



Deft Data at Netflix: Using Amazon S3 and Amazon Elastic MapReduce for Monitoring at Gigascale

Roy Rapoport

November 14, 2013



© 2013 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified, or distributed in whole or in part without the express consent of Amazon.com, Inc.

A Word About Me ...

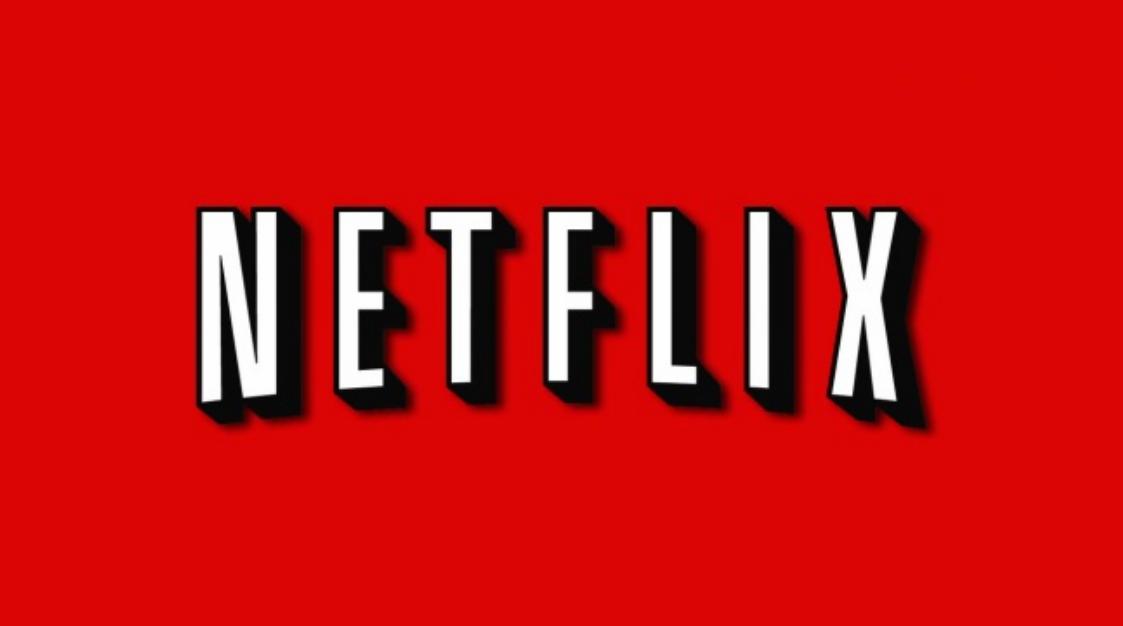
- About 20 years in technology
- Systems engineering, networking, software development, QA, release management
- Time at Netflix: 1599 days(4y:4m:15d)
- Before at Netflix: Service Delivery in the IT/Ops, troubleshooter, Builder of Python Things[tm]
- Current role: Cloud Monitoring
 - We build platforms
 - Sometimes we make them easy to use



A Word About Netflix ...

Just the Stats

- 16 years
- 2000+ employees
- 40 million users
- 5×10^9 hours/quarter

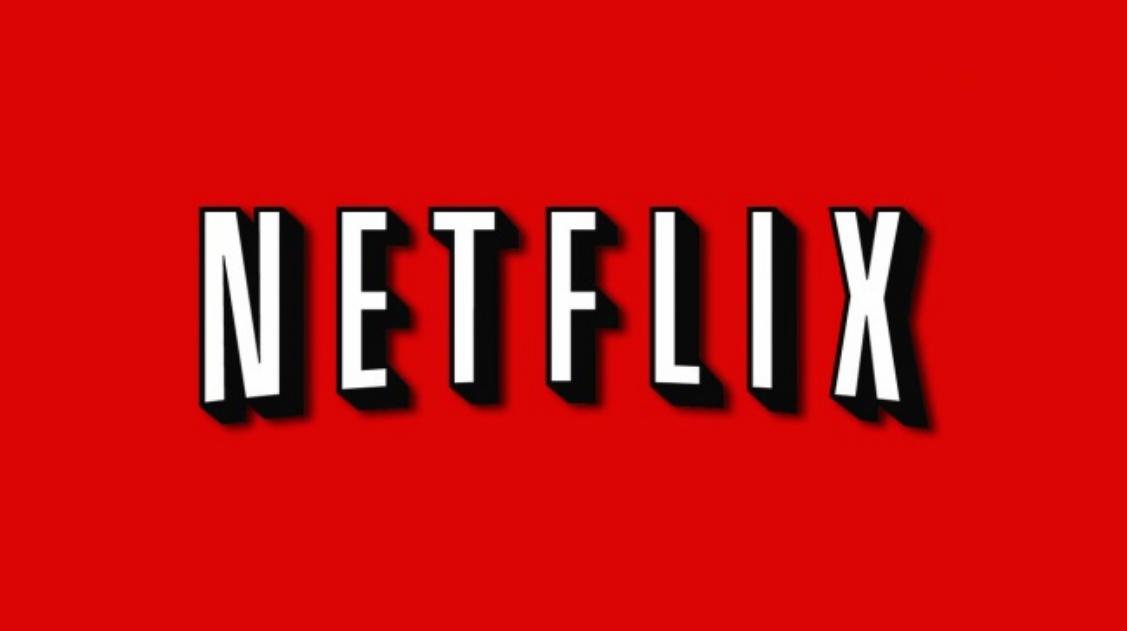


NETFLIX

A Word About Netflix ...

Freedom and Responsibility Culture

- Optimize speed of innovation
- Constrain availability
- Cost will be what cost will be
- Hire smart (experienced) people
- Get out of their way
- Anti-process bias

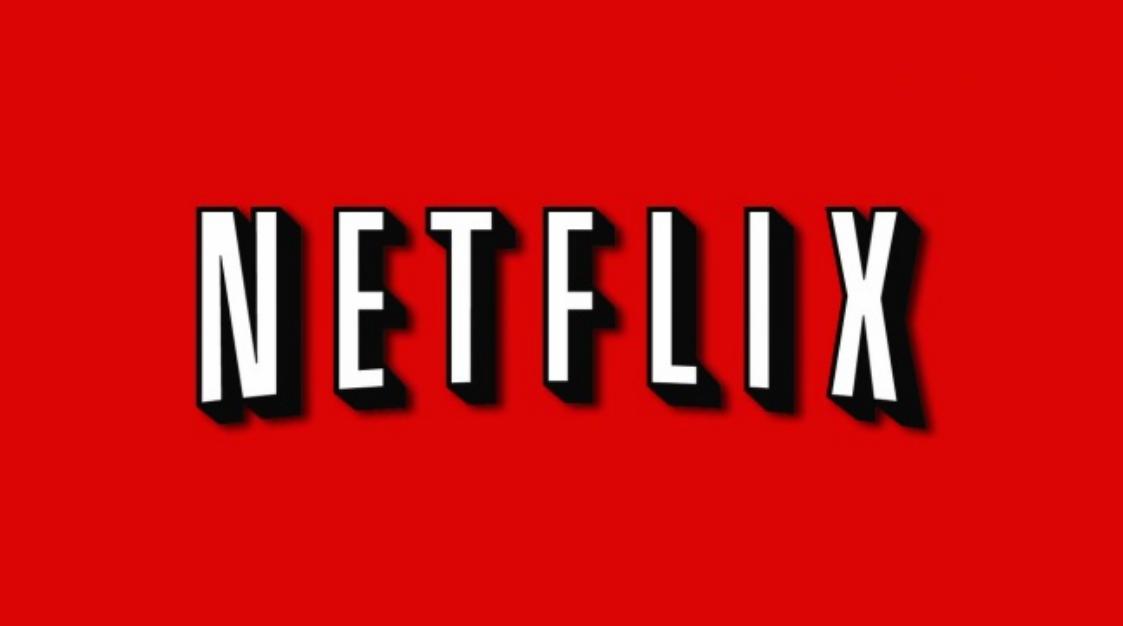


NETFLIX

A Word About Netflix ...

Technology and Operations

- Service Oriented Architecture
- Decentralized Operations. You
 - Build
 - Test
 - Deploy
 - Set up alerting and monitoring
 - Wake up at 2AM

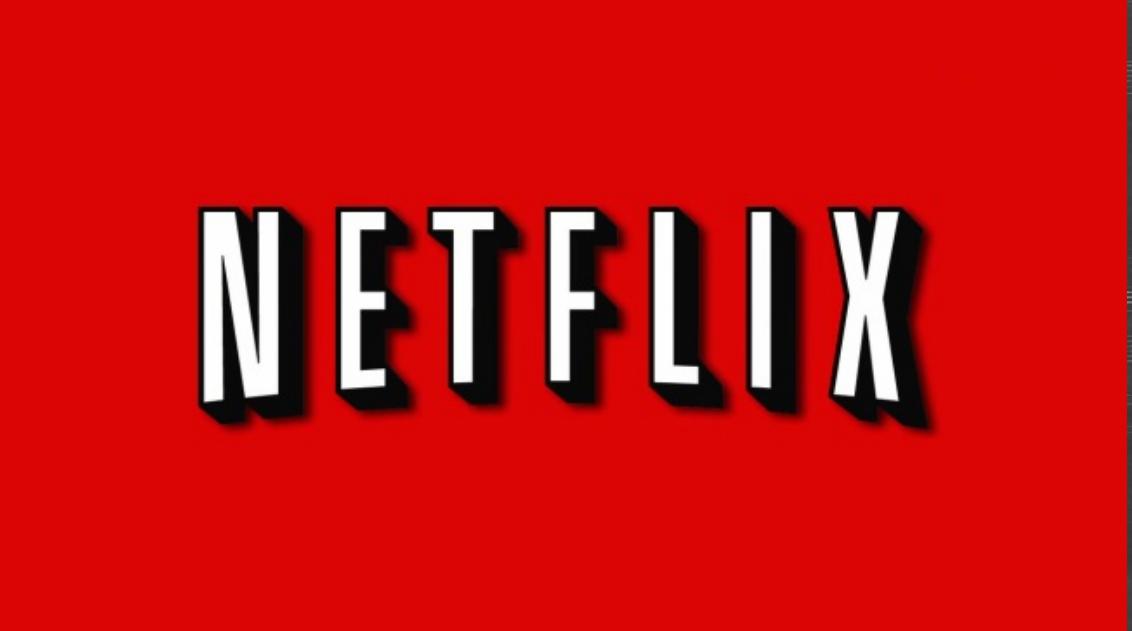


NETFLIX

A Word About Netflix ...

Technology and Operations

- AWS-based for 100% of streaming*
- Huge expansion
 - Customer Growth
 - New markets
 - Metrics

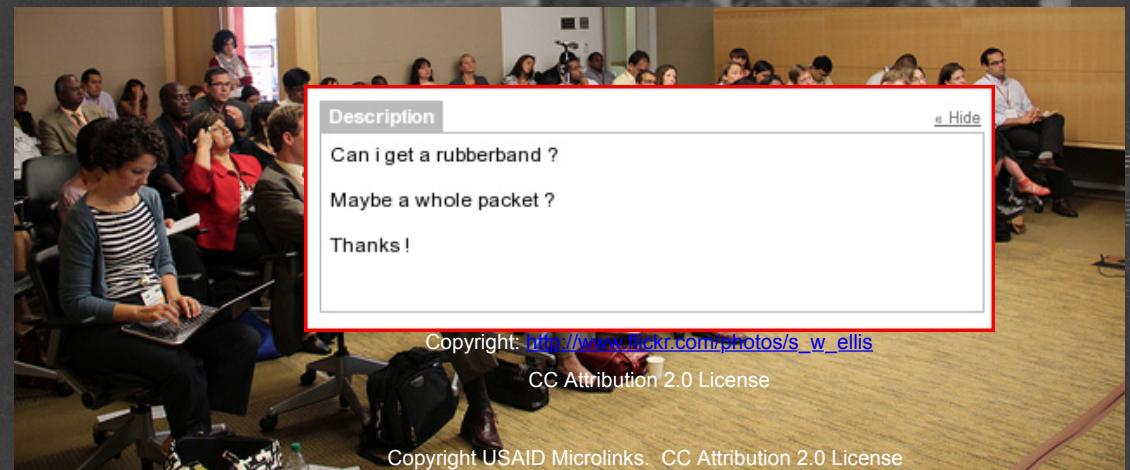


NETFLIX

In the Old Days ...

Our Old Alerting System

- Enterprise IT Solution
- Managed by the Enterprise IT Alerting People
- File Tickets
- Send alerts to NOC
- Completely separate from telemetry system



Copyright: State Library of Victoria Collections. CC Attribution 2.0 License

In the Old Days ...

In the Old Days

Our Old Telemetry System

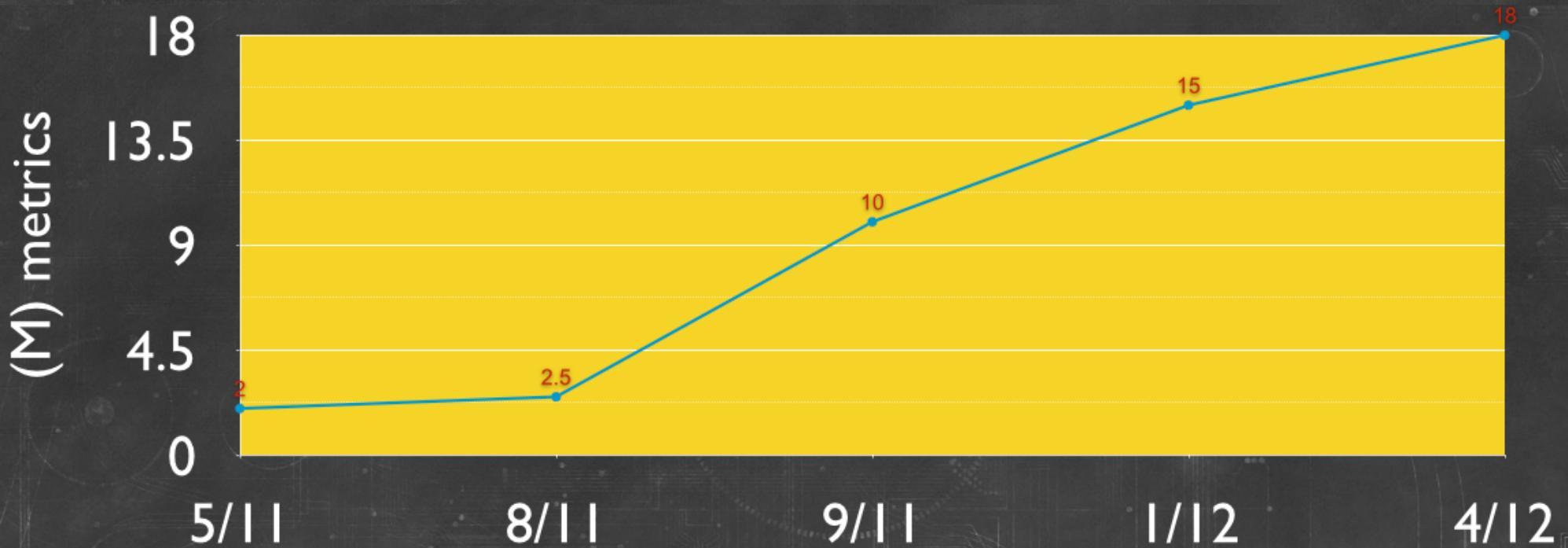
- Spare-time effort by a lone sysadmin
- Loved by developers
- Custom TCP protocol
- RRD file back-end storage
- Mostly Perl
- Datacenter-bound (and limited)
- Starting to falter under metrics growth



Copyright: <http://www.flickr.com/photos/acme>
CC Attribution 2.0 License

Copyright: State Library of Victoria Collections. CC Attribution 2.0 License

Speaking of Growth



Speaking of Growth

By Way of Comparison

- Every person in the world
 - twice
- Every smartphone in the world
 - ten times



So We Built Something Better

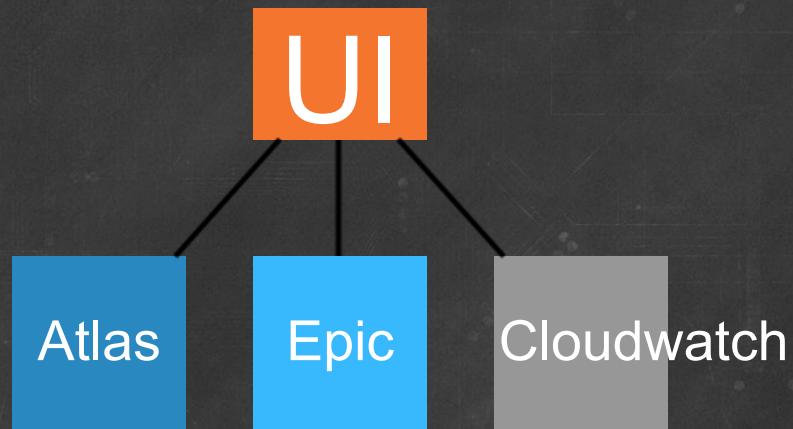


Copyright: <http://www.flickr.com/photos/76651030@N02/>

CC Attribution 2.0 License

So We Built Something Better

UI Layer Fronts Multiple Systems



Copyright: <http://www.flickr.com/photos/76651030@N02/>

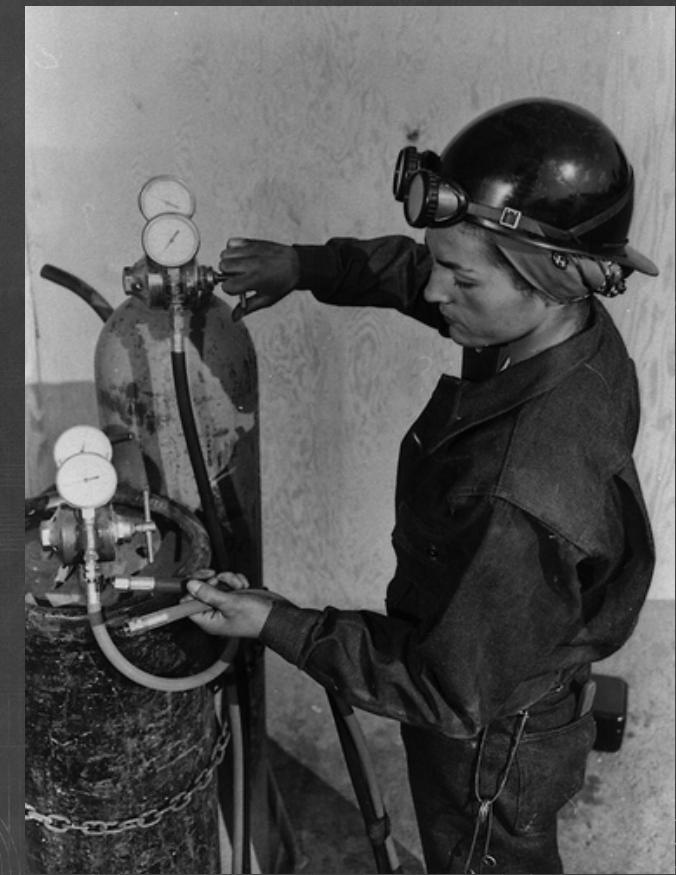
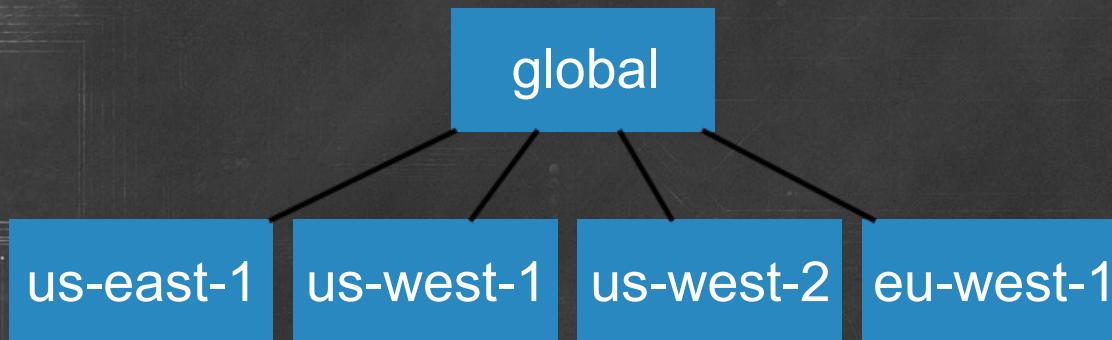
CC Attribution 2.0 License

So We Built Something Better



Clear Regional Separation

- And aggregation



Copyright: <http://www.flickr.com/photos/76651030@N02/>
CC Attribution 2.0 License

So We Built Something Better

Localized Node/Metric Identification

Before:

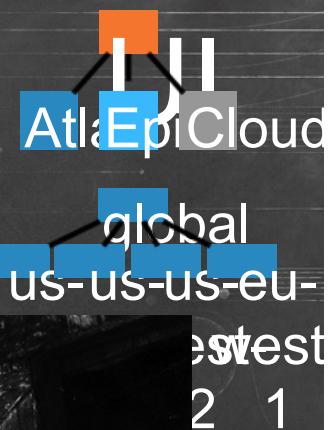
Here's a metric!

I think You're Bob

Now:

I'm Bob.
Here's a metric!

OK!



Copyright: <http://www.flickr.com/photos/76651030@N02/>

CC Attribution 2.0 License

So We Built Something Better

What's a Metric?

- com.netflix.eds.nccp.successful.requests.uiversion.nccprt-authorization.devtypid-101.clver-PHL_0AB.uiver-UI_169_mid.geo-US
- **256 characters aren't enough!**
- **This is better:**

nf.ami	ami-aa5166ef	nf.node	i-097c0e52	uiversion	UI_169_mid
nf.app	wp	nf.region	us-west-1	action	authorization
nf.cluster	wp-batch	nf.zone	us-west-1b	devtype	101
nf.asg	wp-batch-v163	class	nccp	clvers	PHL_0AB
nf.country	us	type	request	geo	us

Atlanta Cloud
global
us-us-us-eu-
east-west
Here's test
I think
1 1 You're 1
metric Bob
Book!
Here's
a
metric!

So We Built Something Better

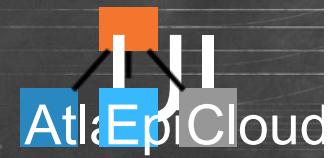
Powerful queries

- Make the complex possible
- Make the simple ... sort of hard



Copyright: Kurt Moerman

CC Attribution 2.0 License



global
us-us-us-eu-
east-west
Here's test
I think
1 1 You're 1
met Bob
Book
Please
a
metric!

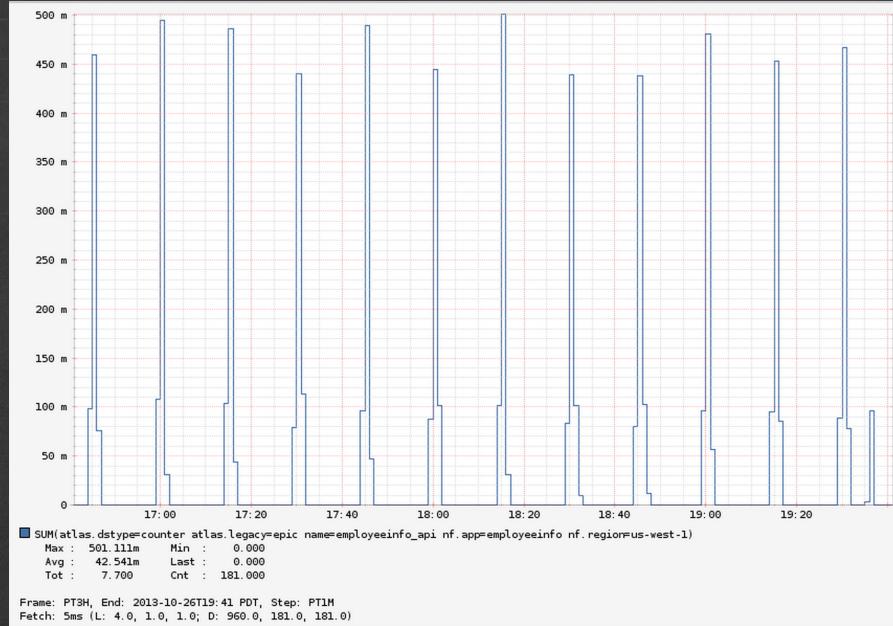
So We Built Something Better

Powerful queries

- Make the complex possible
- Make the simple ... sort of hard

<http://atlas/api/v1/graph?>

```
q=nf.region,us-west-1,:eq,nf.app,employeeinfo,:eq,:and,name,employeeinfo_api,:eq,:and,:sum
&e=now-5m&s=e-3h
```



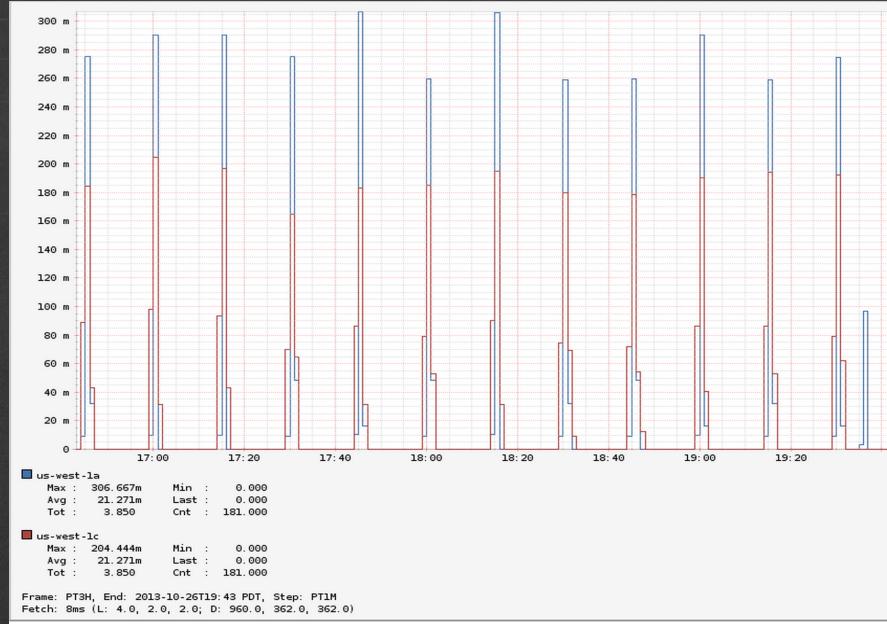
So We Built Something Better

Powerful queries

- Make the complex possible
- Make the simple ... sort of hard

<http://atlas/api/v1/graph?>

q=nf.region,us-west-1,:eq,nf.app,employeeinfo,:eq,:and,name,employeeinfo_api,:eq,:and,:sum,(,nf.zone,),:by &e=now-5m&s=e-3h

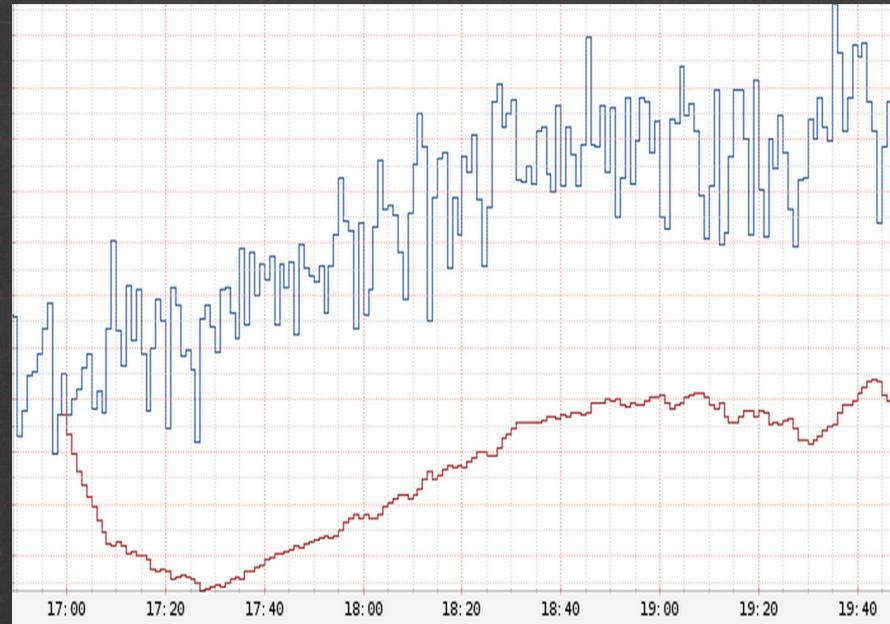


So We Built Something Better

Powerful queries

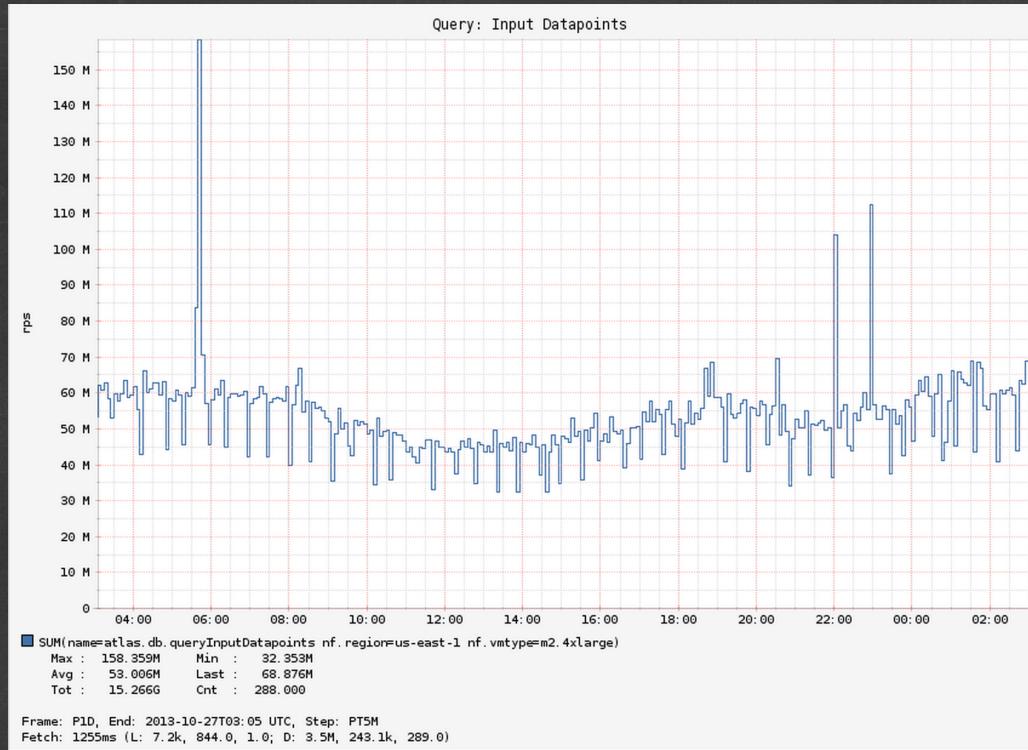
- Make the complex possible
- Make the simple ... sort of hard

```
http://atlas/api/v1/graph?q=sps.nf.cluster,\(,nccp-legacy,nccp-modern,\):in,nccprt,\(,NCCPLicense,com.netflix.streaming\_nccp\_request\_license,\):in,:and,stat,SuccessfulRequests,:eq,:and,device.rollup,3ds,:eq,:and,:sum,:set,entering\_trough,sps,:get,1h,:offset,0.95,:mul,sps,:get,:gt,:set,smoothed,sps,:get,10,0.1,0.02,:count,2,:ge,:vspan,60,:alpha,\$\(device.rollup\),:legend
```



Ridiculous Read Volume:

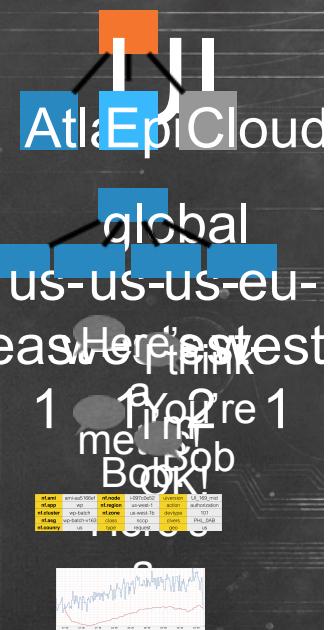
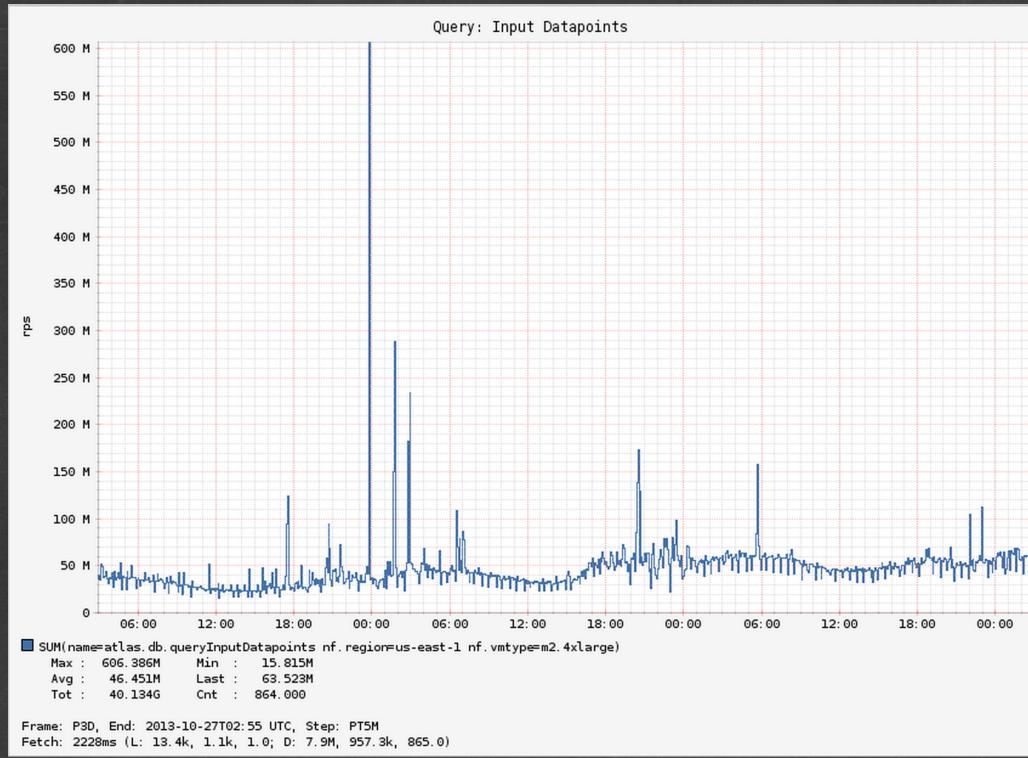
- Engage
 - Graphs and Dashboards
 - Alerting
 - Automated Canaries
 - Capacity Analytics
 - Special Projects
 - BI



So We Built Something Better

Ridiculous Read Volume:

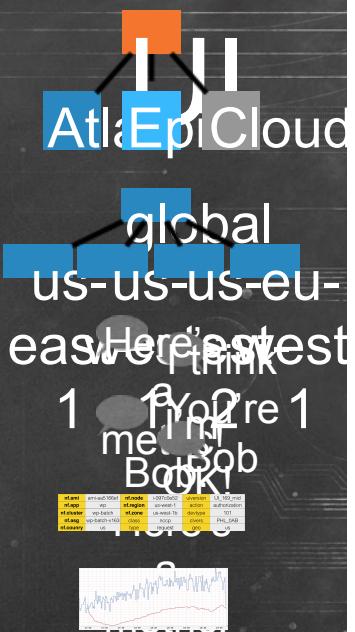
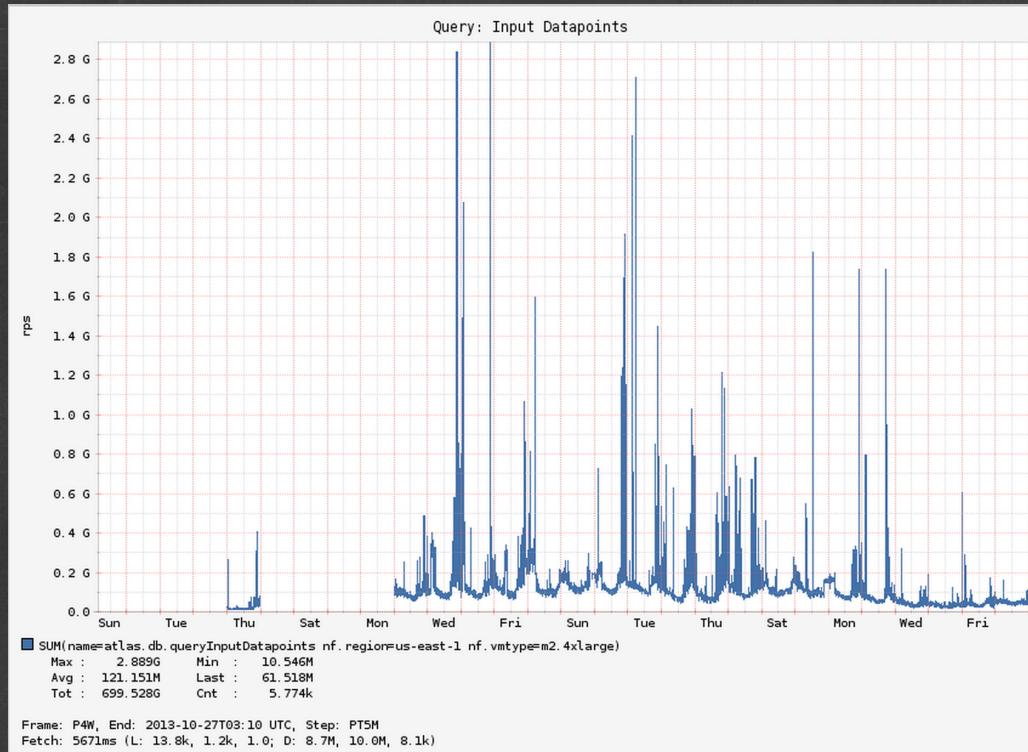
- Engage
- Graphs and Dashboards
- Alerting
- Automated Canaries
- Capacity Analytics
- Special Projects
- BI



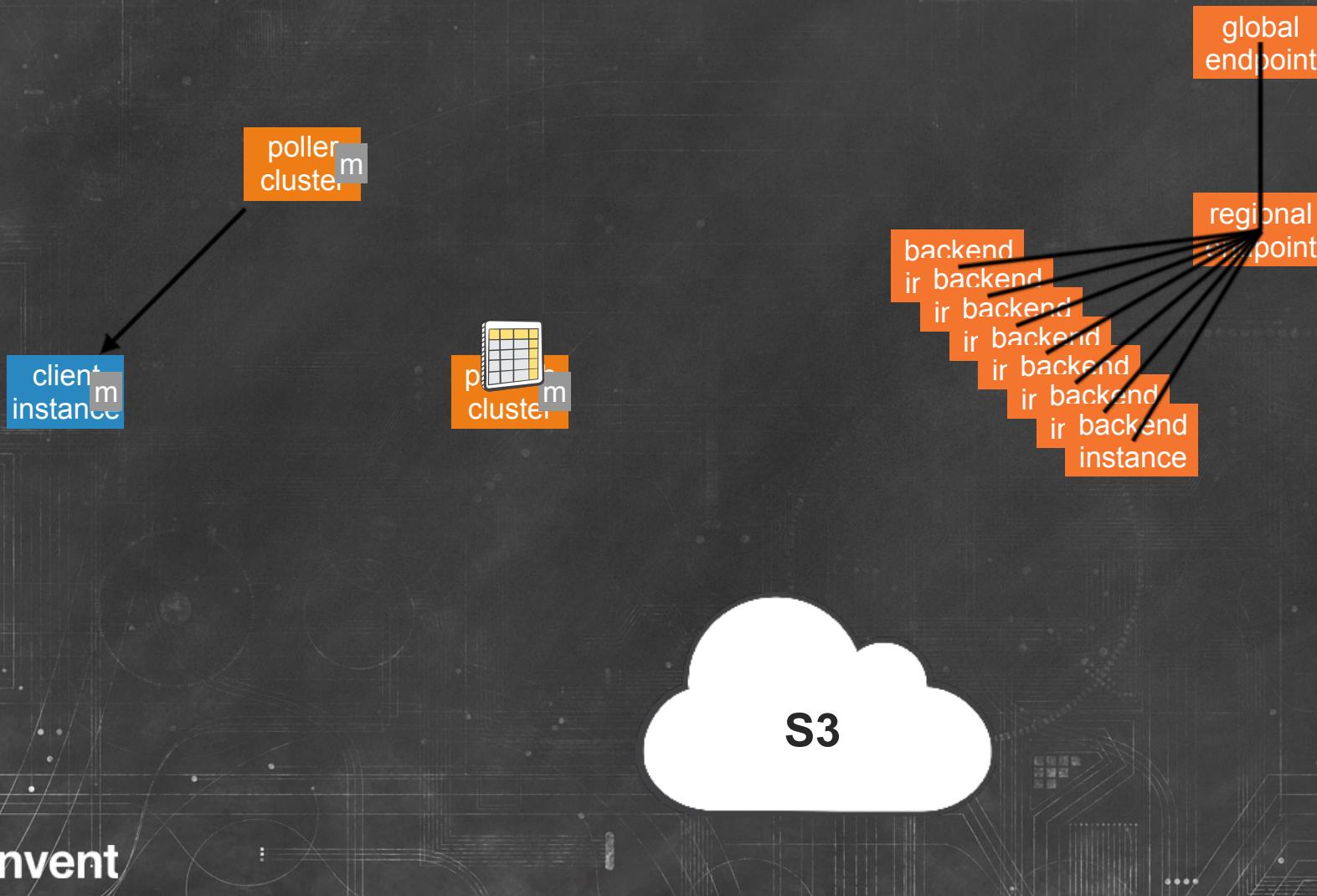
So We Built Something Better

Ridiculous Read Volume:

- Engage
- Graphs and Dashboards
- Alerting
- Automated Canaries
- Capacity Analytics
- Special Projects
- BI



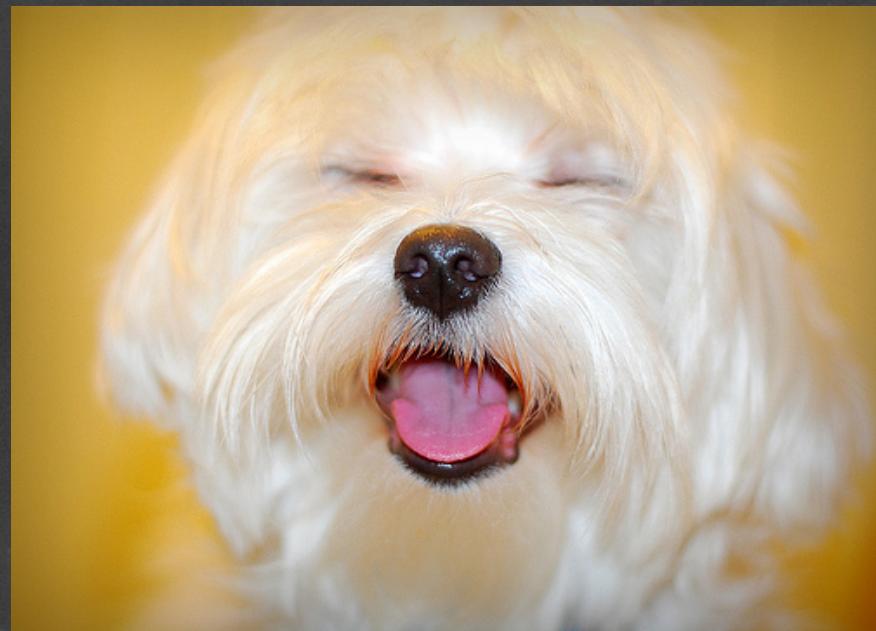
So We Built Something Better



That Sounds Great!

Surely there are no problems

- Speed is hard
- Speed at volume is harder
- We looked at spinning disks
- Memory's the way to go
- m2.4
- This is operational data
 - People want it available, fast
 - Operations have short memories
 - Some days are special



Copyright: <http://www.flickr.com/photos/lainetrees/>

CC Attribution 2.0 License

That Sounds Great!

Surely there are no problems

- Speed is hard
- Speed at volume is harder
- We looked at spinning disks
- Memory's the way to go
- m2.4
- This is operational data
 - People want it available, fast
 - Operations have short memories
 - Some days are special

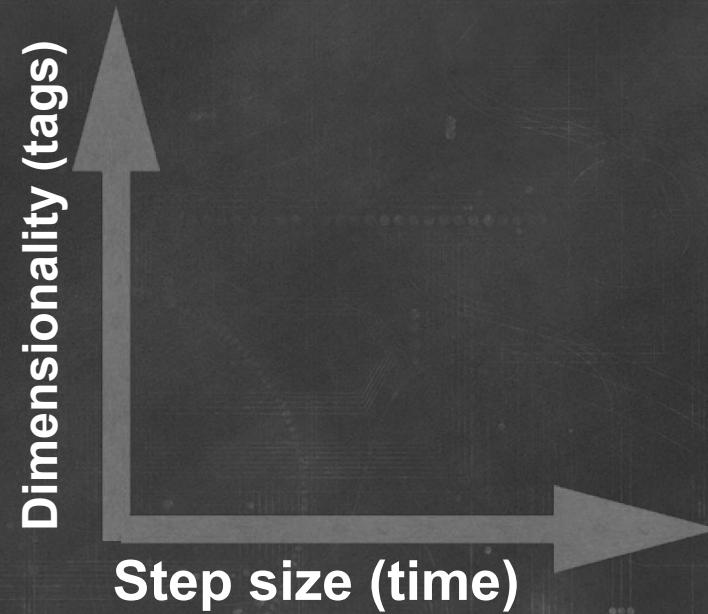


Copyright: <http://www.flickr.com/photos/amenk/>

CC Attribution 2.0 License

That Doesn't Sound Great!

- If only we could reduce it ...
- “Reduce”? Get it? Get it?
- Our granularity is two-dimensional
- We can reduce on either dimension
- Some tags make sense for very rapid reduction
 - Hystrix
 - nf.node
 - Sometimes a lot (vhs)
 - Sometimes a little (Cassandra)



A Reductive Approach

- For a series of values, reduce and keep:
 - minimum
 - maximum
 - total
 - count
- Example:
 - 3,5,9,14,20: min 3, max 20, tot 51, count 5
- Allows for sense of scale
- Allows for arbitrary further reduction w/o loss of precision

Reduction: Policy

- Policy-driven EMR engine
- Four possible actions
 - preserve
 - drop
 - consolidate
 - rollup

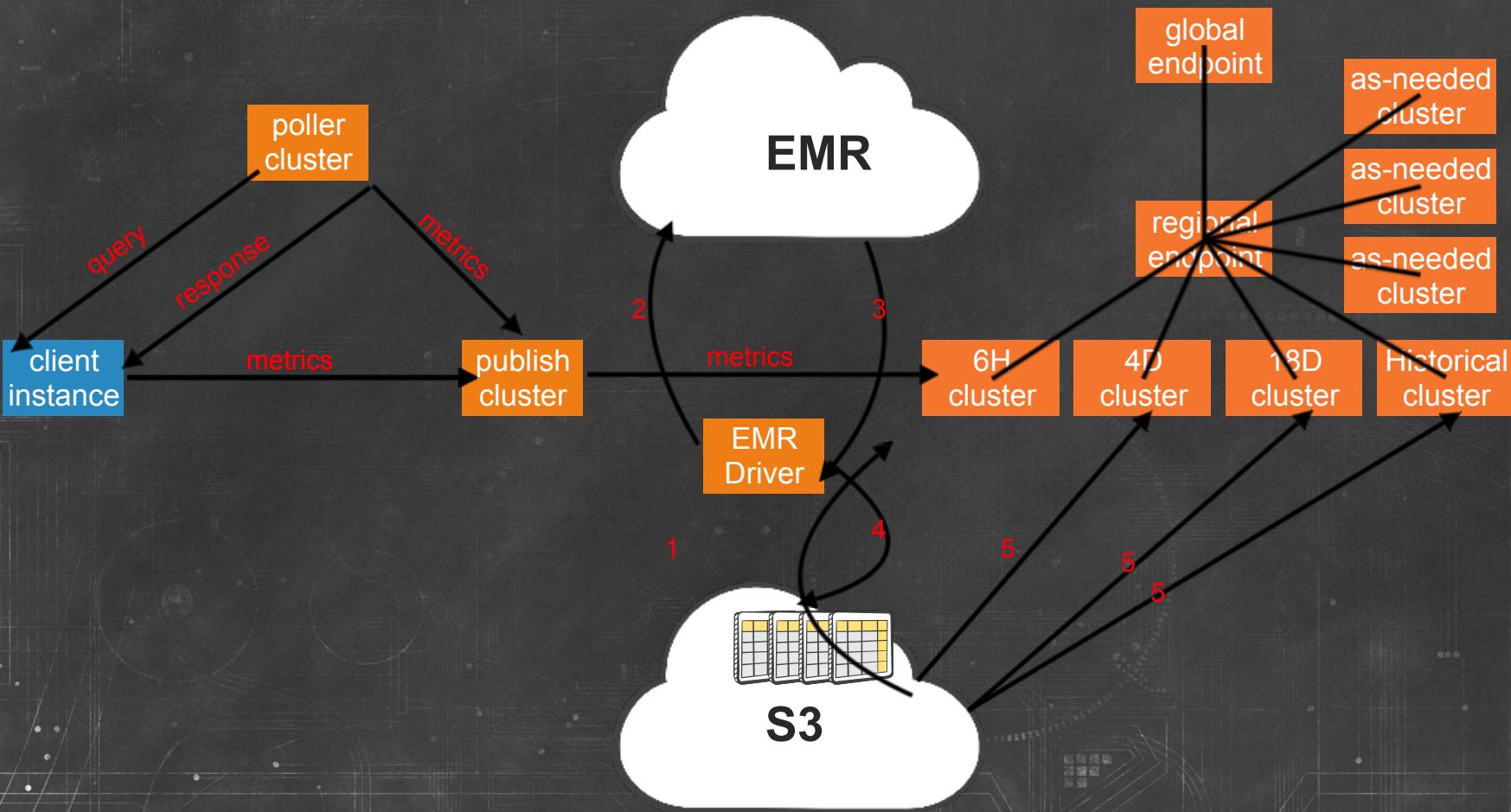


Copyright: <http://www.flickr.com/photos/bagaball/>

CC Attribution 2.0 License

Reduction: Policy

```
{  
  "rules" : [  
    { "operations" : [{"op" : "drop"}],  
      "query" : "nf.app,api,:eq,class,  
      (,LastMinuteFailRatio,SLA,NetflixSimpleDBService,),:in,:and"  
    },  
    { "operations" : [{  
        "config" : { "keys" : [ "nf.node", "device", "nf.country" ]  
      },  
        "op" : "rollup" }],  
      "query" : ":true"  
    }  
  ]  
}
```



Reduction: Benefits

- Indefinite storage in S3
- Fear of commitment achievement: **Unlocked**
- Can be aggressive about hiding metrics
- High granularity for special days
- Automated for regular operations*
- Not in critical path for visibility SLA
- Firewalls accidental metric explosions
- Huge efficiency gains



Copyright: http://www.flickr.com/photos/dr_pete/

CC Attribution 2.0 License

Reduction: Efficiency

With Bottom-Mounted Freezer
Without Through-The-Door Ice Service

Model LSC25990B Capacity: 25 Cubic Feet

Compare the Energy Use of this Refrigerator with Others Before You Buy.

This Model Uses
499kWh/year

Energy use (kWh/year) range of all similar models

Uses Least Energy
499

Uses Most Energy
499

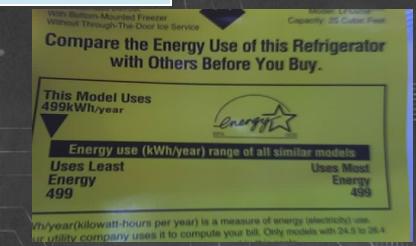
kWh/year(kilowatt-hours per year) is a measure of energy (electricity) use. Your utility company uses it to compute your bill. Only models with 24.5 to 28.4



Copyright: <http://www.flickr.com/photos/sebrenner/>
CC Attribution 2.0 License

Reduction: Efficiency

	6H	4D	18D	HISTORY
Time Horizon	6 Hours	4 Days	18 Days	3 Months
Size	600	512	180	12
Instances Per Hour	100	5	0	0
% Reduction	0	95	100	100



**THIS GREEN BAND HAS BEEN APPROVED FOR
APPROPRIATE AUDIENCES
BY THE AMERICAN GREEN BAND ASSOCIATION OF AMERICA**

www.bandratings.com

www.gbaoa.org

Previews

- Self-service for special requests
- Different instance types
 - cr1.8xlarge
 - hi1.4xlarge
- Multi-tiered metric visibility

THIS **GREEN BAND** HAS BEEN APPROVED FOR
APPROPRIATE AUDIENCES

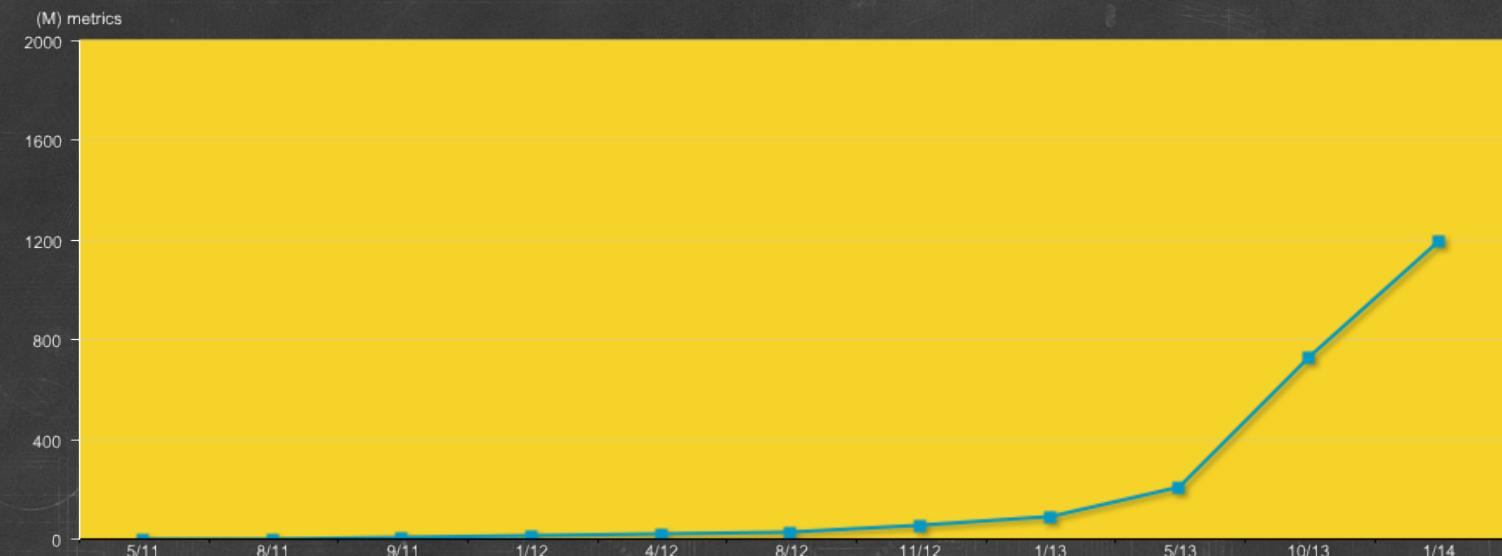
BY THE AMERICAN GREEN BAND ASSOCIATION OF AMERICA

www.bandratings.com

www.gbaoa.org

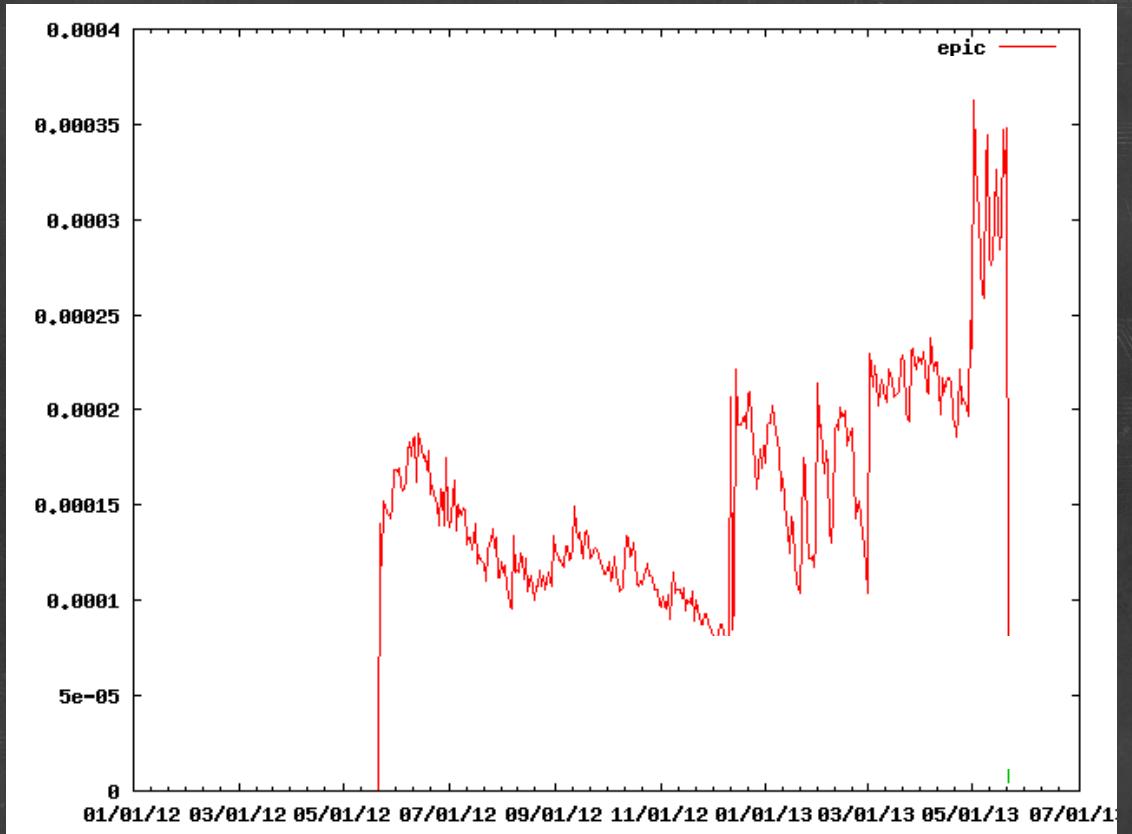
Copyright: <http://www.flickr.com/photos/creativealan/>
CC Attribution 2.0 License

Growth Redux



And a Last Word About Costs

- Priorities Reminder
 - Speed of Innovation
 - Availability
 - Cost
- Never intended to lower costs
 - Cloud migration
 - Additional features
 - Massive Performance



EMR?

FTW



Please give us your feedback on this presentation

BDT-302

As a thank you, we will select prize winners daily for completed surveys!

Thank You

