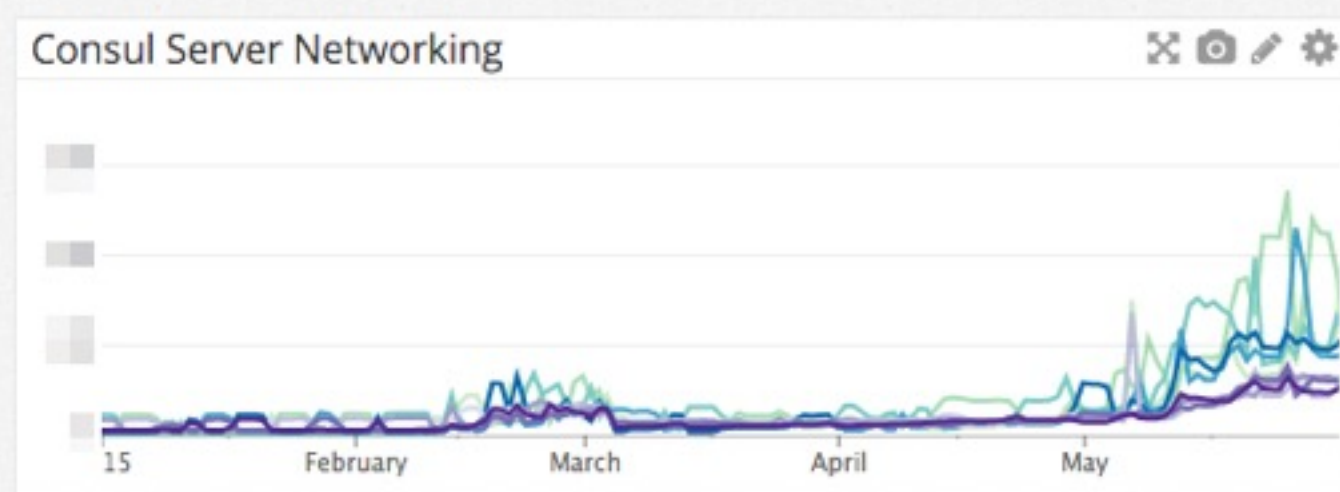
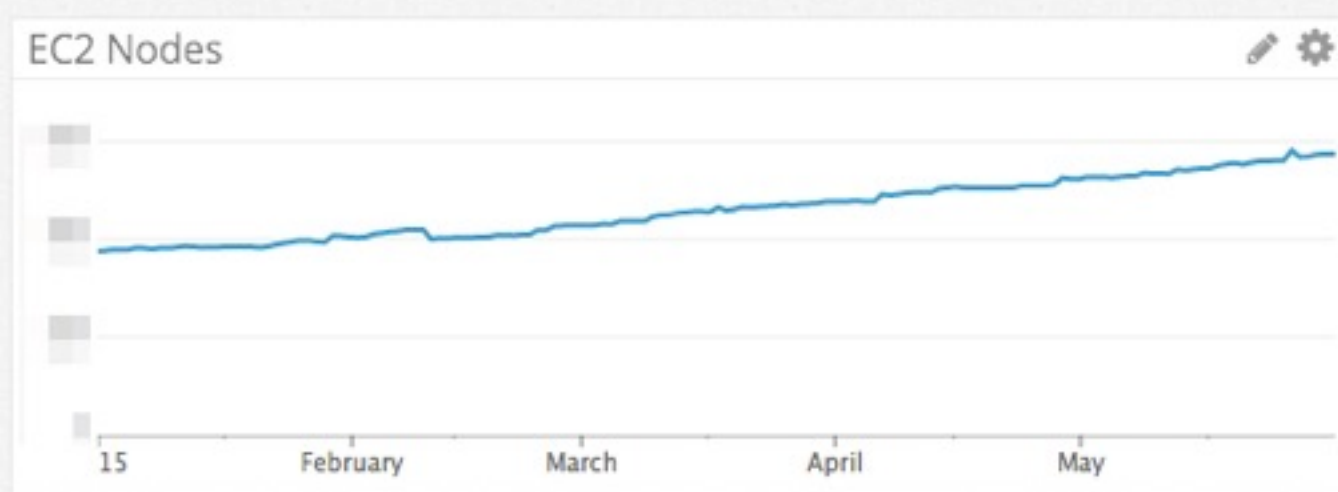
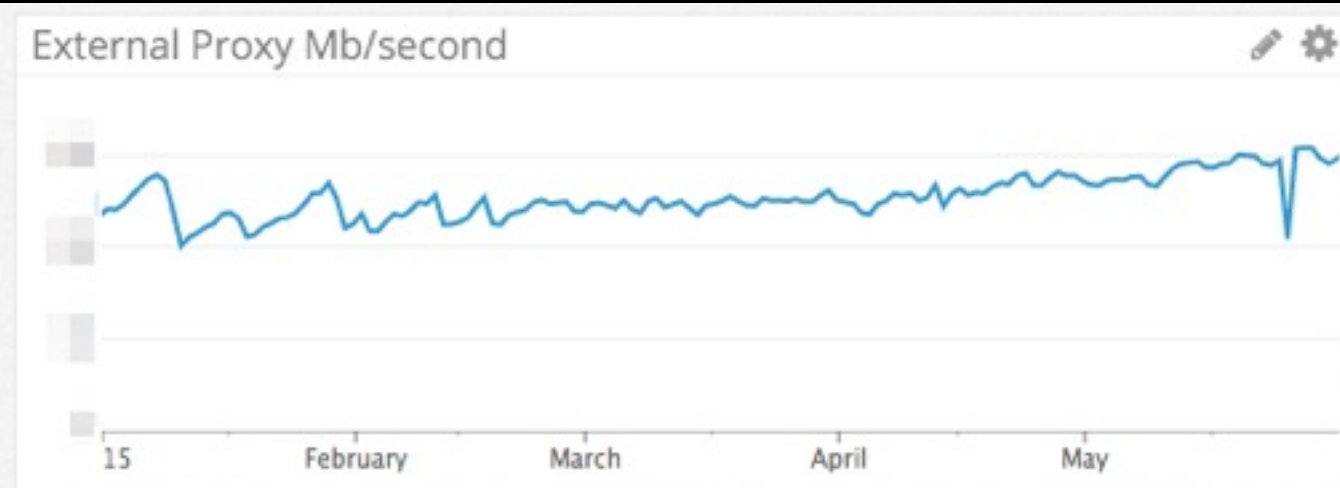
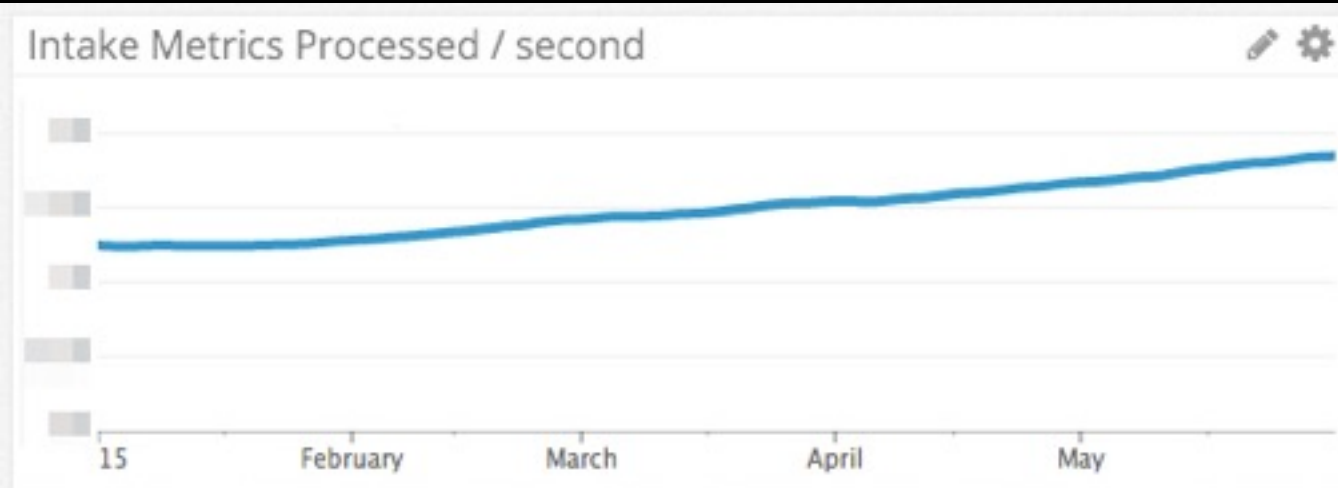
LEADERSHIP TRANSITIONS

Leadership Transition Event



ALMOST 600 NODES

MAY 2015



WE'RE GETTING SERIOUS NOW

SERVICE REGISTRATION

EXPAND

bunk	24 passing
cassandra	60 passing
cole	2 passing
consul	5 passing
ctx-pshard0	4 passing
ctx-pshard1	6 passing
ctx-pshard2	6 passing
ctx-pshard3	6 passing
ctx-pshard4	4 passing
datadog	2068 passing
delancie-backend	1 passing
dogweb-backend	2 passing

cassandra

TAGS

az-us-east-1b, az-us-east-1e, az-us-east-1a

NODES

i-9119fe7e	10.186.190.71	2 passing
Service 'cassandra' check service:cassandra		passing
Serf Health Status serfHealth		passing
i-f501a304	10.95.168.48	2 passing
Service 'cassandra' check service:cassandra		passing
Serf Health Status serfHealth		passing
i-60dcca9b	10.35.179.159	2 passing
Service 'cassandra' check service:cassandra		passing
Serf Health Status serfHealth		passing

CURL/HTTP LOOKUP

SERVICE DISCOVERY

```
ssh 961
[staging]darron@i-86402c2e:~$ curl -s http://127.0.0.1:8500/v1/catalog/service/consul |
jq .
[
  {
    "Node": "i-1dc56ea3",
    "Address": "10.186.190.143",
    "ServiceID": "consul",
    "ServiceName": "consul",
    "ServiceTags": [],
    "ServiceAddress": "",
    "ServicePort": 8300
  },
  {
    "Node": "i-bc13b90c",
    "Address": "10.233.144.121",
    "ServiceID": "consul",
    "ServiceName": "consul",
    "ServiceTags": [],
    "ServiceAddress": "",
    "ServicePort": 8300
  },
  {
    "Node": "i-ccce987b",
    "Address": "10.47.151.240",
```


SERVICE DISCOVERY

```
ssh 361
[staging]darron@i-86402c2e:~$ dig consul.service.consul

; <<>> DiG 9.9.5-3ubuntu0.6-Ubuntu <<>> consul.service.consul
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 27619
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 3, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;consul.service.consul.      IN      A

;; ANSWER SECTION:
consul.service.consul.  10      IN      A      10.47.151.240
consul.service.consul.  10      IN      A      10.233.144.121
consul.service.consul.  10      IN      A      10.186.190.143

;; Query time: 1 msec
;; SERVER: 127.0.0.1#53(127.0.0.1)
;; WHEN: Wed Jan 06 01:39:13 UTC 2016
;; MSG SIZE  rcvd: 150

[staging]darron@i-86402c2e:~$
```

IN AND OUT OF THE SERVICE CATALOG

WOULD IT FLAP?

```
1  {
2    "service": {
3      "name": "datadog",
4      "tags": [
5        "consul-server",
6        "az-us-east-1c"
7      ],
8      "check": {
9        "interval": "60s",
10       "script": "/bin/true"
11     }
12   }
13 }
```



service.json hosted with ❤ by GitHub

[view raw](#)

IN AND OUT OF THE SERVICE CATALOG

NO. IT DID NOT FLAP

```
1  {
2    "service": {
3      "name": "cassandra",
4      "port": 9160,
5      "tags": [
6        "az-us-east-1a"
7      ],
8      "check": {
9        "interval": "10s",
10       "script": "nodetool info"
11     }
12   }
```

cassandra-consul.json hosted with ❤ by GitHub

[view raw](#)

WORRIED ABOUT SPEED

USING DNS

FRONTED CONSUL'S DNS RESOLVER

DNSMASQ

```
ssh 361
[staging]darron@i-86402c2e:/etc/dnsmasq.d$ cat 10-consul
server=/consul/127.0.0.1#8600
[staging]darron@i-86402c2e:/etc/dnsmasq.d$ sudo cat /etc/consul
.d/ttl.json
{
  "dns_config": {
    "allow_stale": true,
    "service_ttl": {
      "*": "10s"
    }
  }
}
[staging]darron@i-86402c2e:/etc/dnsmasq.d$
```

CONSUL-TEMPLATE

CONSUL_DNS_BACKUP

(THE HOSTS FILE)

EVEN IN STAGING

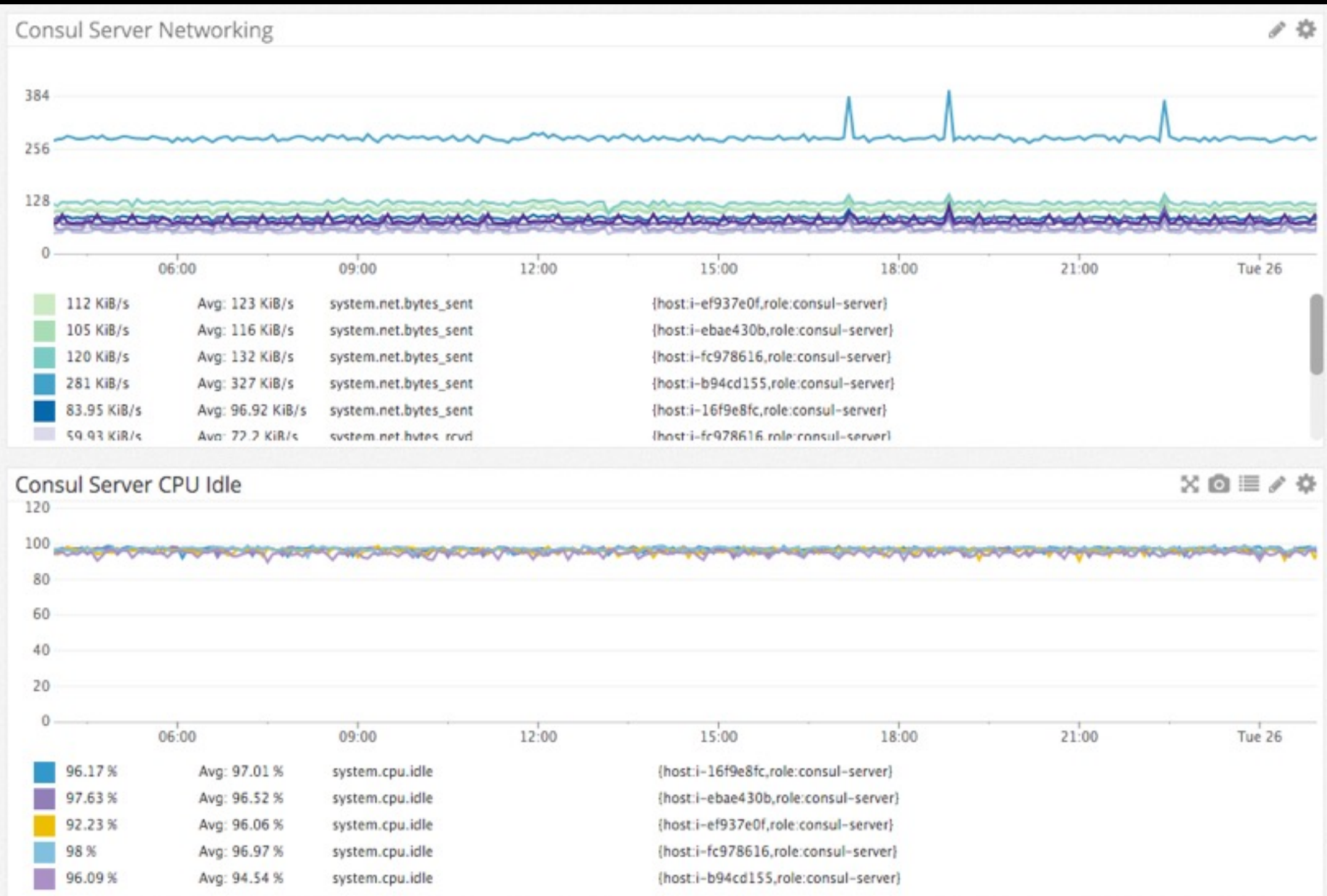
NOT SUCCESSFUL



BUILT ON THE SERVER NODES

USE THE KV STORE TO
DISTRIBUTE.

IT WORKS REALLY WELL



IT WAS A BIT HAIRY

WITHOUT RATE LIMITING



NONE AT ALL

NO LEADERSHIP
TRANSITIONS

"LET'S CLEAN THIS UP"

THE VERY NEXT DAY

Add pre-commit config to keep it clean

[Browse files](#)

JSON files are now pretty and standardized.

 master (#48)

committed on Jun 11

1 parent [d925030](#) commit [1d6da99fdb97675a26c8e88d5010b269c802b002](#)

 Showing **92 changed files** with 1,030 additions and 430 deletions.

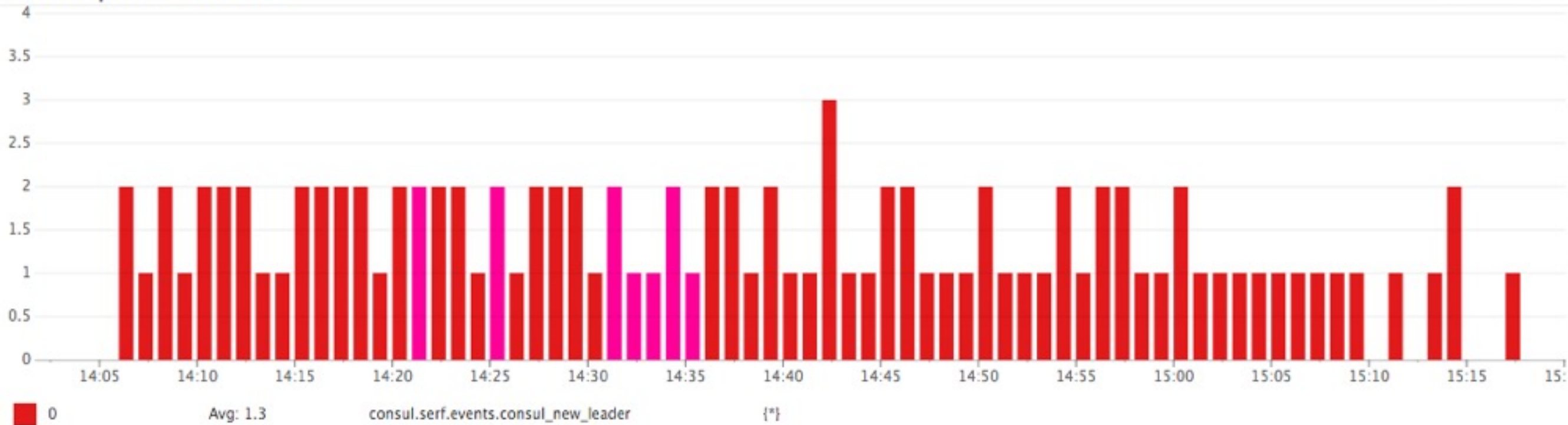
Unified

Split

CAUSING LEADERSHIP TRANSITIONS

READ-PRESSURE

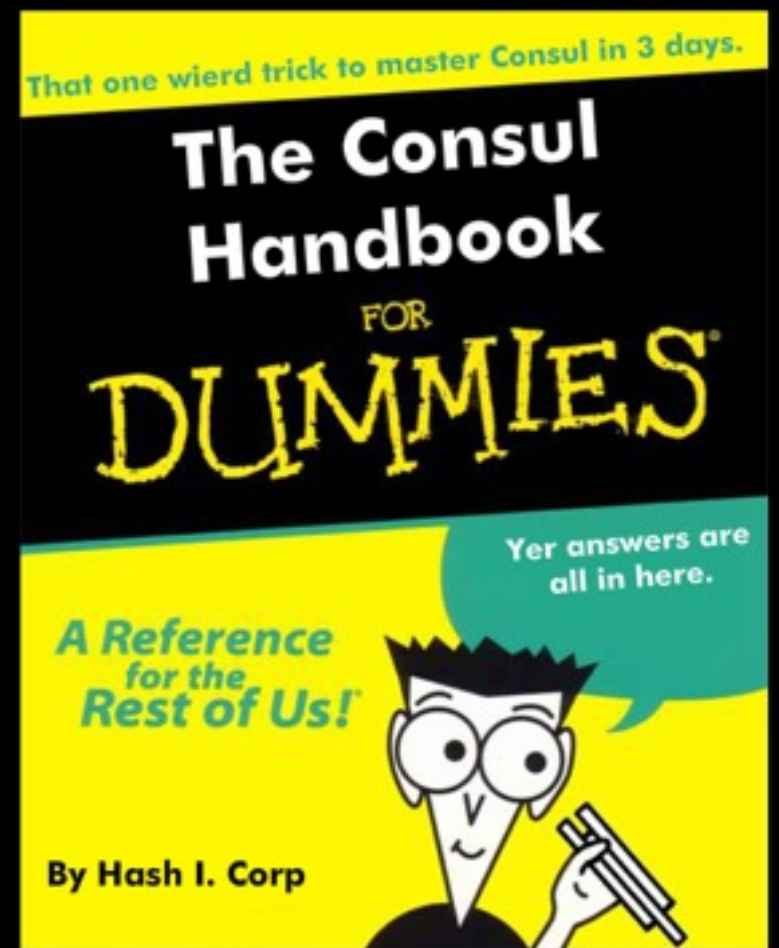
Leadership Transition Event



THE EDGE WAS A LITTLE BLOODY

CONSUL IS NEW

THERE WAS VERY LITTLE
REAL WORLD INFORMATION ABOUT IT.



DON'T DO THIS - WITH CONSUL 0.5.X

LOTS OF SMALL KEYS

 any_filter_array	enable any_filter_array everywhere	4 months ago
 api_log_info_writes	[prod][api log] Disable info writes	2 months ago
 appdirect_marketplace	[prod] enable appdirect endpoints	3 months ago
 automute_filter_aws_tags	[Prod] Enabling EC2 Service checks for 40% of orgs	5 months ago
 aws_auto_scaling_tags_fix	updates tree structure	6 months ago
 aws_billing_linked_accounts	[AWS] Billing - add a feature flag to keep billing metrics related to...	5 months ago
 aws_cloud_provider_tag	updates tree structure	6 months ago
 aws_custom_cw_metrics_n...	updates tree structure	6 months ago
 aws_ec2_check_spurious_t...	[ec2] check spurious tags for everyone	a day ago
 aws_ec2_custom_tags_issue	[AWS] enable workaround for missing ec2 custom tags	2 months ago
 aws_ec2_filter_with_accou...	updates tree structure	6 months ago
 aws_ec2_metrics_all_stats	updates tree structure	6 months ago
 aws_ecs_cluster_tags	Revert "Revert "[ECS] Turn cluster tags for everyone""	5 months ago
 aws_elb_host_tags	updates tree structure	6 months ago
 aws_elb_metrics_all_stats	[AWS] ELB - retrieve all stats for an org	3 months ago
 aws_elb_new_hostname	Removed Campain 17774 and Groundwork 21953 from aws new host name list	2 months ago
 aws_filter_aggregates	enabled aws_filter_aggregtes for org 38736	22 days ago

WE UPSIZED OUR VMS

ONCE WE UNDERSTOOD

Showing 2 changed files with 3 additions and 3 deletions.

Unified Split

4 site-cookbooks/dd-provision/files/default/provision/Rakefile

View



@@ -824,7 +824,7 @@ namespace :provision do

824	824	desc 'Provision the first Consul bootstrap host'
825	825	task :consul_bootstrap, :az do _t, args
826	826	roles = BASE_ROLES + ['consul-bootstrap']
827		- size = IS_PROD ? 'm3.large' : 'm3.medium'
	827	+ size = IS_PROD ? 'c3.xlarge' : 'm3.medium'
828	828	az = args[:az] pick_az
829	829	provision(roles, 'consul-bootstrap', UBUNTU_1404_HVM_AMI, size, az, PROD_BACKEND, FIRST_EPHEMERAL)
830	830	puts 'Provisioned the first Consul bootstrap host.'



@@ -833,7 +833,7 @@ namespace :provision do

833	833	desc 'Provision a Consul server host'
834	834	task :consul_server, :az do _t, args
835	835	roles = BASE_ROLES + ['consul-server']
836		- size = IS_PROD ? 'm3.large' : 'm3.medium'
	836	+ size = IS_PROD ? 'c3.xlarge' : 'm3.medium'
837	837	az = args[:az] pick_az
838	838	provision(roles, 'consul-server', UBUNTU_1404_HVM_AMI, size, az, PROD_BACKEND, FIRST_EPHEMERAL)
839	839	puts 'Provisioned a new Consul server host.'



2 site-cookbooks/dd-provision/metadata.rb

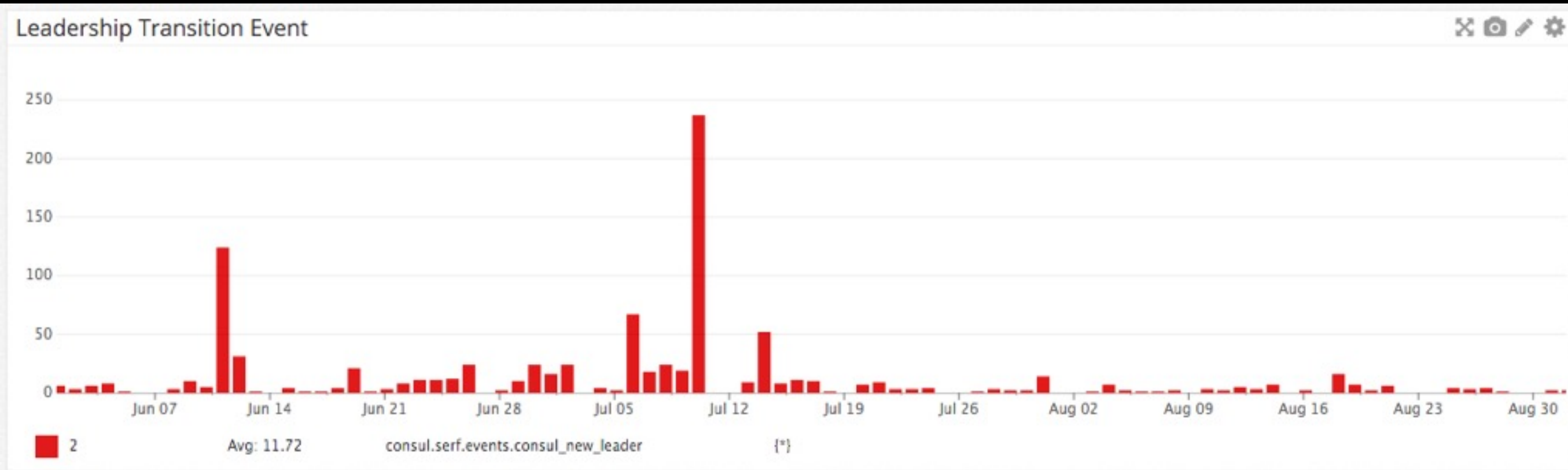
View



@@ -4,4 +4,4 @@

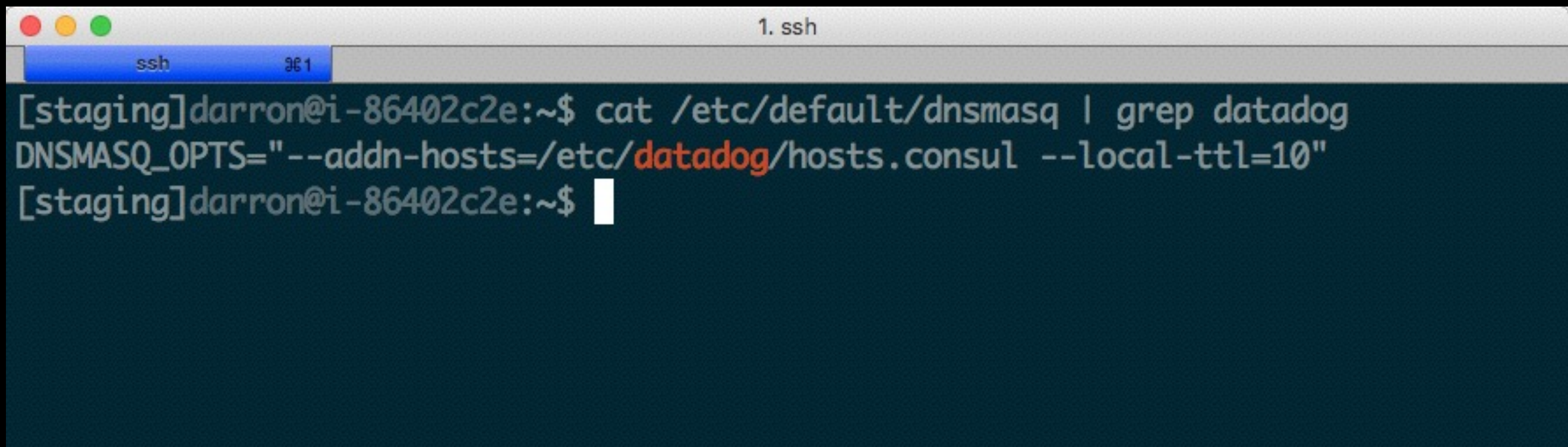
INSTALLED LARGER SERVER NODES

THINGS QUIETED DOWN



THE DNSMASQ THING...

HOW DID IT WORK?



```
1. ssh  
ssh 961  
[staging]darron@i-86402c2e:~$ cat /etc/default/dnsmasq | grep datadog  
DNSMASQ_OPTS="--addn-hosts=/etc/datadog/hosts.consul --local-ttl=10"  
[staging]darron@i-86402c2e:~$
```

ASKS ONCE EVERY 10 SECONDS

DNSMASQ HONORS THE
CONSUL TTL

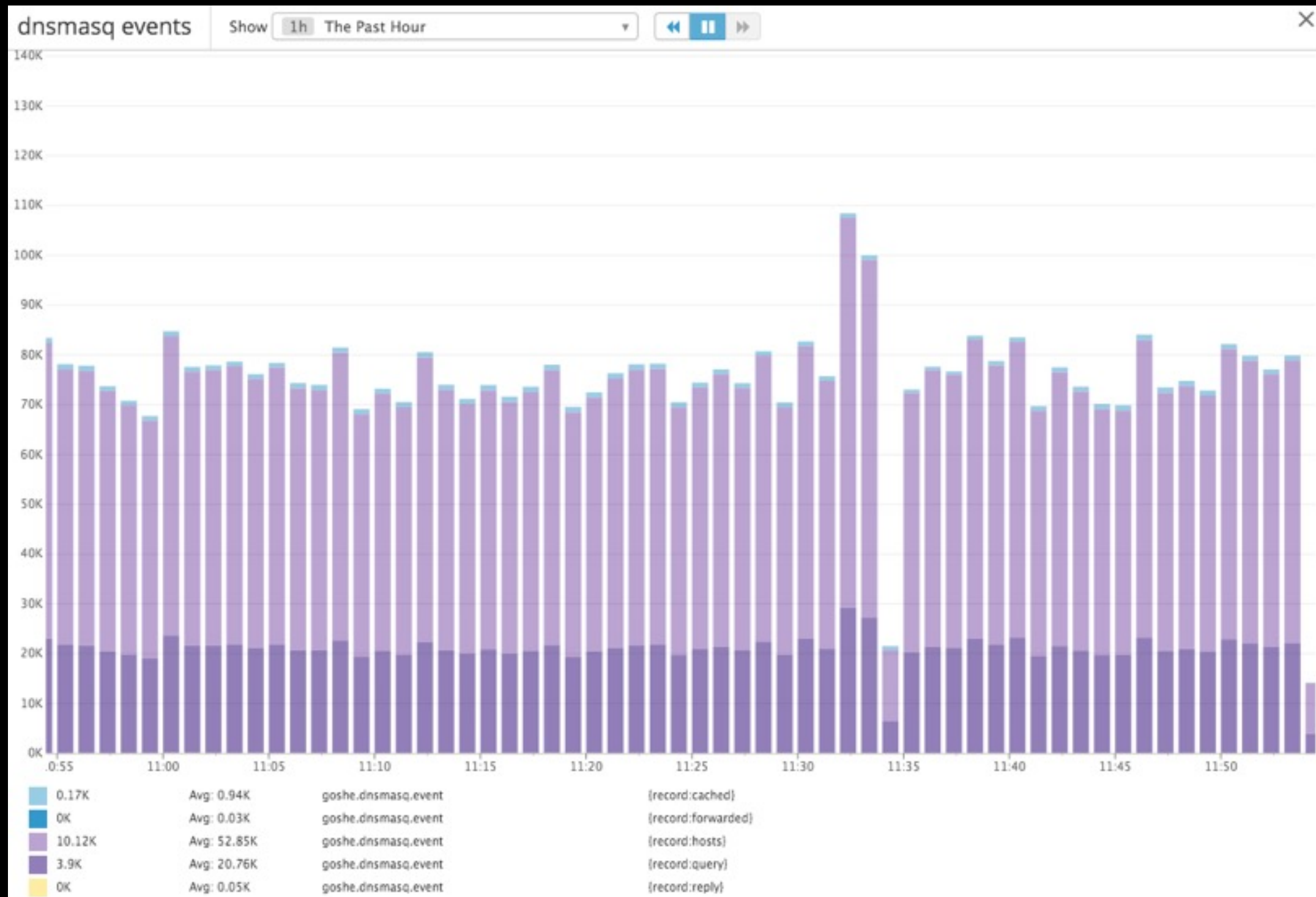
WELL

IT RESPONDS QUICKLY

```
ssh  %1
[staging]darron@i-1dc56ea3:~$ ./main bunk.datadog.service.consul
bunk.datadog.service.consul. 10 IN A 10.7.201.208 in Rtt: 643.448µs
bunk.datadog.service.consul. 10 IN A 10.148.10.224 in Rtt: 643.448µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.datadog.service.consul
bunk.datadog.service.consul. 5 IN A 10.148.10.224 in Rtt: 211.277µs
bunk.datadog.service.consul. 5 IN A 10.7.201.208 in Rtt: 211.277µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.datadog.service.consul
bunk.datadog.service.consul. 2 IN A 10.7.201.208 in Rtt: 223.807µs
bunk.datadog.service.consul. 2 IN A 10.148.10.224 in Rtt: 223.807µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.datadog.service.consul
bunk.datadog.service.consul. 10 IN A 10.148.10.224 in Rtt: 688.294µs
bunk.datadog.service.consul. 10 IN A 10.7.201.208 in Rtt: 688.294µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.service.consul
bunk.service.consul. 10 IN A 10.7.201.208 in Rtt: 88.793µs
bunk.service.consul. 10 IN A 10.148.10.224 in Rtt: 88.793µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.service.consul
bunk.service.consul. 10 IN A 10.148.10.224 in Rtt: 221.247µs
bunk.service.consul. 10 IN A 10.7.201.208 in Rtt: 221.247µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.service.consul
bunk.service.consul. 10 IN A 10.7.201.208 in Rtt: 225.727µs
bunk.service.consul. 10 IN A 10.148.10.224 in Rtt: 225.727µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.service.consul
bunk.service.consul. 10 IN A 10.148.10.224 in Rtt: 112.676µs
bunk.service.consul. 10 IN A 10.7.201.208 in Rtt: 112.676µs
[staging]darron@i-1dc56ea3:~$ ./main bunk.service.consul
bunk.service.consul. 10 IN A 10.7.201.208 in Rtt: 182.770µs
```


DNSMASQ AND THE MAGICAL HOSTS FILE

HOW DID IT WORK?



WEREN'T SURE IF
WE WERE GOING TO CONTINUE

PEOPLE WERE A BIT
SCARED

THEY DROPPED OUT OF THE CATALOG

NODES WERE GOING
DEAF

I WAS VISITING THE
OFFICE & HEARD SOME GRUMBLING

BUT I COULD FINALLY
DUPLICATE THE
PROBLEM

HUGE THANKS TO JAMES AND ARMON FOR ALL THEIR HELP!

HASHICORP LENT A HAND



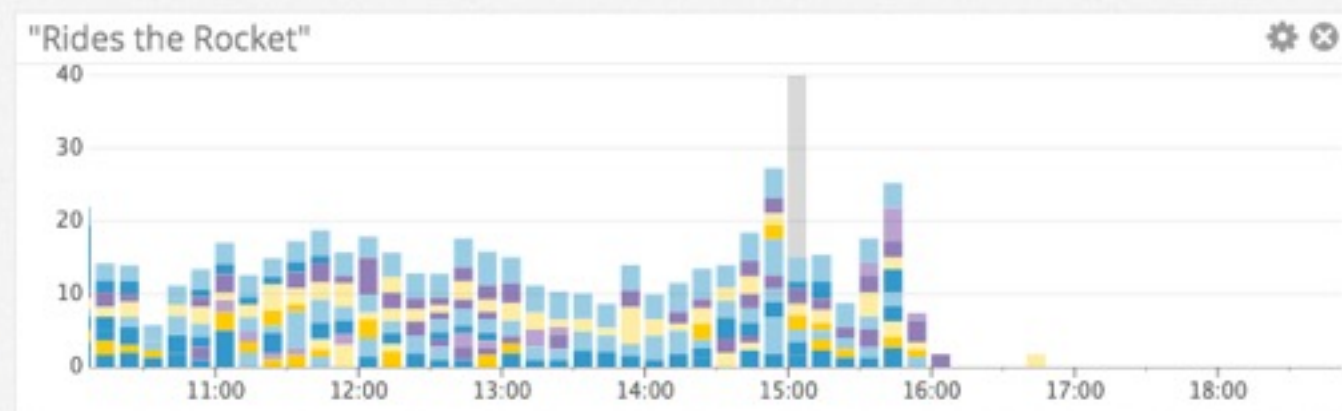
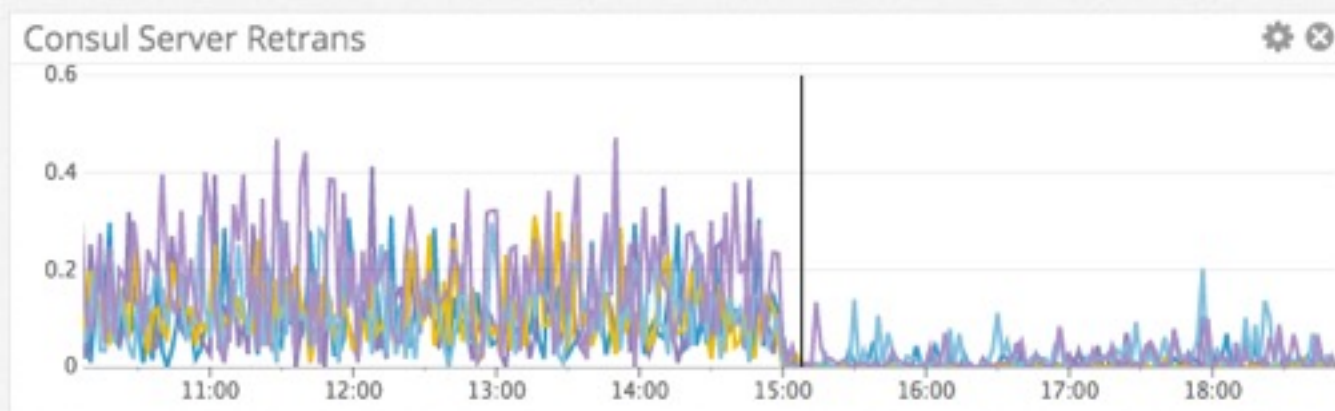
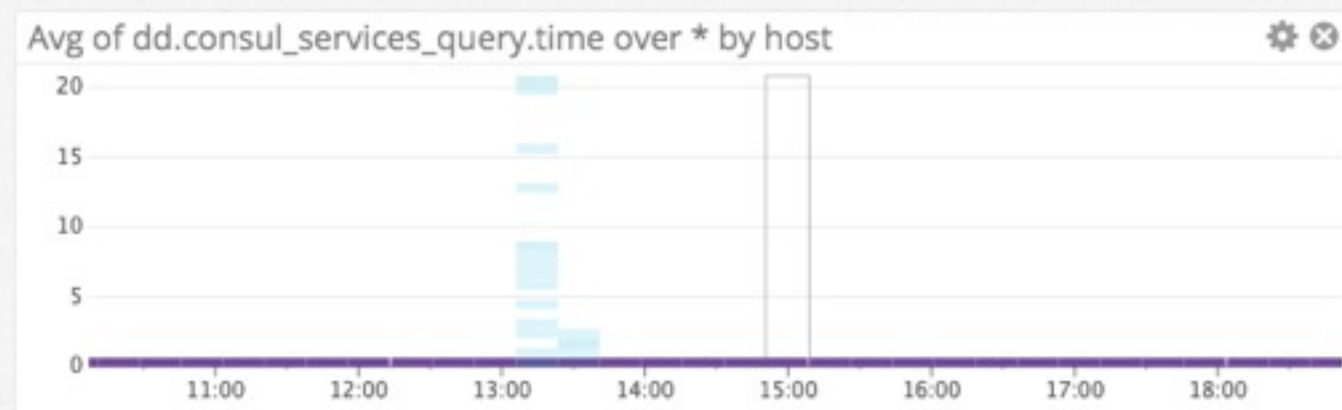
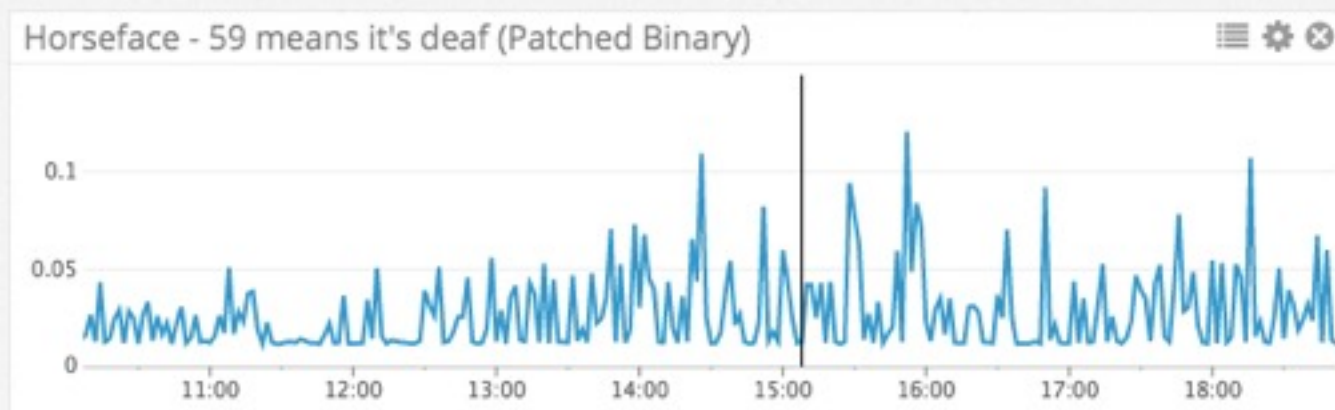
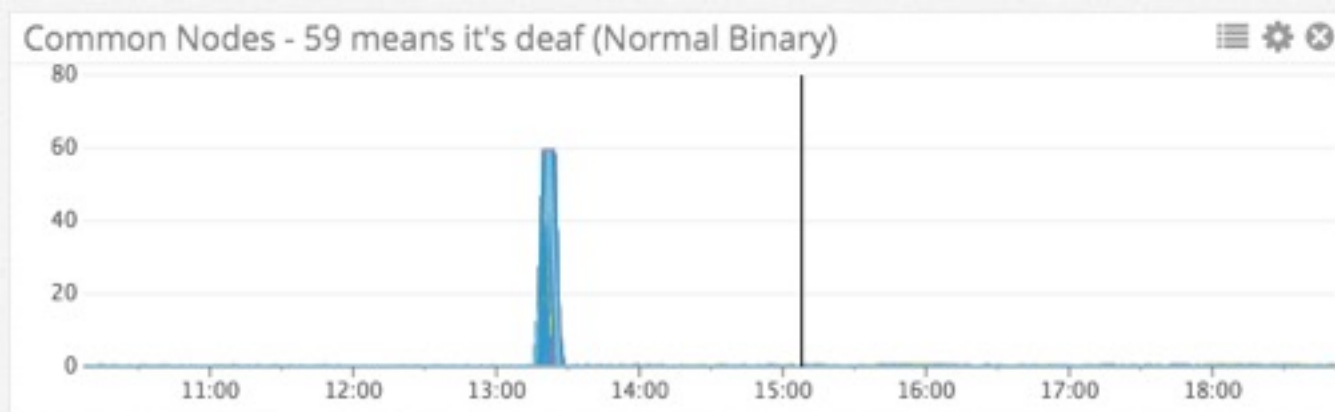
BUT THE KEY WAS QUAKE RELATED.

2 DEADLOCKS FIXED

```
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.220079] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.224758] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.229763] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.234064] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.238479] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.244155] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.250585] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:29 i-6d1f93c1 kernel: [2388778.254697] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:35 i-6d1f93c1 kernel: [2388784.860091] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:39 i-6d1f93c1 kernel: [2388788.615393] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:39 i-6d1f93c1 kernel: [2388788.620076] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:39 i-6d1f93c1 kernel: [2388788.624920] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:39 i-6d1f93c1 kernel: [2388788.633086] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:39 i-6d1f93c1 kernel: [2388788.637685] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:39 i-6d1f93c1 kernel: [2388788.641706] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:39 i-6d1f93c1 kernel: [2388788.654226] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:40 i-6d1f93c1 kernel: [2388788.658702] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:40 i-6d1f93c1 kernel: [2388788.663259] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:46 i-6d1f93c1 kernel: [2388795.804128] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:52 i-6d1f93c1 kernel: [2388802.348119] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:29:54 i-6d1f93c1 kernel: [2388804.508087] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:30:05 i-6d1f93c1 kernel: [2388815.452115] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:30:21 i-6d1f93c1 kernel: [2388830.684086] xen_netfront: xennet: skb rides the rocket:
Aug 18 16:30:32 i-6d1f93c1 kernel: [2388841.628086] xen_netfront: xennet: skb rides the rocket:
```

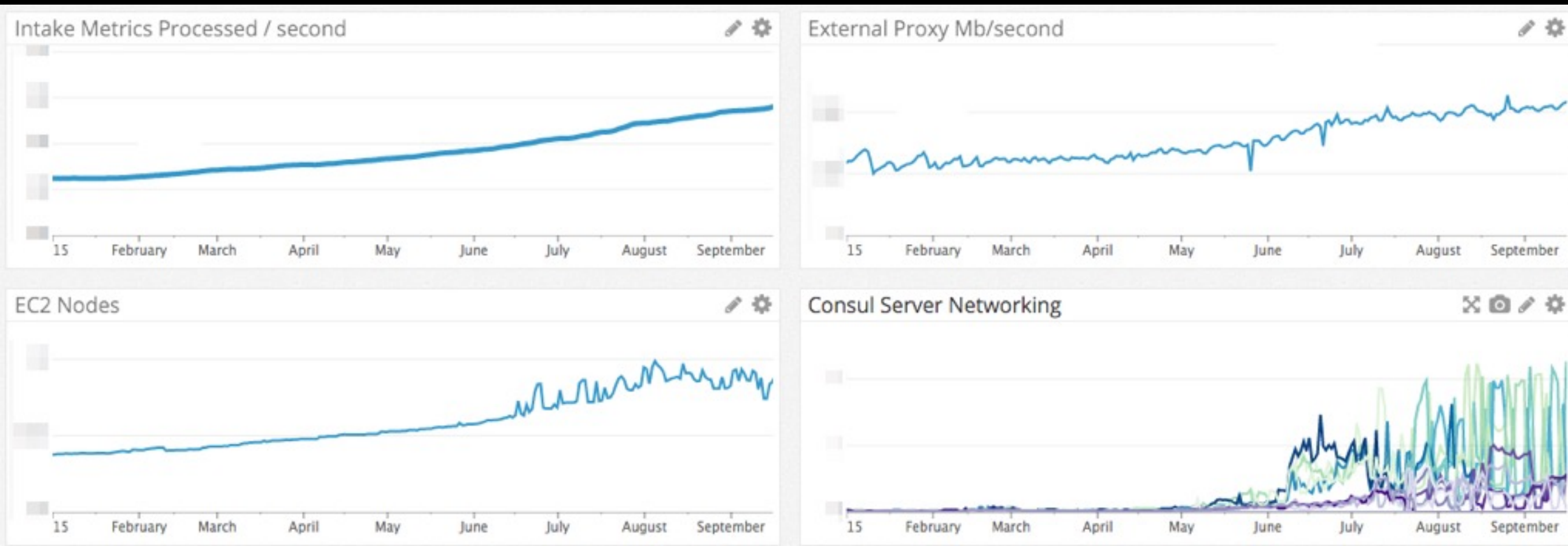

FOR THE MOST PART

AND ALL WAS RIGHT



850 NODES - MOSTLY STABLE - BUT CAUTIOUS

OCTOBER 2015



MADE RETIRING &
SWAPPING SOME SERVICES EASIER

CONSUL-CONFIG

HELPED AS WE GREW

LIKE CONSUL EXEC BUT WITH SECURITY++

STARTED USING EVENTS

```
ssh 361
[staging]darron@i-86402c2e:~$ sudo cat /etc/consul.d/event-watch-apt-update.json
{
  "watches": [
    {
      "type": "event",
      "name": "apt-update",
      "handler": "sifter run -d true -e 'apt-get update'"
    }
  ]
}
[staging]darron@i-86402c2e:~$ consul event -service datadog -name apt-update
Event ID: 8ef0bbc9-b7e4-fbc9-87c1-b82f76d78638
[staging]darron@i-86402c2e:~$
```

RUN 1 OF N PROCESSES - WITH HOT SPARES

CONSUL LOCK



WAITING



WAITING



RUNNING

RUN 1 OF N PROCESSES - WITH HOT SPARES

CONSUL LOCK



WAITING



RUNNING



CRASHES

RUN 1 OF N PROCESSES - WITH HOT SPARES

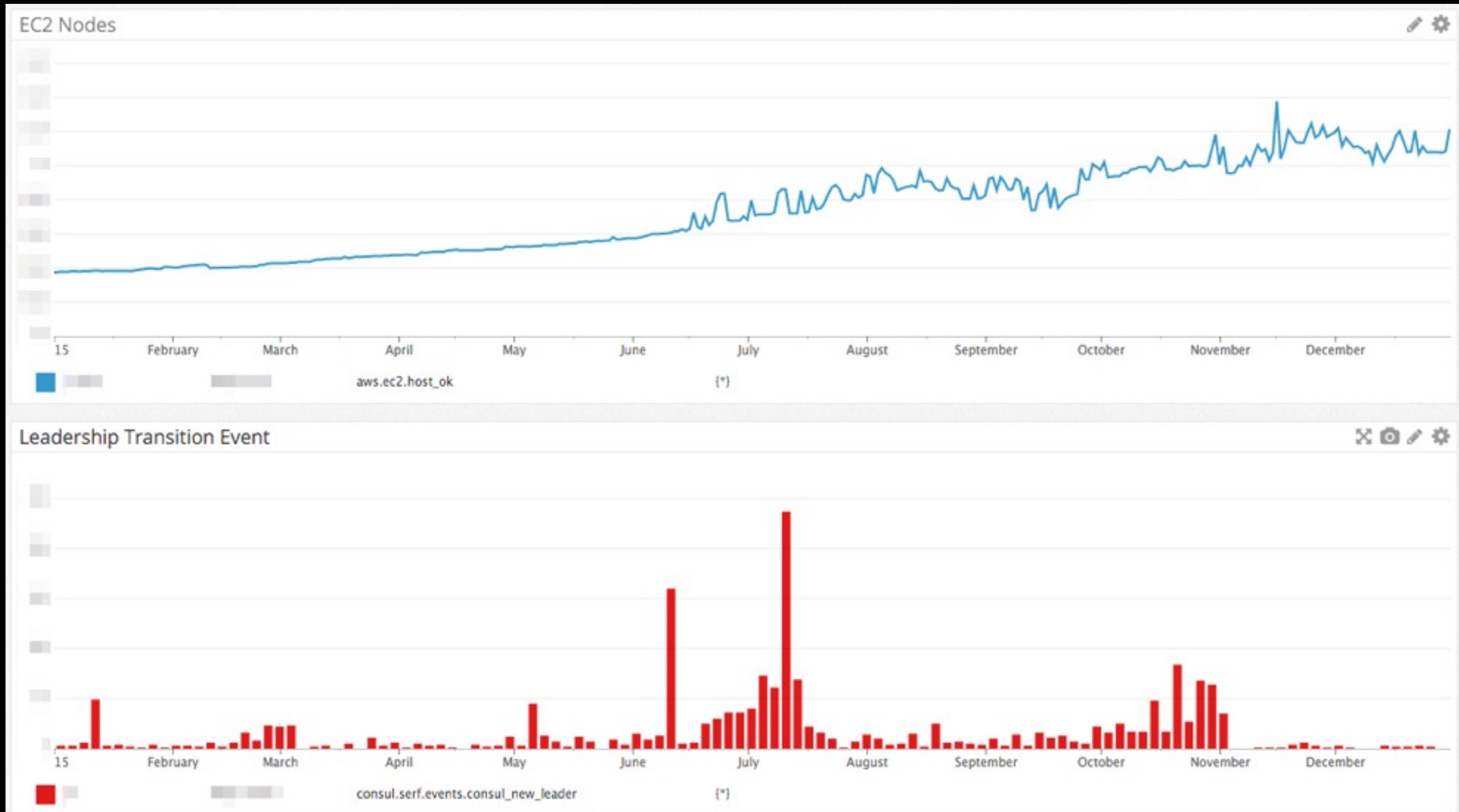
CONSUL LOCK

```
1  description "consul lock example"
2
3  # Defaults set by kernel
4  limit nofile 1024 4096
5
6  emits consullock-up
7
8  start on runlevel [2345]
9  stop on runlevel [!2345]
10
11 env CONSUL_TOKEN=token-goes-here
12
13 exec consul lock -n 1 -token $CONSUL_TOKEN consullock /usr/local/bin/binary goes here
14
15 post-start exec initctl emit consullock-up
16
17 kill signal INT
```

SUPERSIZED OUR SERVERS

WHEN THE LEADERSHIP
TRANSITIONS GREW

AT 1200 NODES - IT'S LIKE WE (ALMOST) TURNED THEM OFF.
C3.2XLARGE DID THE TRICK



ONE FOR 3 MINUTES BECAUSE OF A PACKAGING PROBLEM

2 SMALL OUTAGES

LAST YEAR.

ONE FOR AN HOUR THAT WAS DOCUMENTATION AND
"BROADCAST INPUT TO ALL PANES" RELATED

[HTTP://JOBS.DATADOGHQ.COM/](http://jobs.datadoghq.com/)

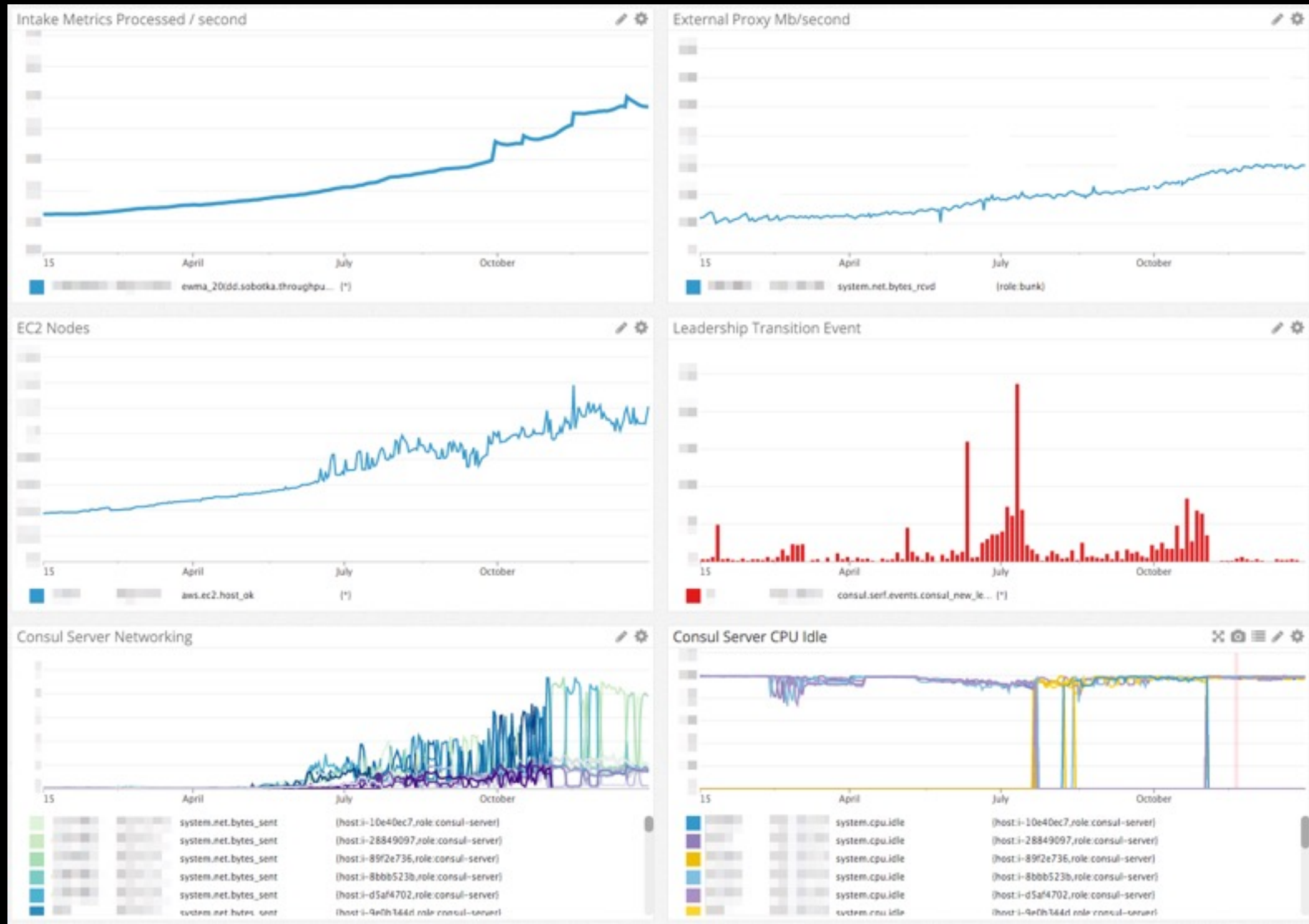
BTW - WE'RE HIRING



DATADOG

WHAT HAVE WE LEARNED?

JANUARY 2016



IT'S YOUR DATACENTER'S BACKBONE

CONSUL IS AWESOME



JUST DO IT

MONITORING IS ESSENTIAL

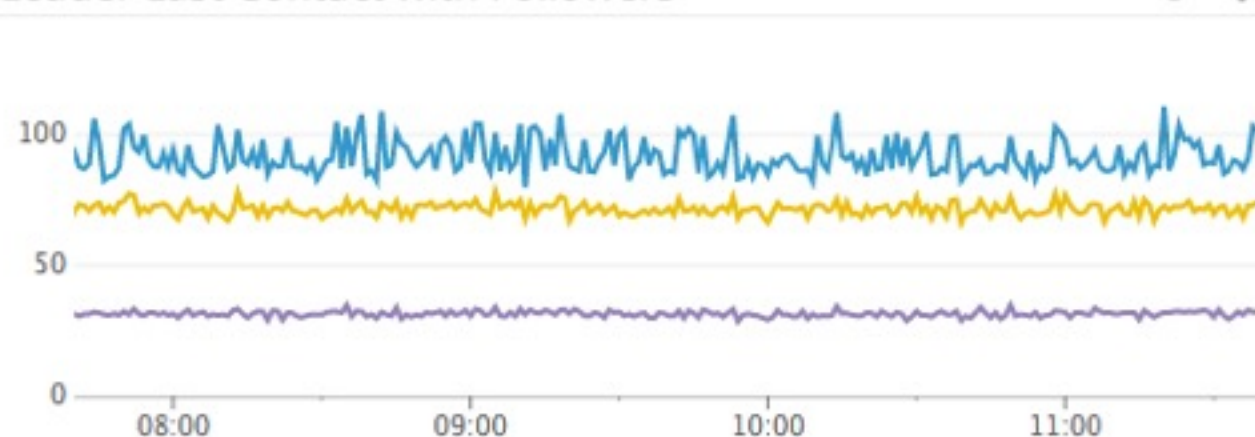
Consul Cluster Join and Failure



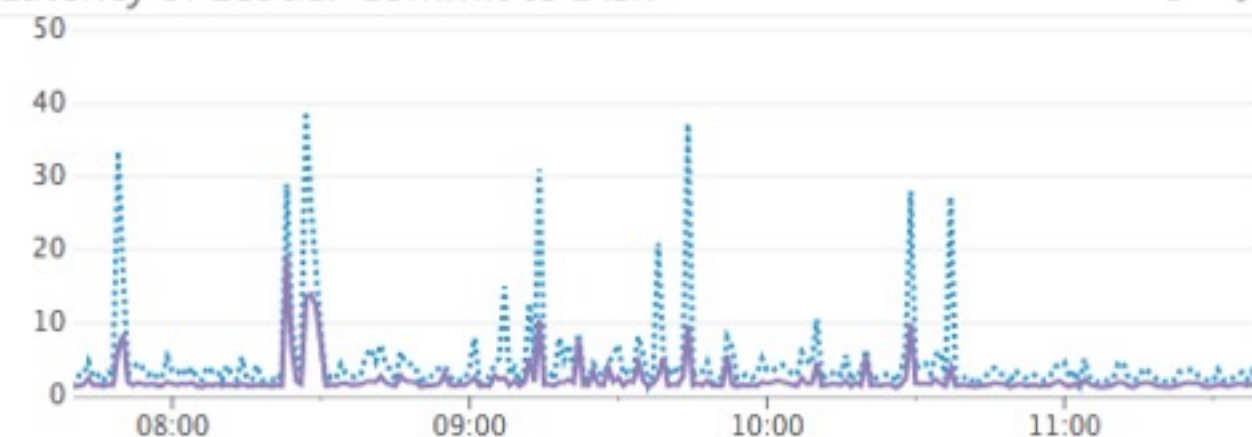
Leader Time to Append Entries



Leader Last Contact with Followers



Latency of Leader Commit to Disk



Consul Raft Commit Time



Consul Raft Apply



TONS OF FIXES AND UPGRADES

UPGRADE TO 0.6.X

Below are the available downloads for the latest version of Consul (0.6.0). Please download the proper package for your operating system and architecture.

You can find the [SHA256 checksums for Consul 0.6.0](#) online and you can [verify the checksums signature file](#) which has been signed using [HashiCorp's GPG key](#). You can also [download older versions of Consul](#) from the releases service.

[Download Consul Web UI](#)



Mac OS X

[32-bit](#) | [64-bit](#)

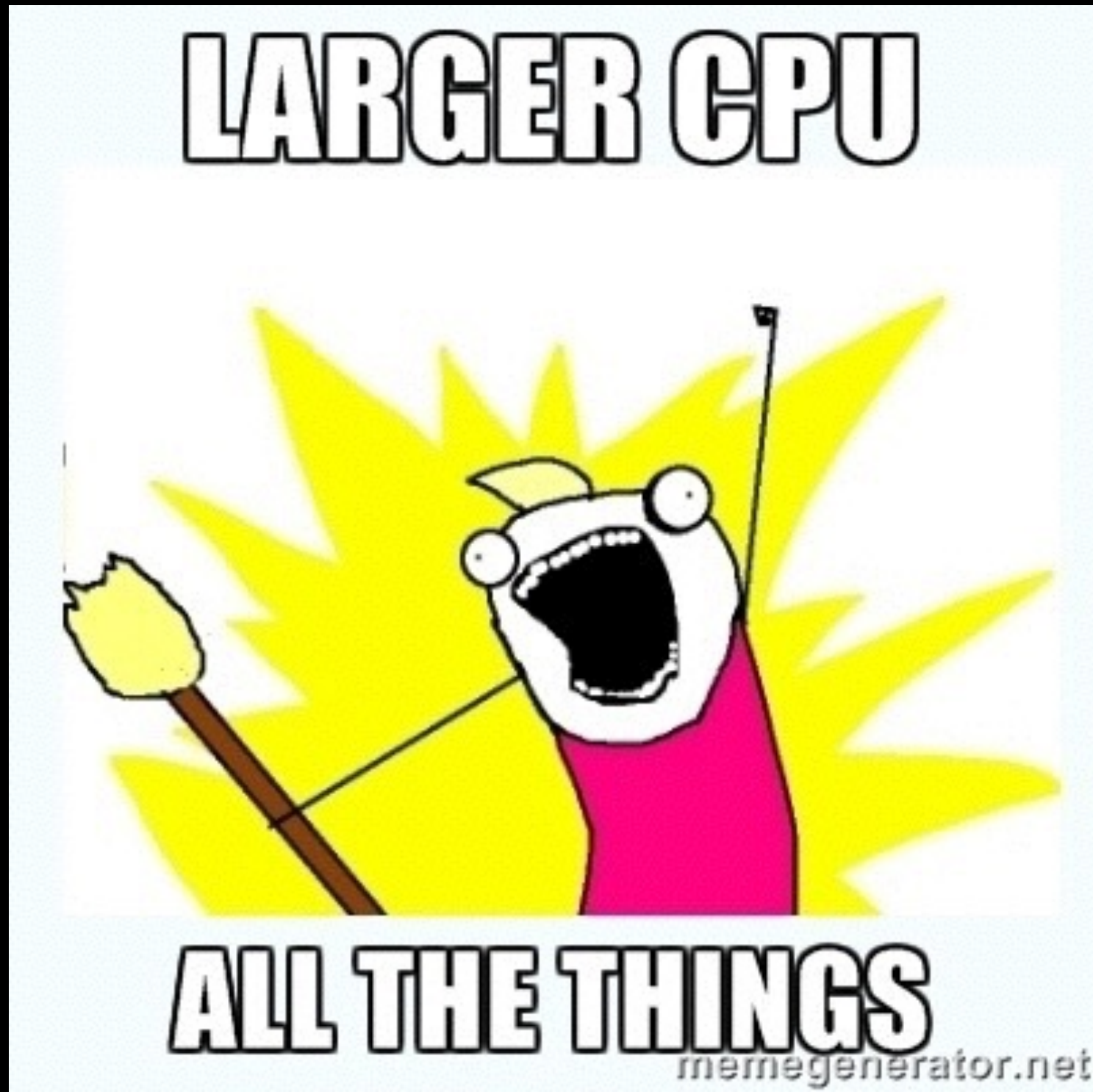
TASTES GREAT, LESS FILLING!

USES LESS MEMORY



FEED IT ALL THE CPUS

CONSUL LOVES CPU



SOME EXAMPLE SIZING

- m3.large ~300 nodes
- c3.xlarge ~500 nodes
- c3.2xlarge ~800 nodes
- As always YMMV.
- 0.6 is more efficient - might be able to be smaller nodes.

BUILD FOR IT - ADD RETRIES - BACKOFF - CIRCUIT BREAKERS

EMBRACE FAILURE

MAKES YOUR WHOLE SYSTEM MORE RESILIENT

DON'T DDOS YOURSELF

WATCH YOUR READ
VELOCITY

RATHER THAN LOTS OF SMALL KEYS

USE FEWER AND
LARGER KEYS

ESPECIALLY IF YOU'RE READING
A LOT OF THEM AT ONCE

FEED IN DATA FROM THE OUTSIDE

LOCK DOWN PARTS OF
THE KV STORE

ACLS ARE YOUR FRIEND

MAKE SURE THEY ONLY FIRE WHEN YOU WANT THEM TO

CONSUL WATCHES
ARE POWERFUL

[HTTPS://GITHUB.COM/DARRON/SIFTER](https://github.com/darron/sifter)

DON'T BUILD CONFIG ON EVERY NODE

IF OUTPUT ISN'T
UNIQUE

USE THE KV STORE TO MOVE THOSE FILES AROUND

THAT'S MY LAST TIP FOR TODAY BUT WE HAVE

ONE MORE THING

USE THE KV STORE TO TRANSPORT
CONFIGURATION FILES

KVEXPRESS

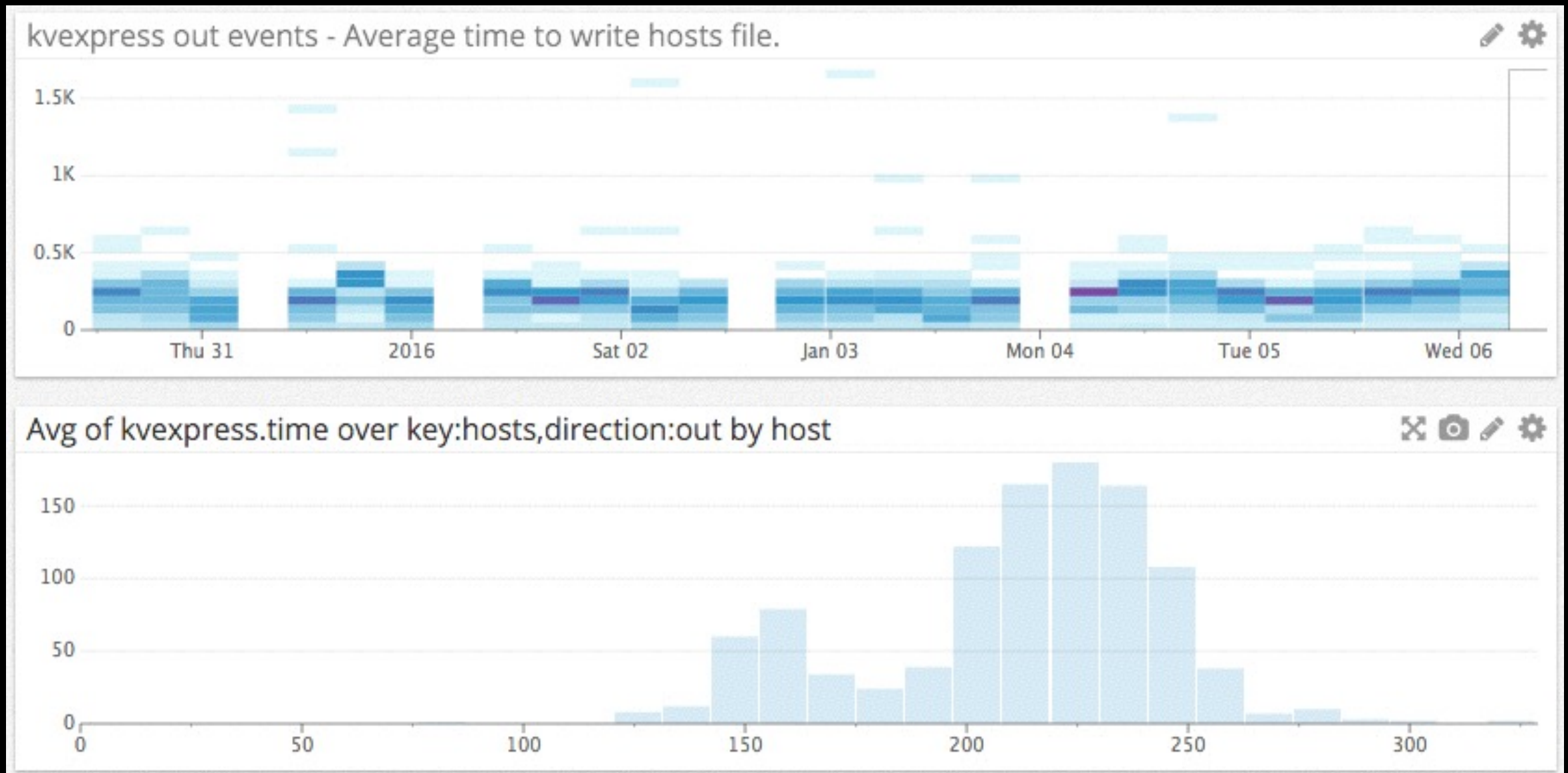
IN & OUT
BOTH DIRECTIONS

MAIN FEATURES

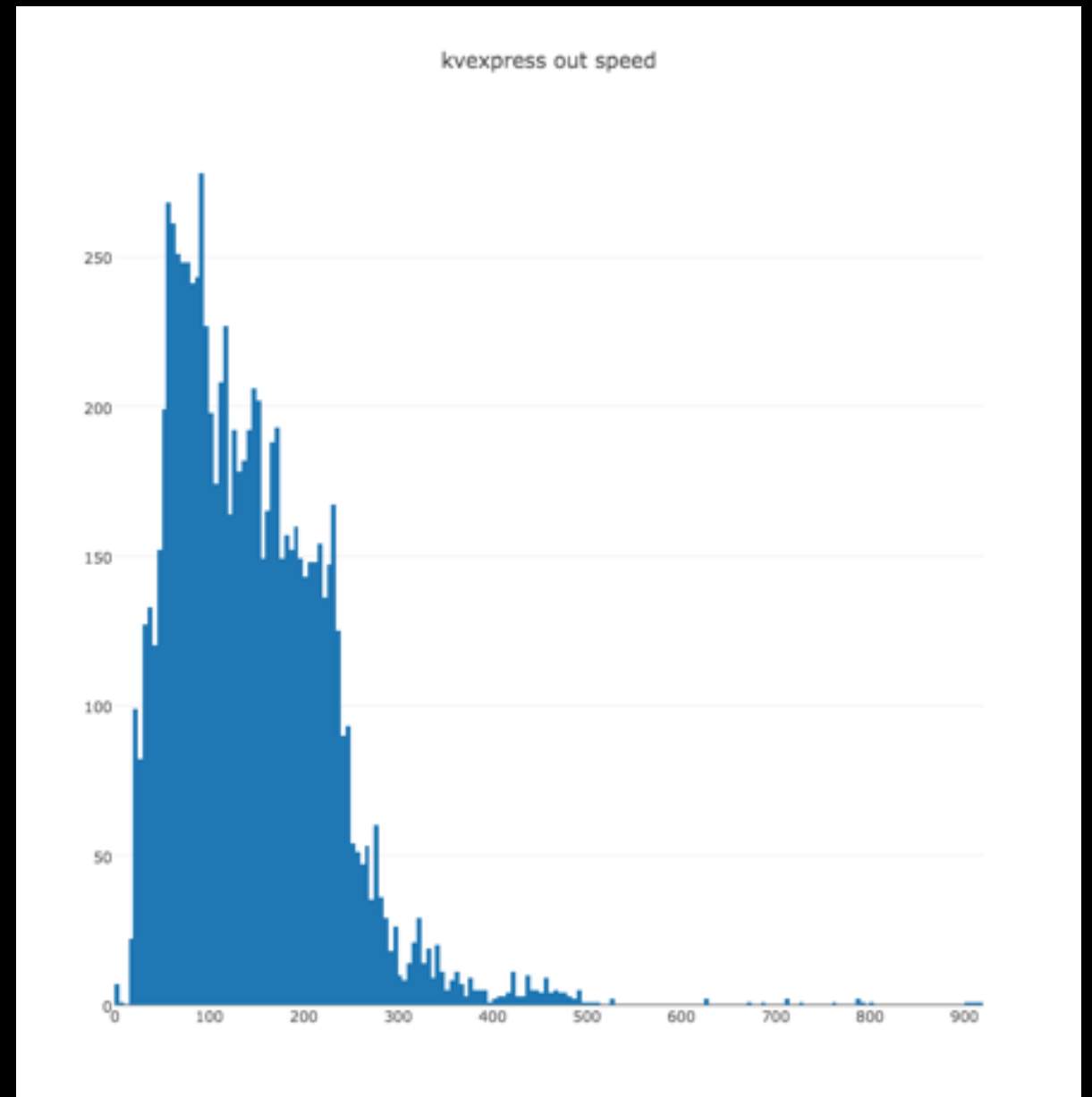
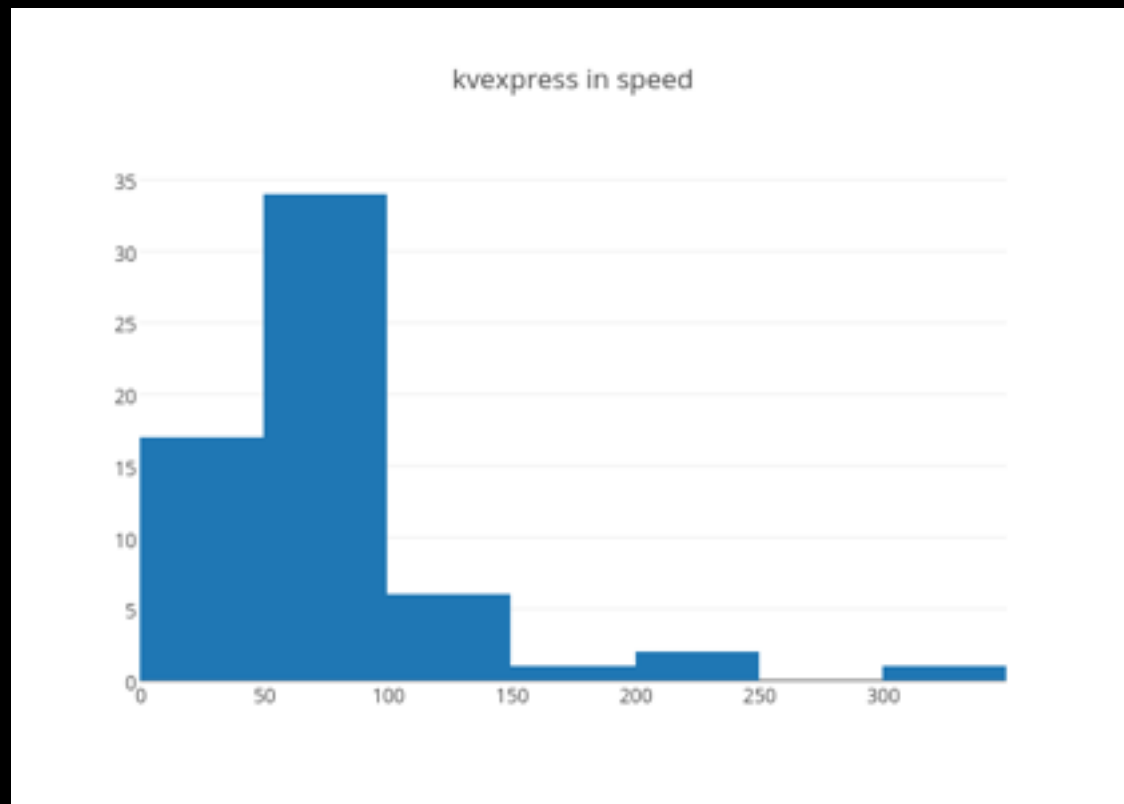
- 10MB Go binary
- Uploads and downloads files under 512KB
- Emits Dogstatsd metrics and Datadog Events
- Files sent == files delivered
- Doesn't re-upload or re-deliver
- Very safe
- Runs commands after delivery

< 500MS TO DELIVER A FILE TO 1000 NODES

IT'S SUPER FAST



SPEED IN AND OUT



WHEN FILES UPDATE

CAN POST DIFFS

Amazon EC2

Amazon RDS

Amazon SNS

Amazon Web Services

Campfire

Capistrano

ChatWork

Chef

Consul

Datadog

Docker

Elasticsearch

Feed

FlowDock

Github

HAProxy

Jenkins

Monitor Alert

My Apps

MySQL

New Relic

PagerDuty

Pingdom

RabbitMQ

Sentry

Slack

10 matching events from Jan 8, 8:08AM - Jan 8, 12:08PM



Leave a status update...

Post



Updated: kvexpress/hosts/data

#direction:in

#host:i-d5af4702

#key:kvexpress/hosts/data

...

10.111.199.50 3.lamar.service.consul
10.111.199.50 4.lamar.service.consul
10.111.199.50 lamar.service.consul
-10.111.219.132 az-us-east-1b.kafka-cold.service.consul
-10.111.219.132 kafka-cold.service.consul
10.113.176.243 az-us-east-1b.ctx-pshard1.service.consul
10.113.176.243 ctx-pshard-standby.ctx-pshard1.service.consul
10.113.176.243 ctx-pshard.ctx-pshard1.service.consul
22 mins ago · [Add comment](#) · [Lower priority](#)



Updated: kvexpress/hosts/data

#direction:in

#host:i-d5af4702

#key:kvexpress/hosts/data

...

10.97.128.62 replica.master-db.service.consul
10.97.132.140 az-us-east-1a.cassandra.service.consul
10.97.132.140 cassandra.service.consul
+10.97.134.75 az-us-east-1a.master-db.service.consul
+10.97.134.75 dogweb-backend-standby.master-db.service.consul
+10.97.134.75 dogweb-backend.master-db.service.consul
+10.97.134.75 master-db.service.consul
+10.97.134.75 postgres.master-db.service.consul
+10.97.134.75 standby.master-db.service.consul
10.97.152.234 18.spidly.service.consul
10.97.152.234 19.spidly.service.consul
10.97.152.234 spidly.service.consul
32 mins ago · [Add comment](#) · [Lower priority](#)



Updated: kvexpress/hosts/data

#direction:in

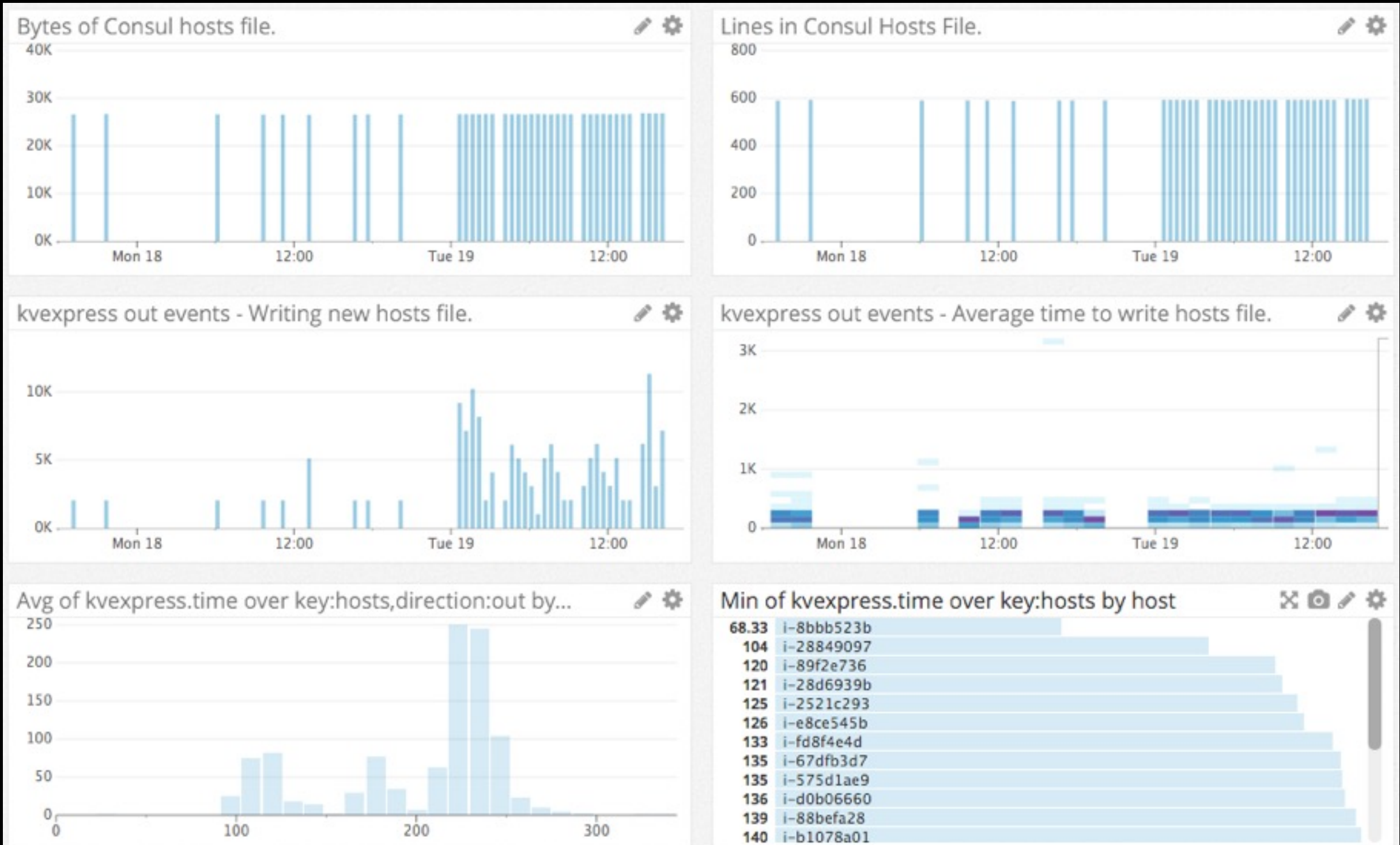
#host:i-d5af4702

#key:kvexpress/hosts/data

...

MEASURE ALL THE THINGS

METRICS FOR ALL



VERY QUICK DEMO

[HTTPS://GITHUB.COM/
DATADOG/KVEXPRESS](https://github.com/datadog/kvexpress)

QUESTIONS?

SERVICE DISCOVERY IN THE CLOUD

CONSUL @ SCALE14x

THANKS!

DARRON@FROESE.ORG

@DARRON

[GITHUB.COM/DARRON](https://github.com/darron)