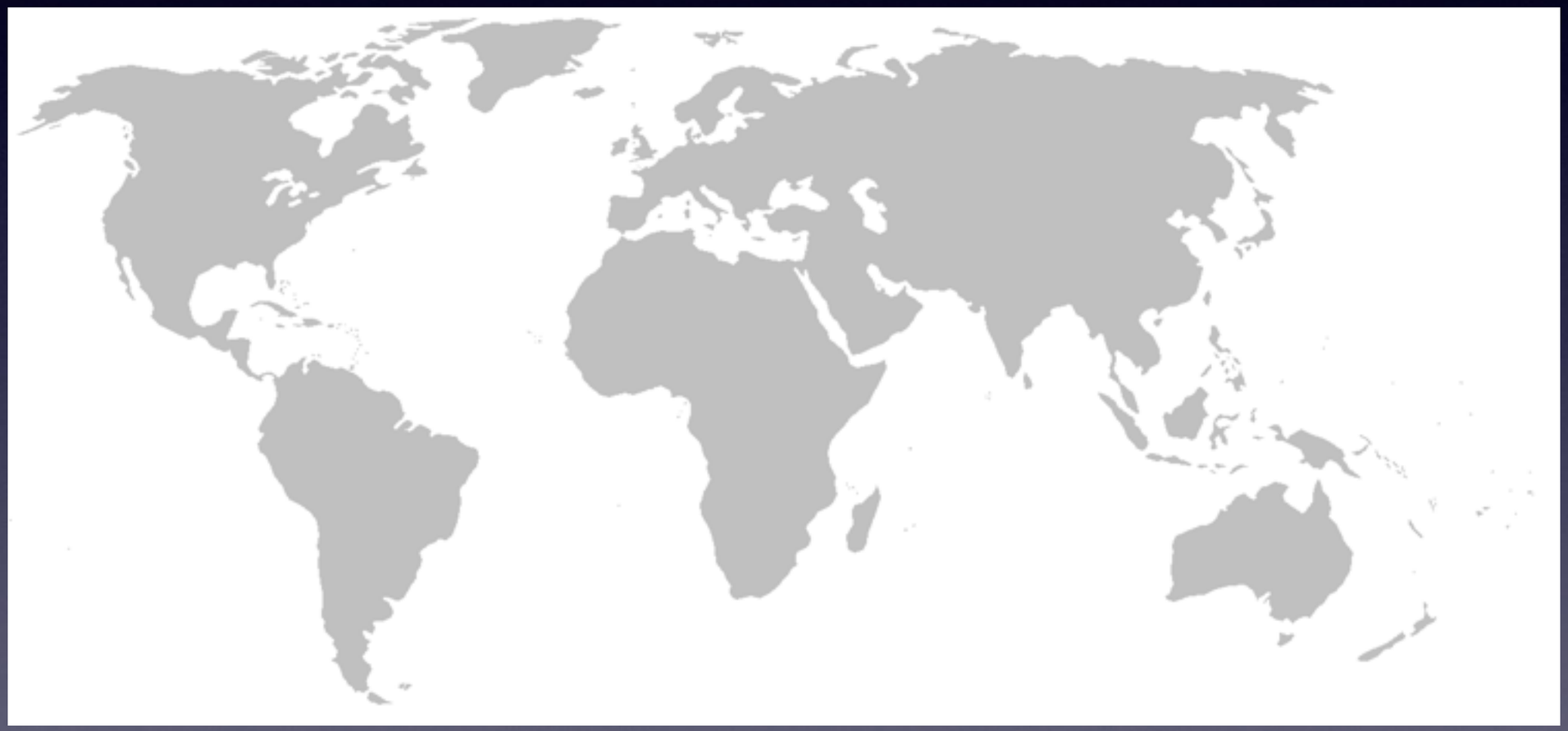
Mantis in Action

Neeraj Joshi and Justin Becker
6/12/2015

Managing a complex operational environment is hard

NETFLIX



NETFLIX

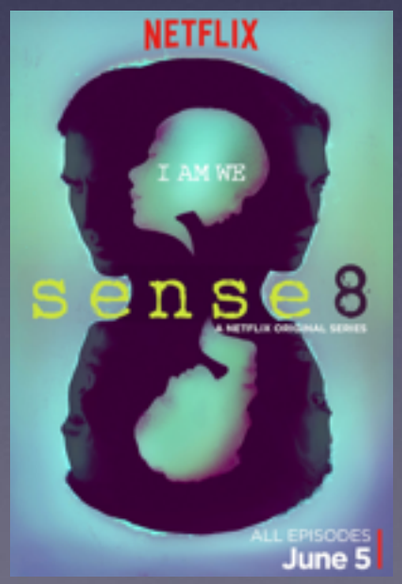

Roku
Netflix Player by Roku
Microsoft
Xbox 360
LG
BD390
Sony
PlayStation 3
SONY
VIZIO
SV422XVT
Nintendo
Wii
Apple
iPad
Samsung
LED C8000
Netflix Ready Devices
From: May2008
To: May2010
instant streaming ready
NETFLIX

NETFLIX


A NETFLIX ORIGINAL SERIES
HOUSE of CARDS


NETFLIX
A NETFLIX ORIGINAL SERIES
MARVEL
DAREDEVIL
ALL EPISODES
April 10


NETFLIX
A NETFLIX ORIGINAL SERIES
ORANGE is the new BLACK
BACK WITH CONVICTION
NEW EPISODES JUNE 6

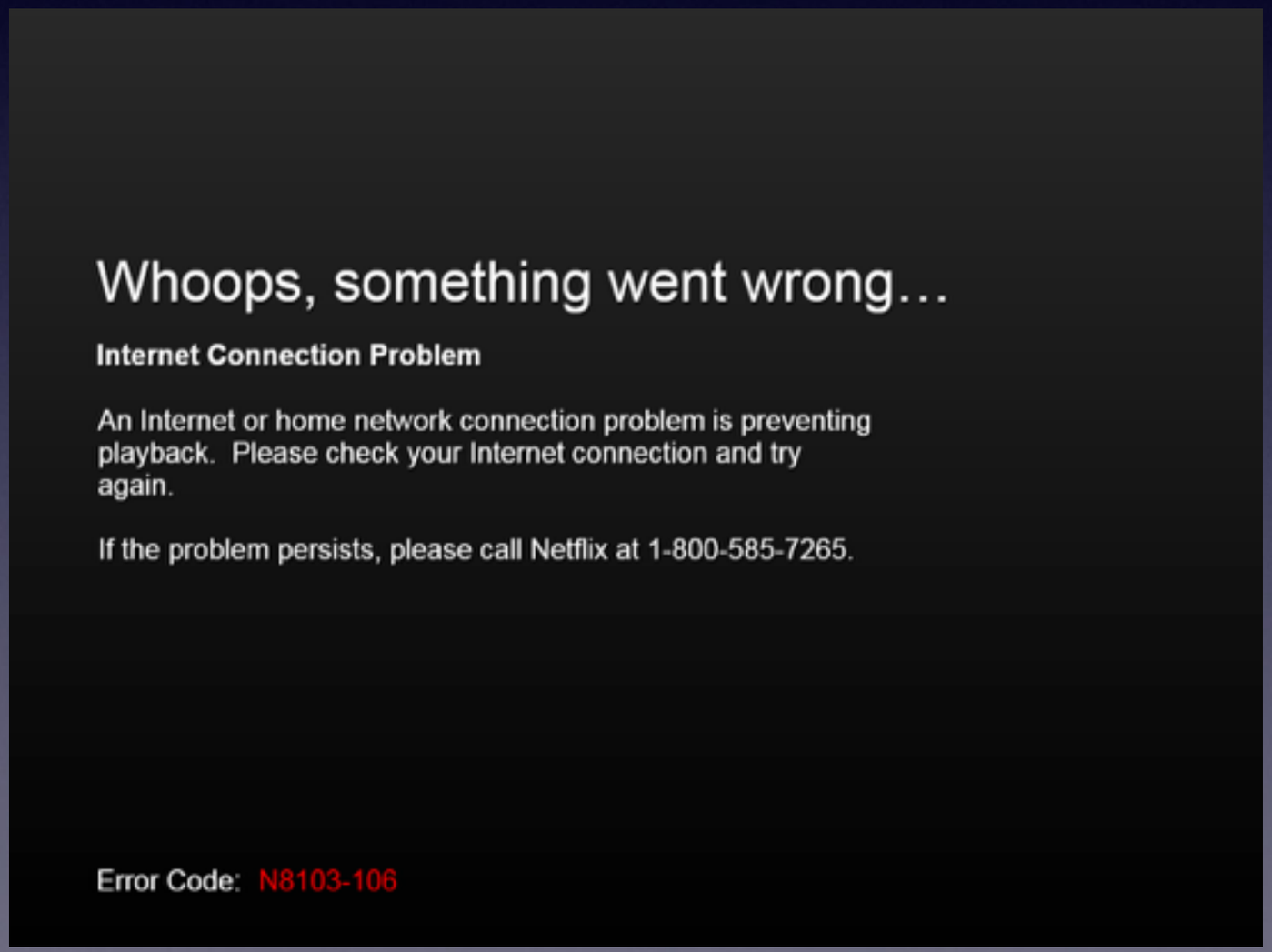

NETFLIX
I AM WE
sense8
A NETFLIX ORIGINAL SERIES
ALL EPISODES
June 5

Developing an understanding of what is going on

Knowing what works

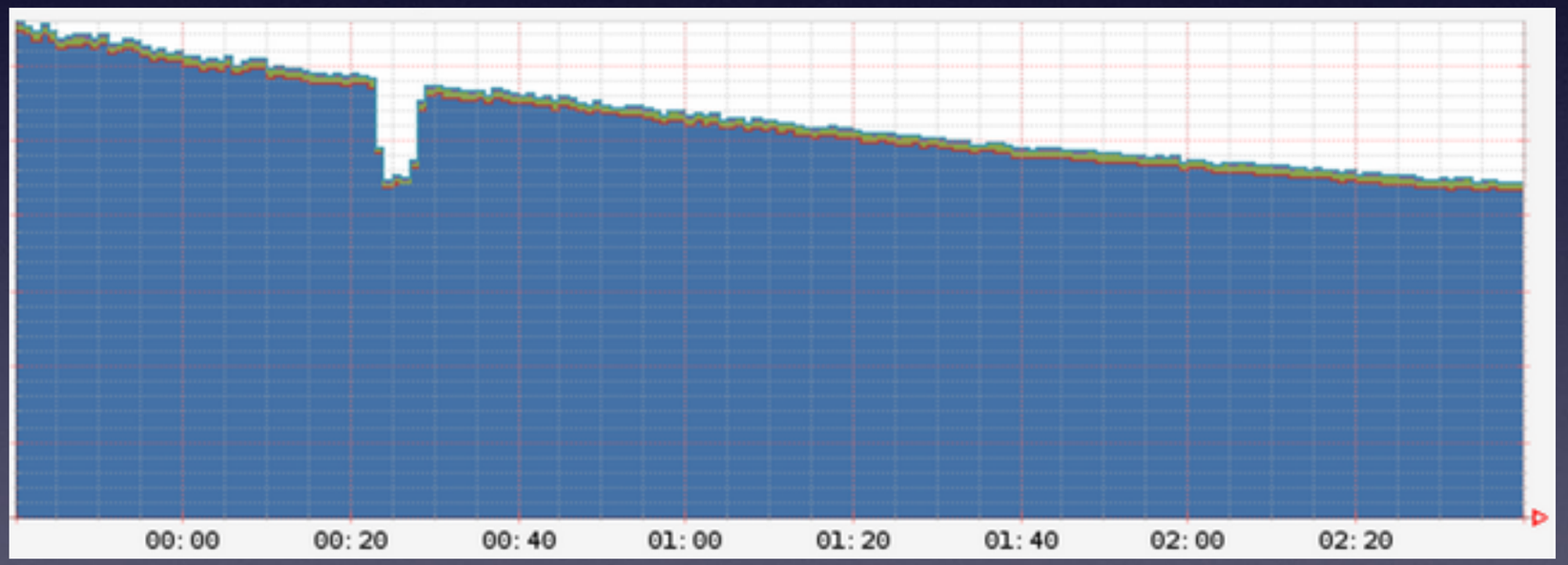
Developing an understanding of what is going on

Identify what doesn't work



Developing an understanding of what is going on

Determining impact when doesn't work



Developing a deep understanding is hard, due to complexity

- Hundreds of software services
- Processing billions of requests
- For millions of users
- Operating in multiple data centers
- Across the globe

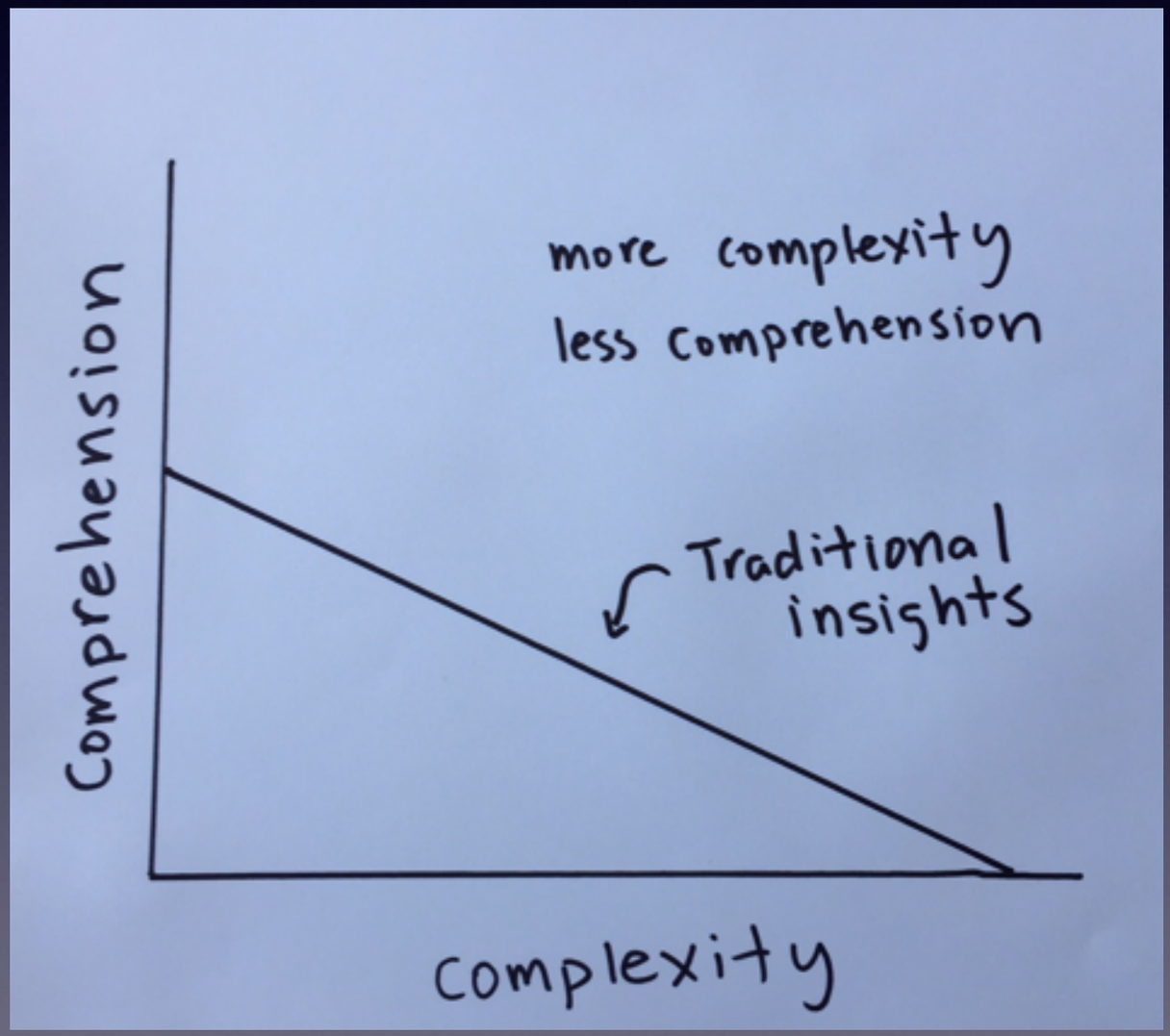
Help us (operators) make sense out of complexity, need tools

specifically, we need…

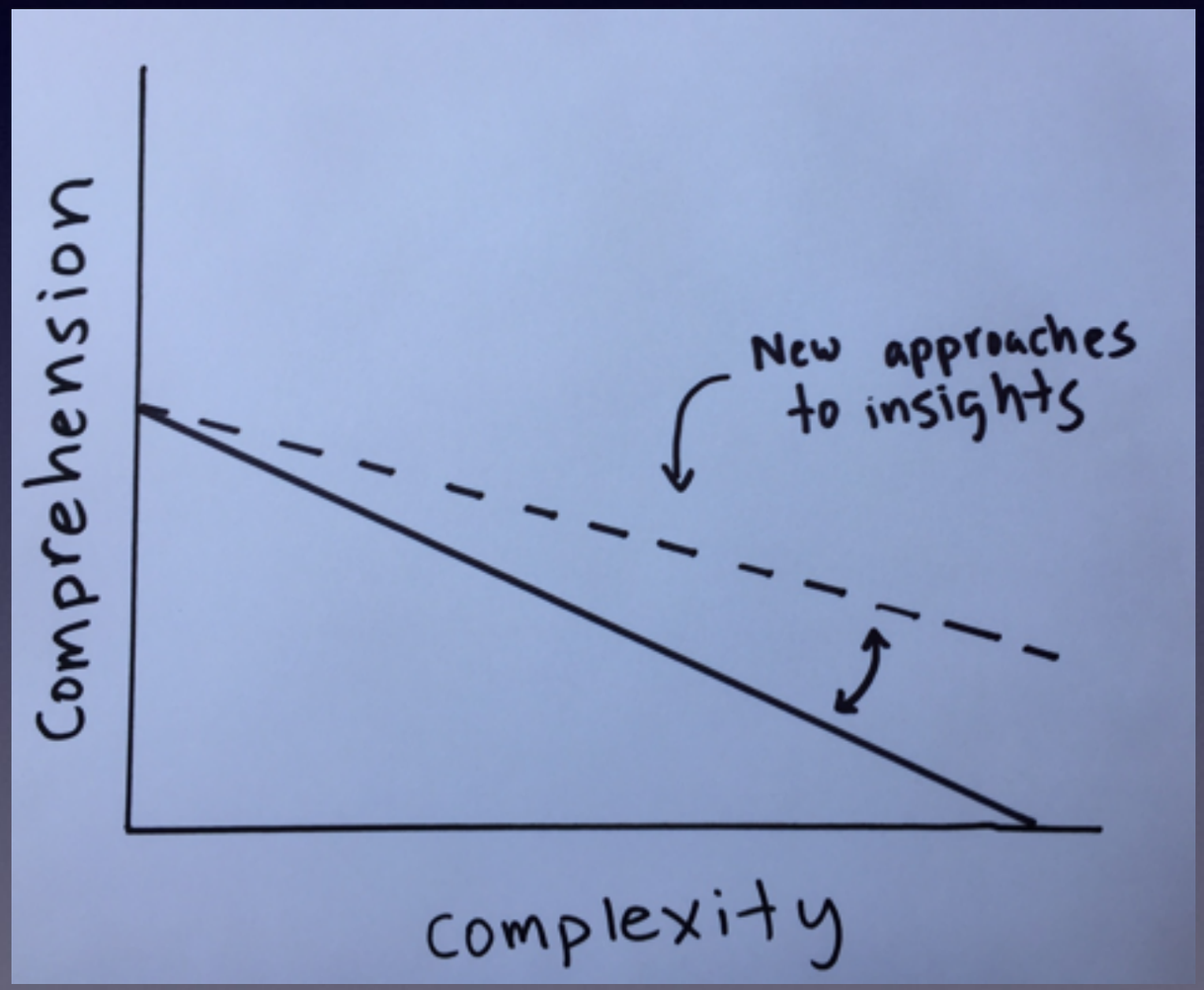
Insight tools

…to help comprehend what is going on in our operational environments

Make the case a relationship exists, between complexity and comprehension



So, in order to manage complex environments, need to rethink insights, *shift the curve*



Identified three insight ‘patterns’ to help shift the curve

- Long tail analysis
- Real time tracking and trending
- Ad hoc investigation

Grouped three patterns into a new effort

Scalable Insights Initiative

Goal is to help us manage (comprehend) our environments given an increase in complexity

Mention specific insights tools to help shift curve

- Realtime Data Explorer
- Realtime Search
- Realtime Application Monitoring

Mention other generic insights jobs to help shift curve

- Short term historical anomaly detection
- Threshold-based anomaly detection
- Realtime metrics generation

Overview of scalable insights in action

- 65-75 total jobs running in 3 regions
- Global access to data
- Processing 4.7 million events per second at peak
- 20 specific data sources and generic adapters
- Ability to startup jobs in ~5 seconds

Demo

EDGE | BLACKBOX | INFRASTRUCTURE | SERVICES | THREAD POOLS | ZUUL

us-east-1 prod jbeck@netflix.com Give Feedback

No Facets Selected +

OVERALL SERVICE STATUS (228)

SORT: RPS

Filter by...

Service	Circuit Breakers Open %	RPS	Error %	Successes	Failures	Short Circuited	Timeouts	Rejections	Cache Responses	Thread Group	Isolation Strategy	Latency
PBC_TRACKS_MANIFEST_ELASTIC	0.0	77628.5	0.0	776285	0	0	0	0	0	PLAYBACK_CONTENT_ELA...	SEMAPH...	50%: 0.0 ms 90%: 0.2 ms
GeoLookupCommand	0.0	72004.9	0.004	720037	12	0	15	0	0	GEO_LOOKUP	SEMAPH...	50%: 0.0 ms 90%: 1.0 ms
NDCSubscriberGetCustomerAccountByIde..	0.0	59575.8	0.02	595758	0	0	136	1	4489	SUBSCRIBER	THREAD	50%: 3.4 ms 90%: 12.0 ms
IdentityCookieAuth	0.0	52055.1	0.01	520551	0	0	56	0	28	IDENTITY	THREAD	50%: 0.5 ms 90%: 1.1 ms
CryptexMacVerifyViaSemaphore	0.0	42194.1	0.0	421941	0	0	0	0	0	CRYPTEX	SEMAPH...	50%: 0.0 ms 90%: 0.8 ms
NDCMapReadEVCachePersister	0.0	36443.6	0.002	364436	0	0	0	9	0	MAP	THREAD	50%: 1.3 ms 90%: 2.9 ms
PhsRemoteGetBookmarksCommand	0.0	35264.9	0.03	352617	32	0	0	58	0	PHS	THREAD	50%: 8.6 ms 90%: 24.7 ms
CryptexDecryptViaSemaphore	0.0	33492.9	0.0	334929	0	0	0	0	0	CRYPTEX	SEMAPH...	50%: 0.0 ms 90%: 0.9 ms
NDCABGetAllocations	0.0	32822.7	0.0006	328225	2	0	0	0	1545560	AB	THREAD	50%: 5.2 ms 90%: 10.4 ms
HystrixGetABServiceClient	0.0	24601.6	0.0	246016	0	0	0	0	1080909	AB	SEMAPH...	50%: 0.0 ms 90%: 1.0 ms
ADCIdentityReadFromCookie	0.0	16602.0	0.0006	166020	0	0	1	0	11	IDENTITY	SEMAPH...	50%: 0.0008 m 90%: 1.0 ms
GetCassandraSequenceNumberCommand	0.0	15543.5	0.0	155435	0	0	0	0	0	CASSANDRA_SEQUENCEN...	THREAD	50%: 4.0 ms 90%: 6.9 ms
NDCMapGetCachedLists	0.0	13897.6	0.4	138437	539	0	0	0	14633	MAP	THREAD	50%: 11.9 ms 90%: 26.9 ms
QTGetQTVGenres	0.0	11658.2	0.0	116582	0	0	0	0	0	QT	SEMAPH...	50%: 0.001 ms 90%: 1.0 ms
PDS_KEEP_ALIVE	0.0	10460.9	0.006	104605	4	0	0	2	0	PDS_EVENTS	THREAD	50%: 10.5 ms 90%: 26.8 ms
CinematchGetVisitorPredictions	0.0	9473.3	0.0	94733	0	0	0	0	115481	CINEMATCH_PS	SEMAPH...	50%: 5.0 ms 90%: 27.1 ms
CinematchGetVisitorVideoRatings	0.0	9227.6	0.0	92276	0	0	0	0	41000	CINEMATCH_RS	THREAD	50%: 1.3 ms 90%: 14.4 ms
YellowSquareCustomerPreferenceRetriev...	0.0	9098.6	0.001	90985	1	0	0	0	686250	YELLOWSQUARE	THREAD	50%: 1.0 ms 90%: 3.8 ms
QTGetCharacterBio	0.0	7291.6	0.0	72916	0	0	0	0	0	QT	SEMAPH...	50%: 0.0 ms 90%: 0.07 ms
QTGetCharacterName	0.0	7159.4	0.0	71594	0	0	0	0	2147	QT	SEMAPH...	50%: 0.0 ms 90%: 0.2 ms
PlaylistGet	0.0	7077.1	0.0	70771	0	0	0	0	357634	PlaylistGet	THREAD	50%: 1.0 ms 90%: 13.7 ms
GetTrackId	0.0	5514.6	0.0	55146	0	0	0	0	0	TRACKING	SEMAPH...	50%: 0.0006 m 90%: 0.8 ms
NDCMapGetGallery	0.0	5243.1	0.01	52425	6	0	0	0	16694	MAP	THREAD	50%: 12.0 ms 90%: 69.9 ms
MapGetPlaylist	0.0	4117.4	0.01	41169	5	0	0	0	0	MAP	THREAD	50%: 11.5 ms 90%: 65.4 ms

MANTIS REALTIME EVENTS | GENERIC QUERY BUILDER | SIMPLE QUERY BUILDER | PRODUCTION

Mantis Realtime - This will launch a temporary Mantis job, and connect to it with SSE widget. MARS Wiki

Choose Sources: × APIRequestSource

APIRequestSource

Choose Fields:

Not finding the field you are looking for. You can add your custom field here? Add New Field

Add Criterion

Add Custom Criterion

Query to Mantis:

Subscription Id: :_

Elastic Search TTL(seconds): 86400

Save matched events To Elastic Search: Submit

Verify your query against a sample Event Verify Query

Edit Query: Edit Query

SSE Job Output Start Clear 0 records per second - Sampling:

Content length: 0 (Max 200)

Server-Sent Event via WebSocket

SSE URL:

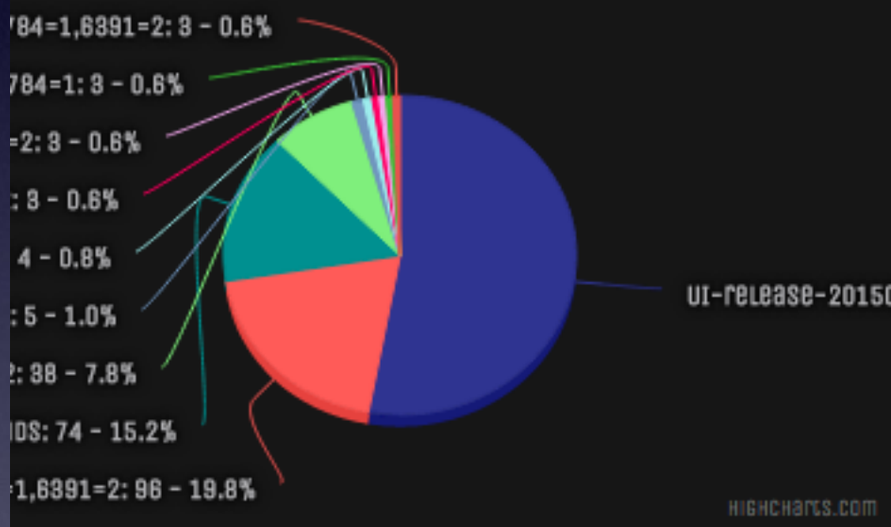
Curl URL:

time

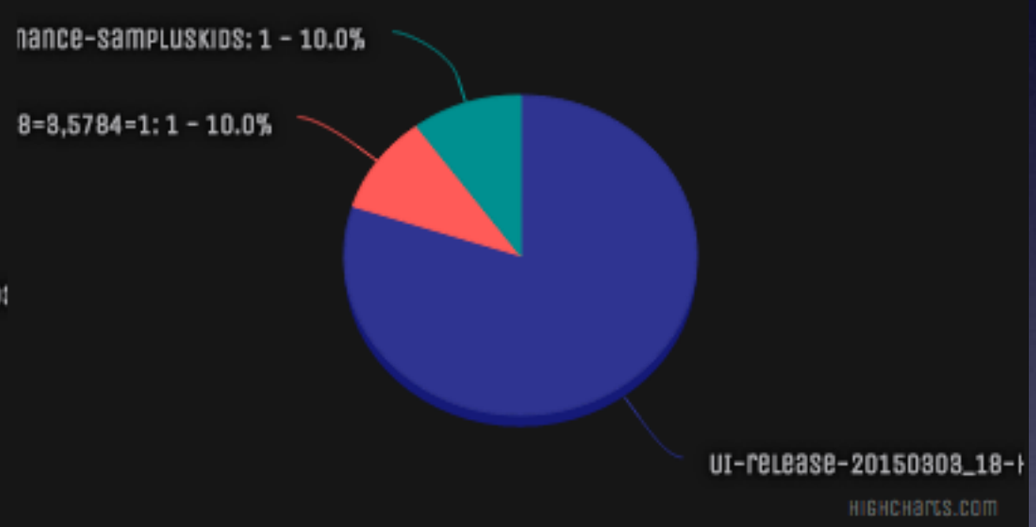
STARTPLAYS PER SECOND (SPS)

12:58:30 12:58:32 12:58:34 12:58:36 12:58:38 12:58:40

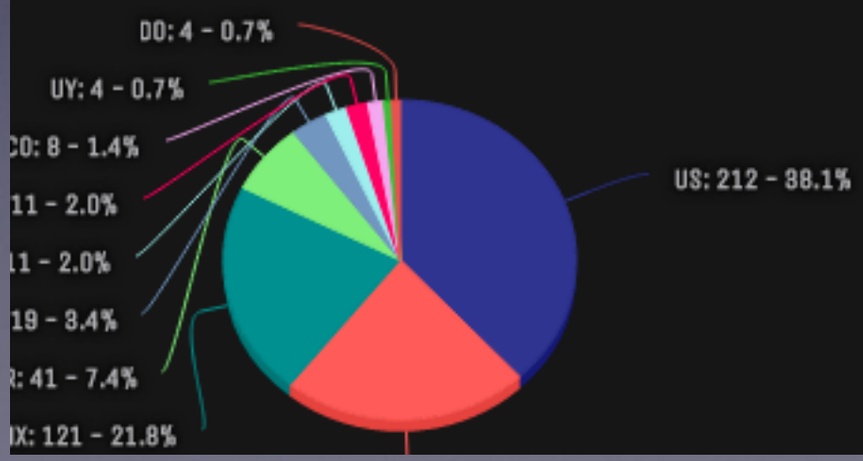
UIVER_INFO



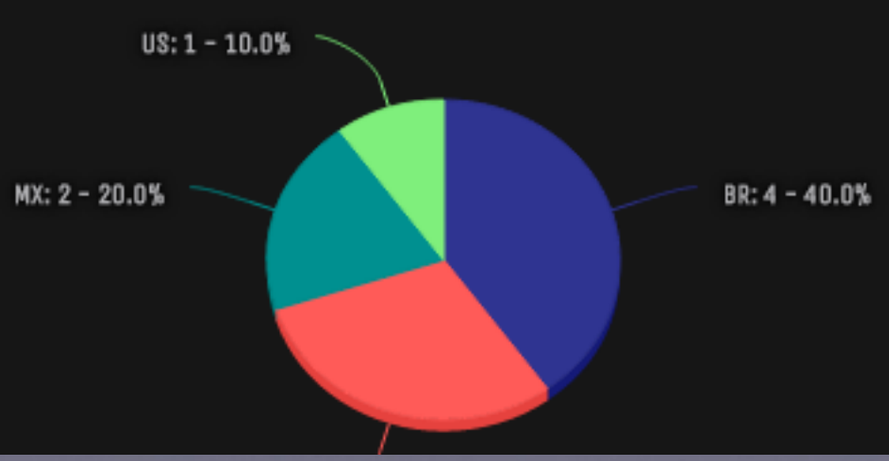
UIVER_ERROR



GEO_DATA.COUNTRY_CODE_INFO



GEO_DATA.COUNTRY_CODE_ERROR



MANTIS REALTIME EVENTS | GENERIC QUERY BUILDER | SIMPLE QUERY BUILDER

Mantis Realtime - Server Send Events - Gets a pre-configured form with frequently queried fields for each available job. MARS Wiki

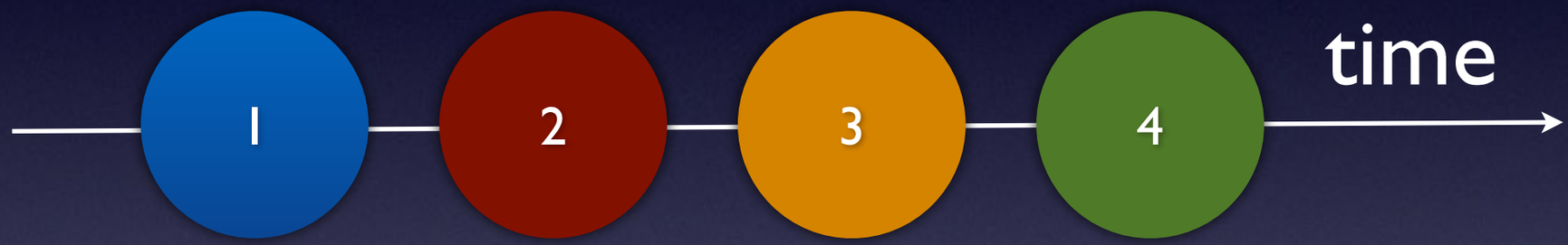
Select Source:	StartPlayLogBlobSource	
Event time:	match	
Customer Id:	match	
Xid:	match	
Movie Id:	match	
Device Id:	match	
Country Code:	match	
ESN:	match	
Error Type:	match	
Type:	match	
CDN ID:	match	

Mantis, a reactive stream processing system

Some Basics Concepts

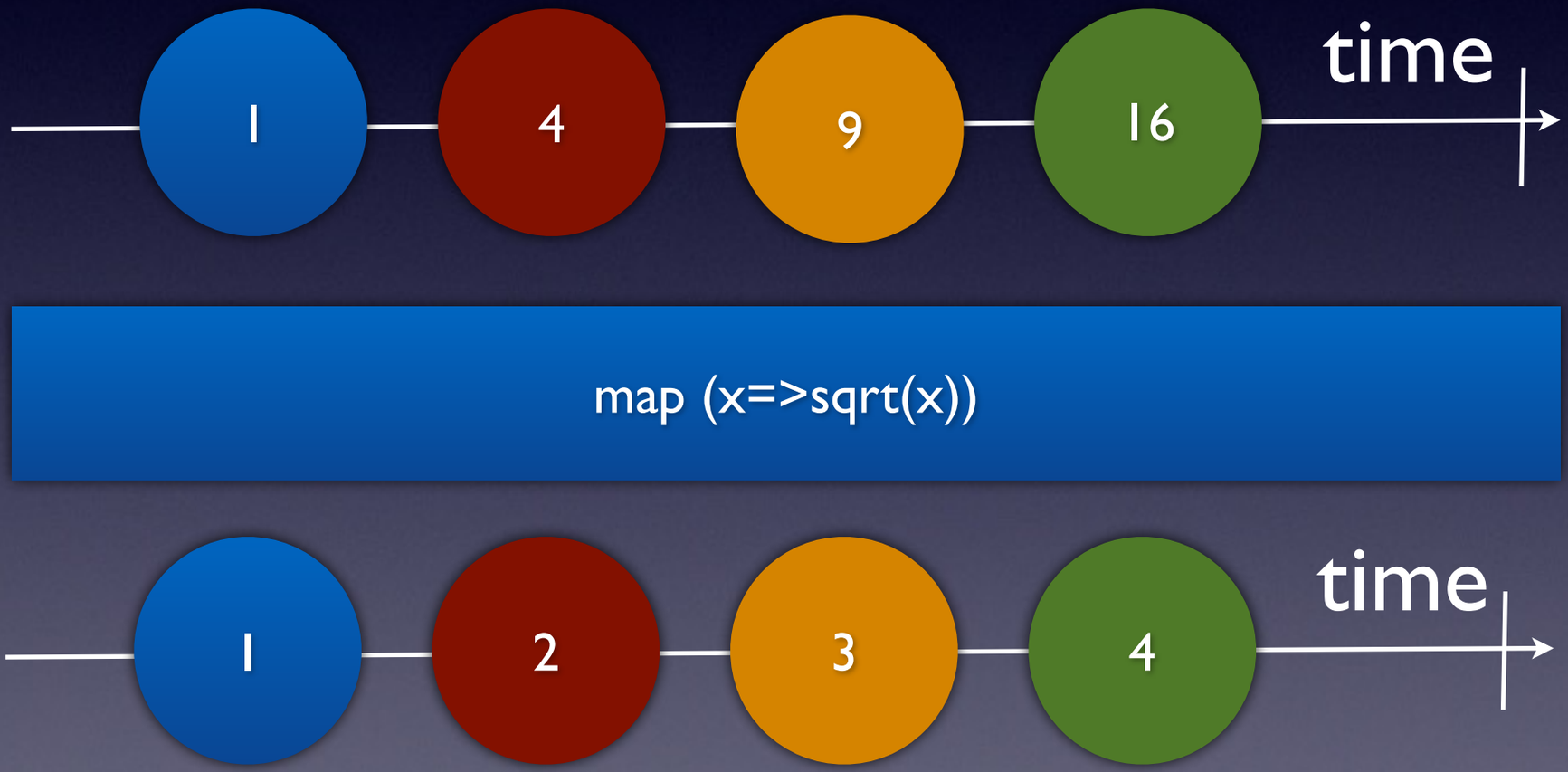
Stream

Sequence of Events



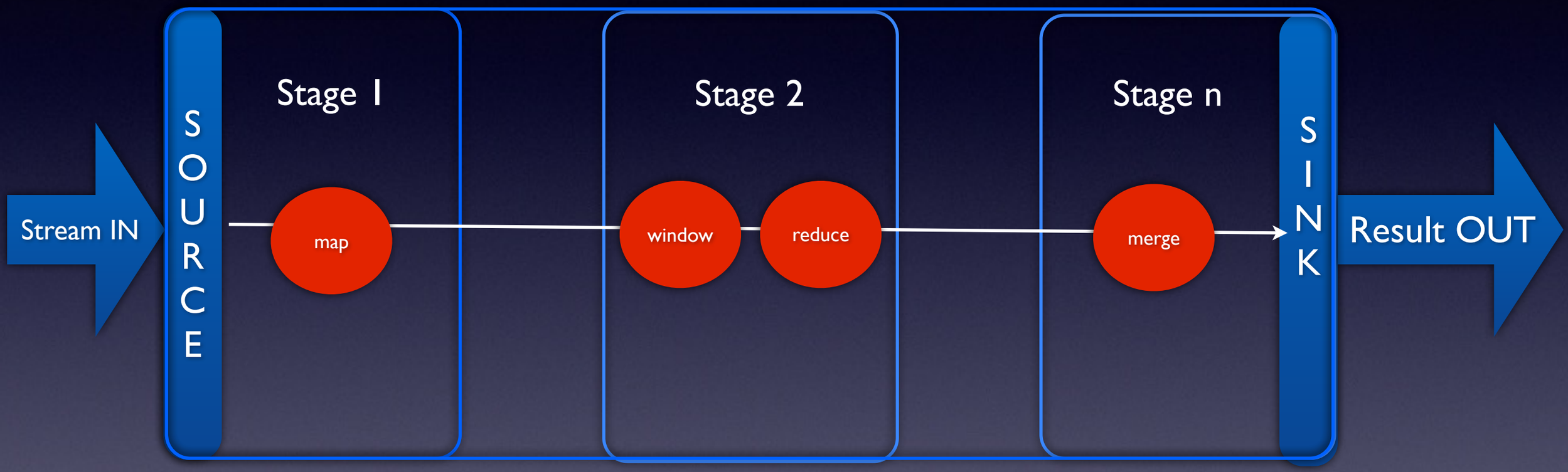
Higher Order Functions

Transformations applied to a Stream
to create a new stream

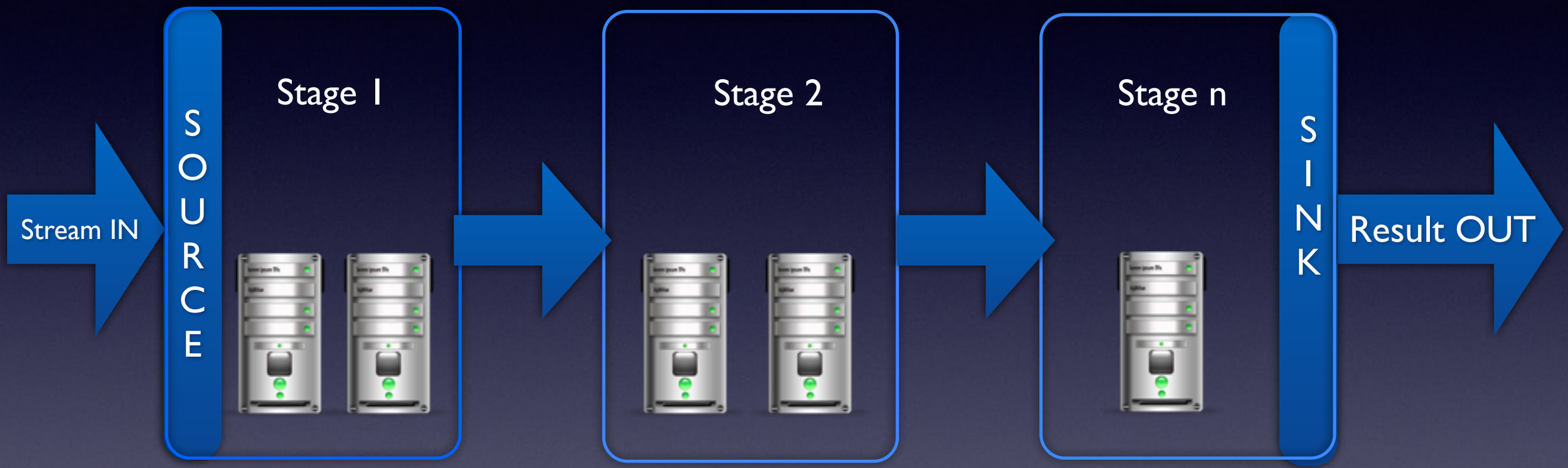


Mantis Job

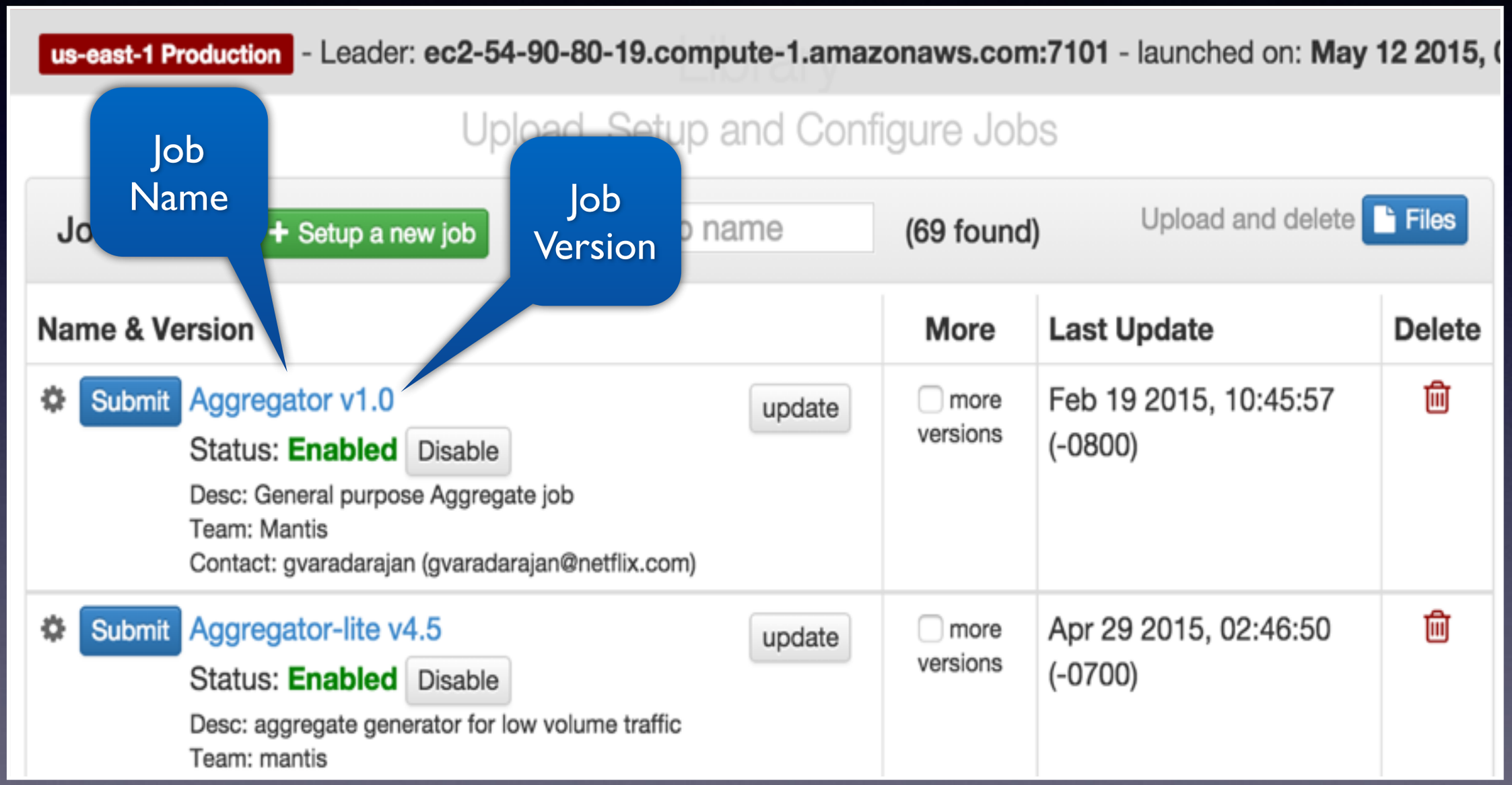
A sequence of functions applied to a stream



Mantis Job



Named Jobs


us-east-1 Production - Leader: ec2-54-90-80-19.compute-1.amazonaws.com:7101 - launched on: May 12 2015,
Upload, Setup and Configure Jobs
Job Name
Job Version
+ Setup a new job
(69 found)
Upload and delete Files
Name & Version
More
Last Update
Delete
Submit
Aggregator v1.0
update
Status: Enabled Disable
Desc: General purpose Aggregate job
Team: Mantis
Contact: gvaradarajan (gvaradarajan@netflix.com)
more versions
Feb 19 2015, 10:45:57 (-0800)
Submit
Aggregator-lite v4.5
update
Status: Enabled Disable
Desc: aggregate generator for low volume traffic
Team: mantis
more versions
Apr 29 2015, 02:46:50 (-0700)

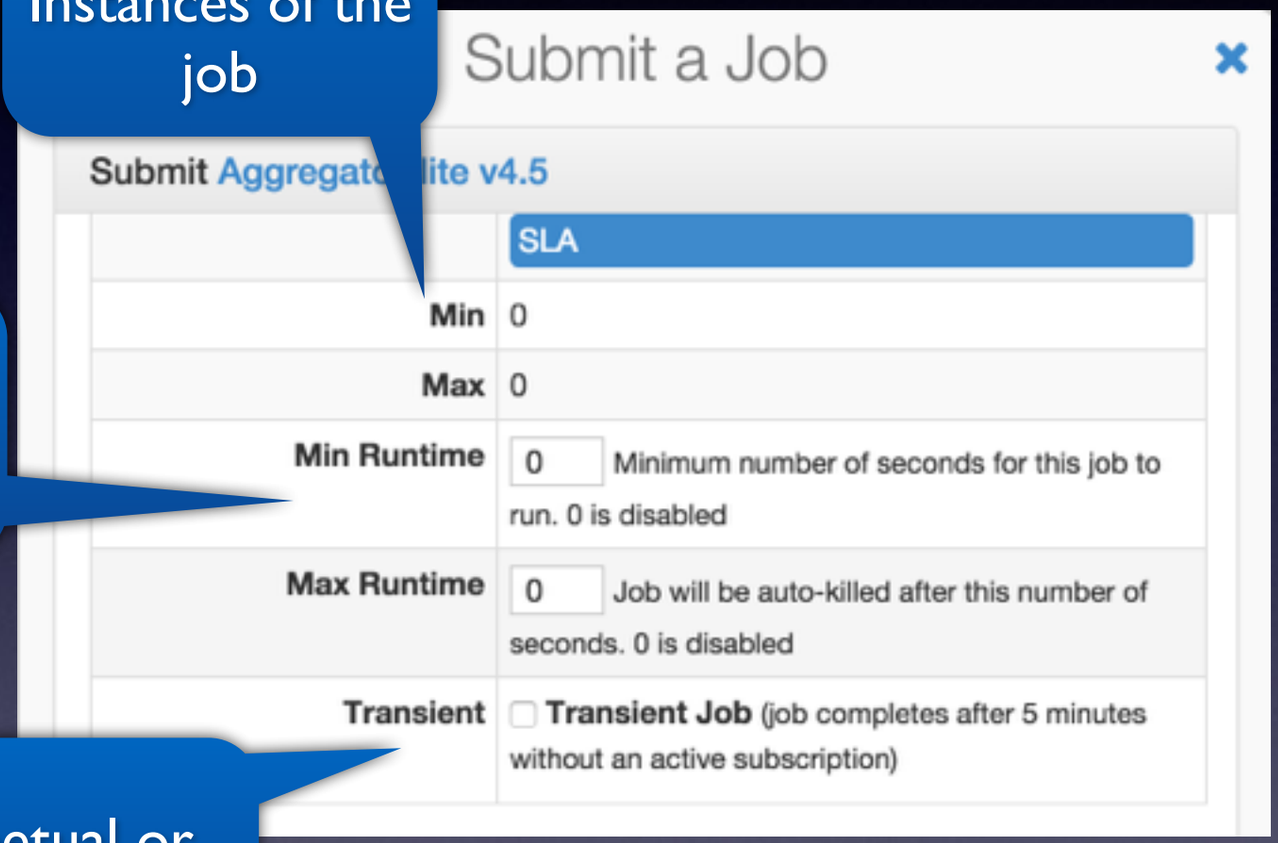
Are Parameterized

Have SLAs

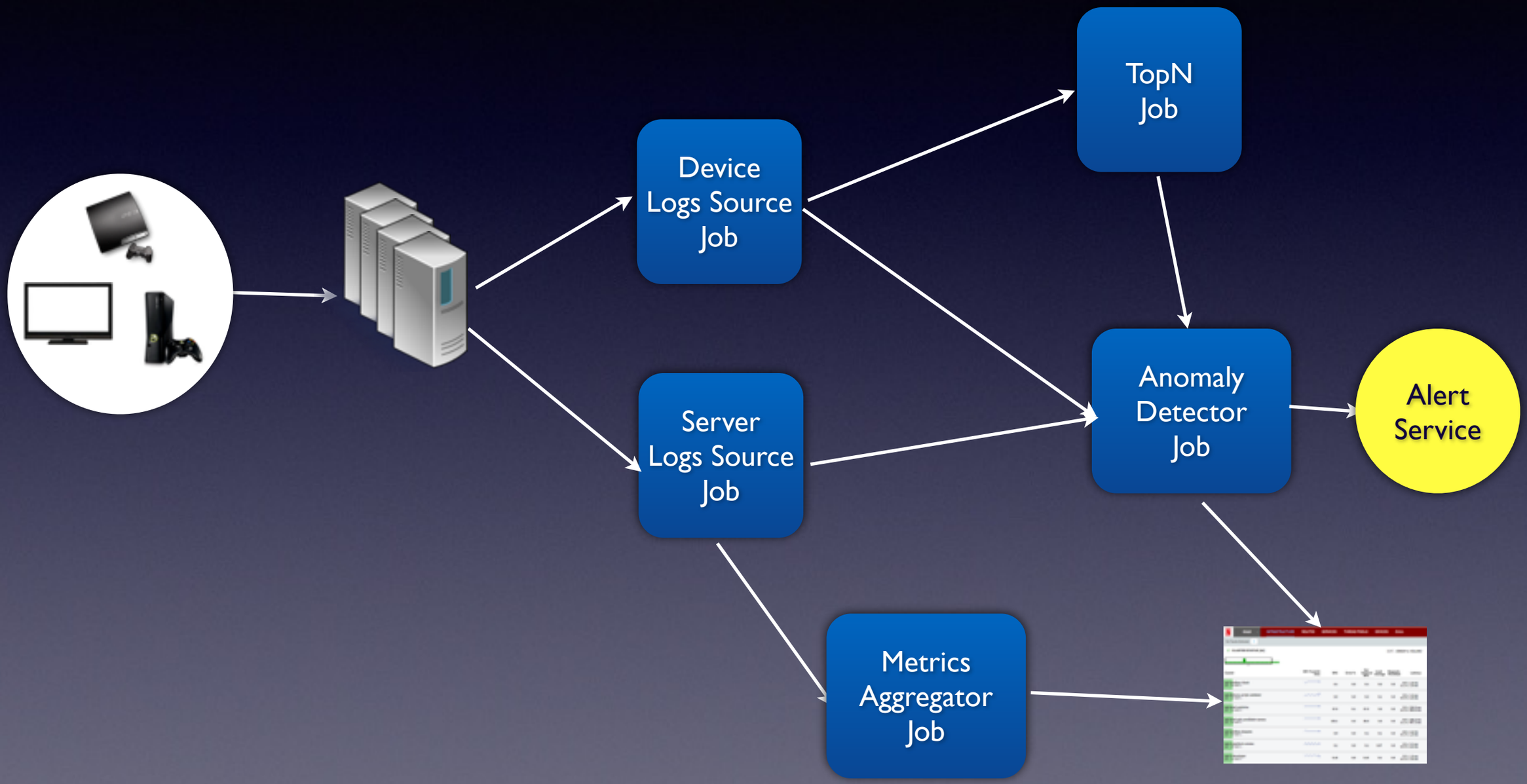
Min/Max Instances of the job

Min/Max runtime

Perpetual or Transient



Can be chained



That is fine but...

How does Mantis meet the Scalable Insights challenge?

Key Requirements

- Cost (Utilization) Sensitive
- Optimize for low latency
- High Throughput
- Resilient

Minimizing Costs

amazon web services™ | EC2

Elastic Clusters

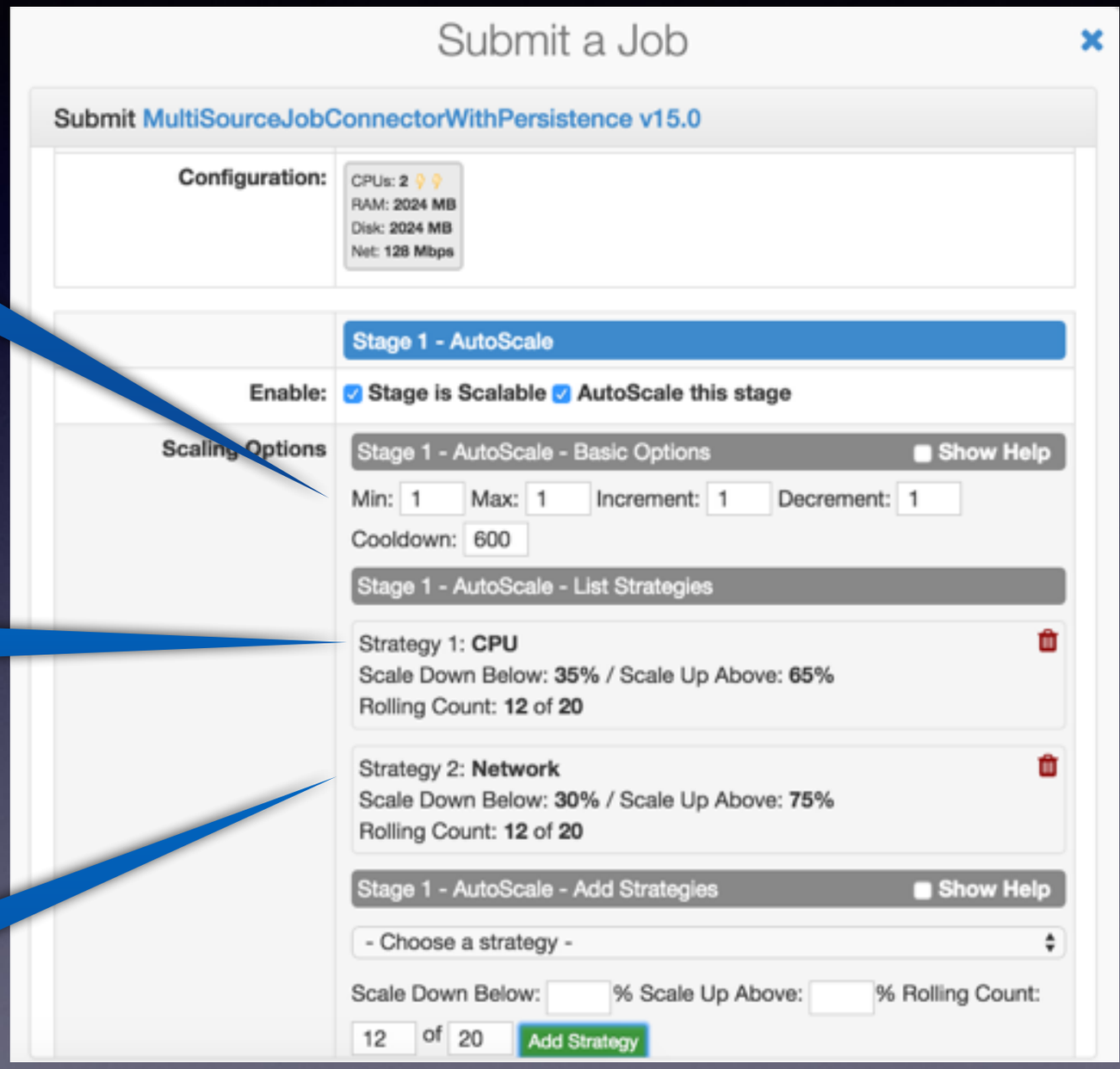
Elastic Jobs

Job Autoscaling

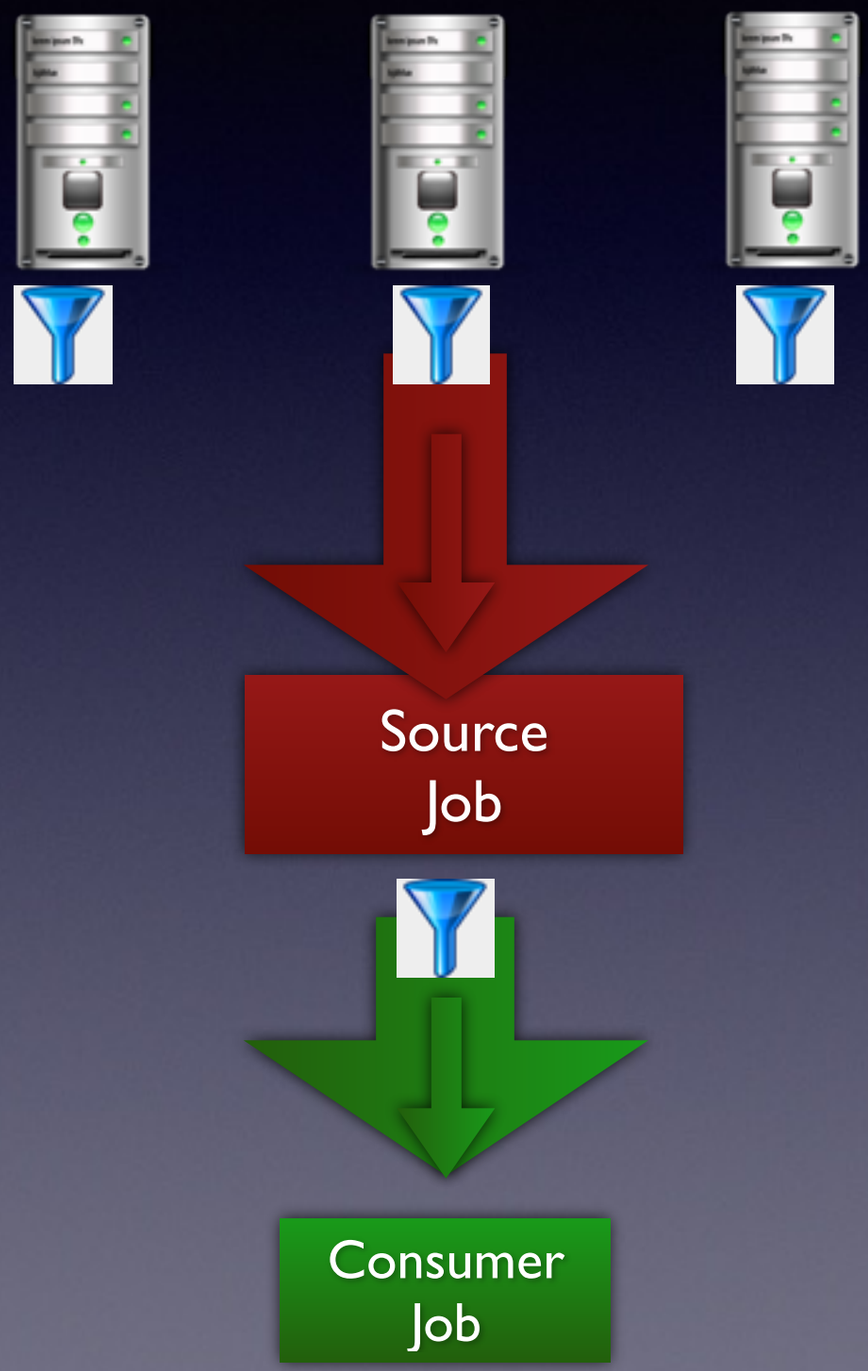
Scaling Config

CPU Strategy

Network Strategy



Filtering at Source



Low latency - High Throughput

To block or not to block?

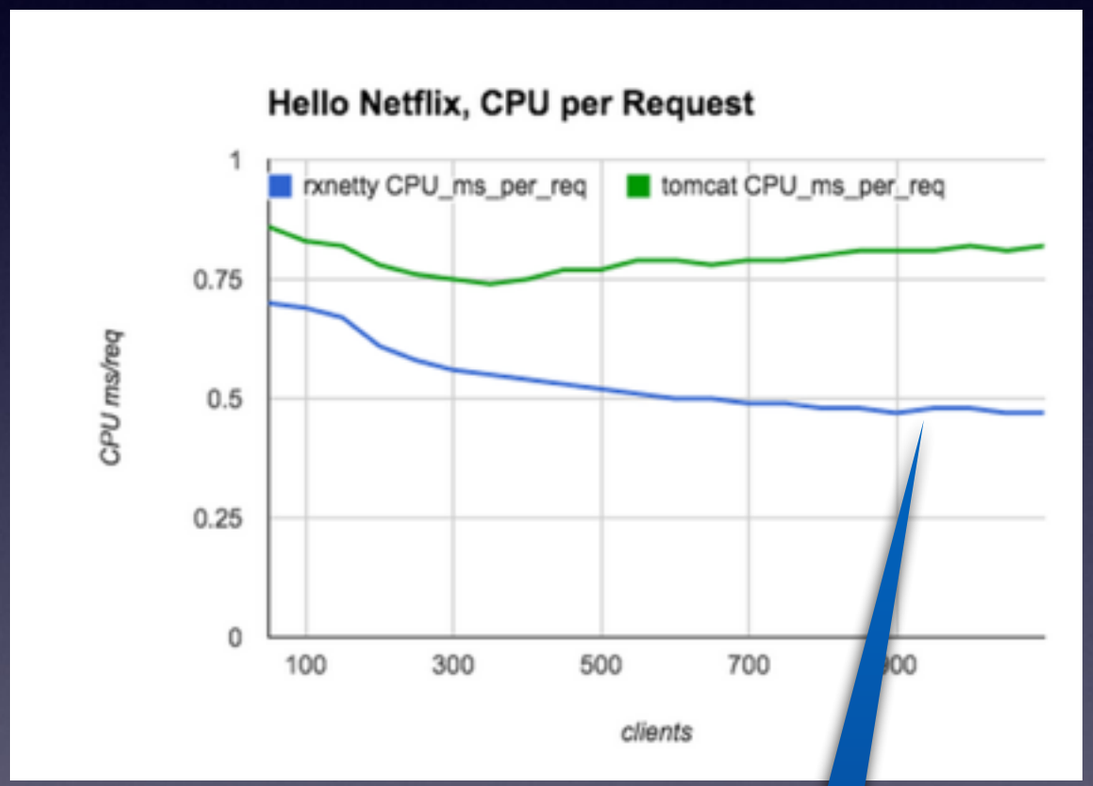
RxNetty (non-blocking) vs Tomcat (blocking)

by Brendan Gregg
@brendangregg

https://docs.google.com/presentation/d/18i-d72m7tD4wKlzm-1PCR8g62l66_9Btbg5-fuFRqf0/edit#slide=id.g761289dab_0_77

CPU consumption

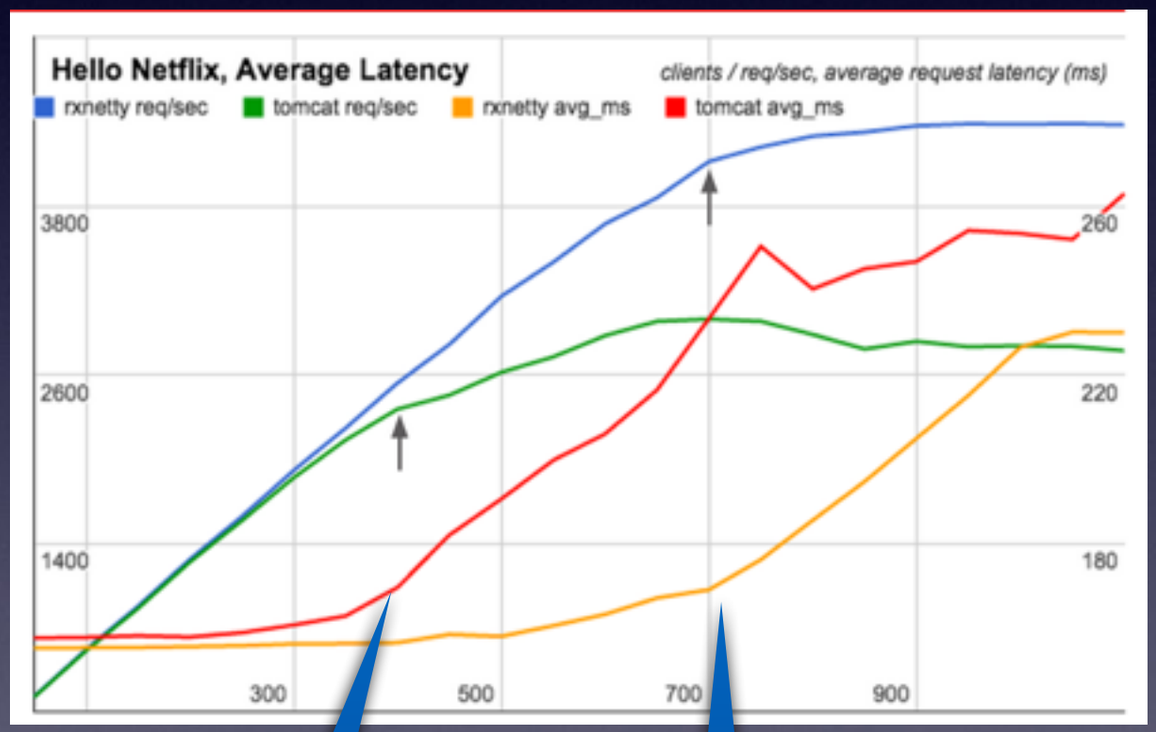
- RxNetty consumes less CPU / request
- Reduced thread migration
- Lower object allocation rate



CPU consumed reduces as load increases

Lower latency

- RxNetty has lower latency under high load
  - fewer lock contentions
  - fewer thread migrations



Async Processing

Non-blocking I/O

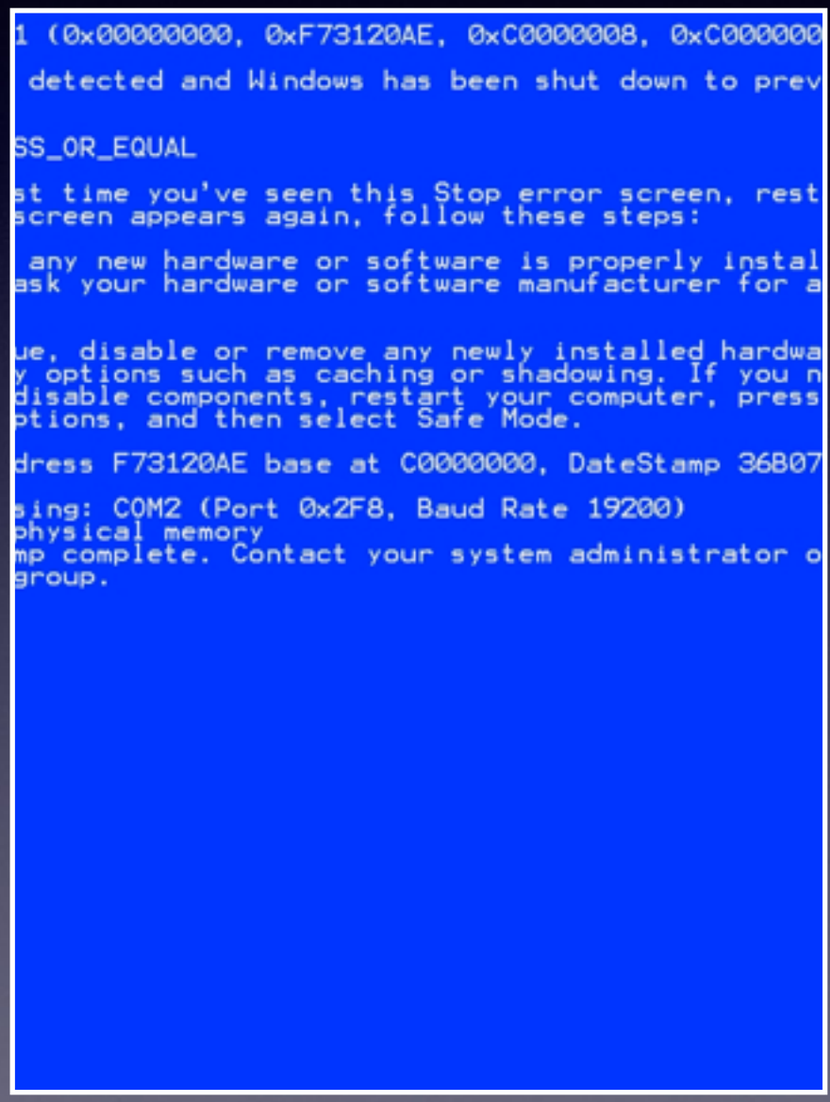
Async Processing

Designed for Resilience

http://signsofpolitics.blogspot.com/2009/03/around-and-about-resilience.html

Server Resilience

- Servers crashes inevitable
- Server health constantly monitored with heartbeats
  - Crashed servers replaced
  - Lost jobs relaunched



Network Resilience

- Long lived connections can fail
- Connection topology is constantly monitored and corrected

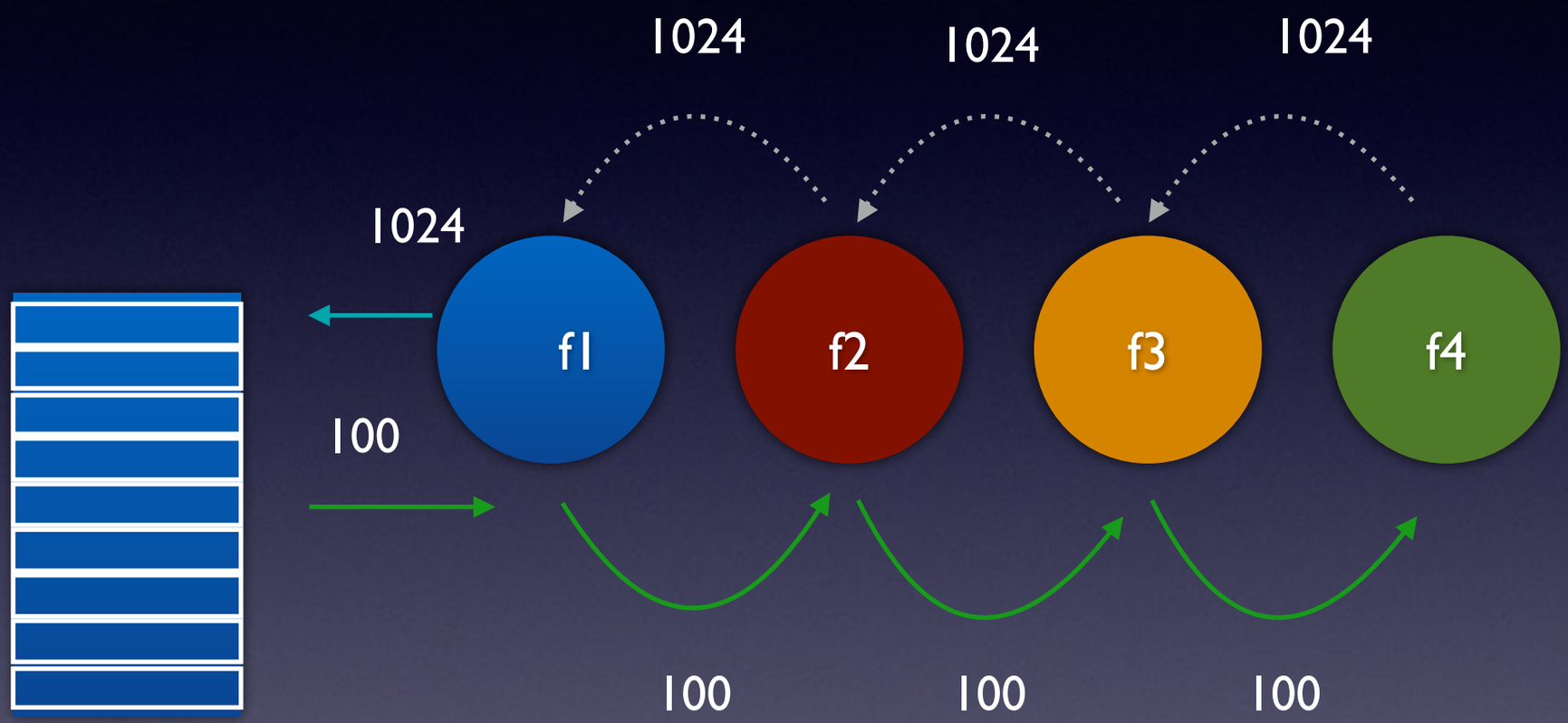
Backpressure

Cold Source

Amazon SQS

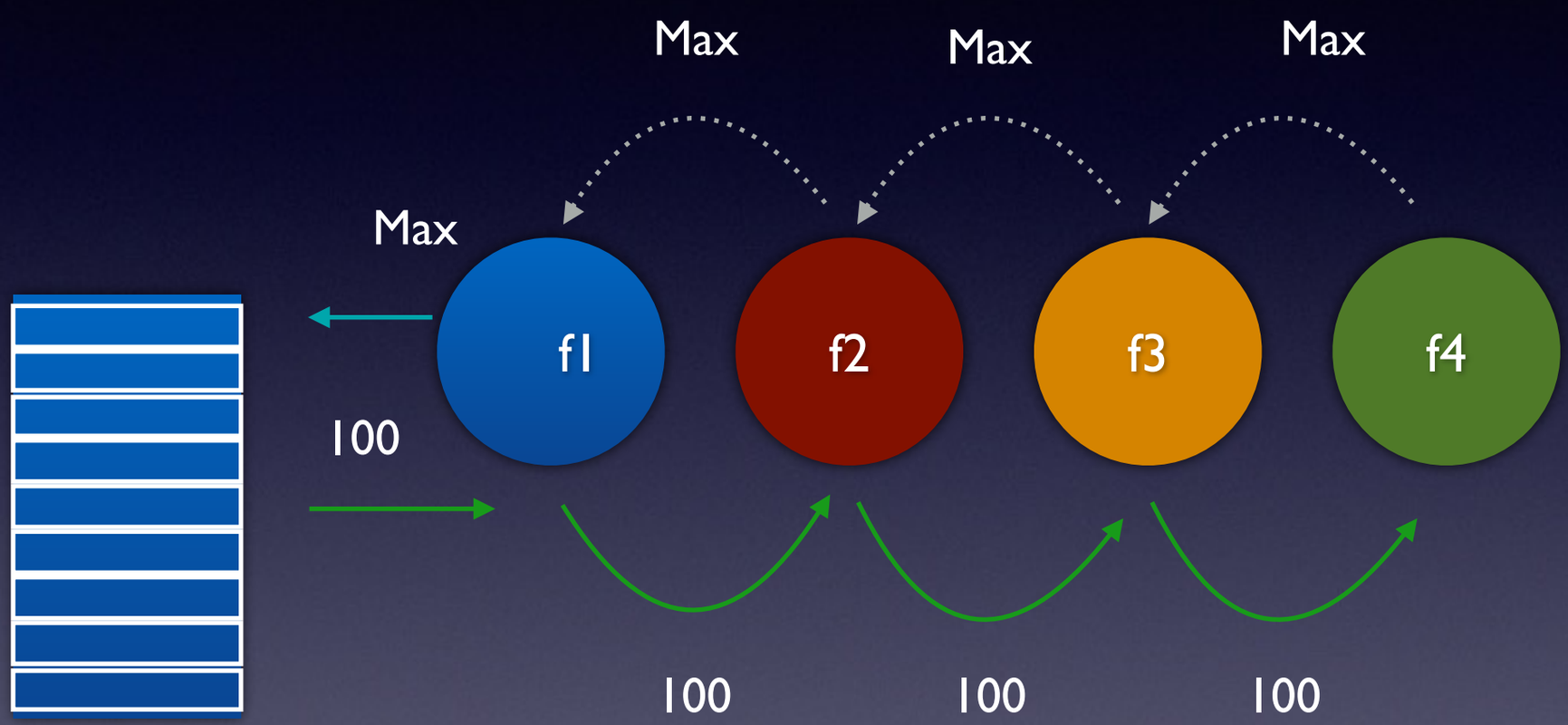
Hot Source

Reactive push-pull (Cold Source)



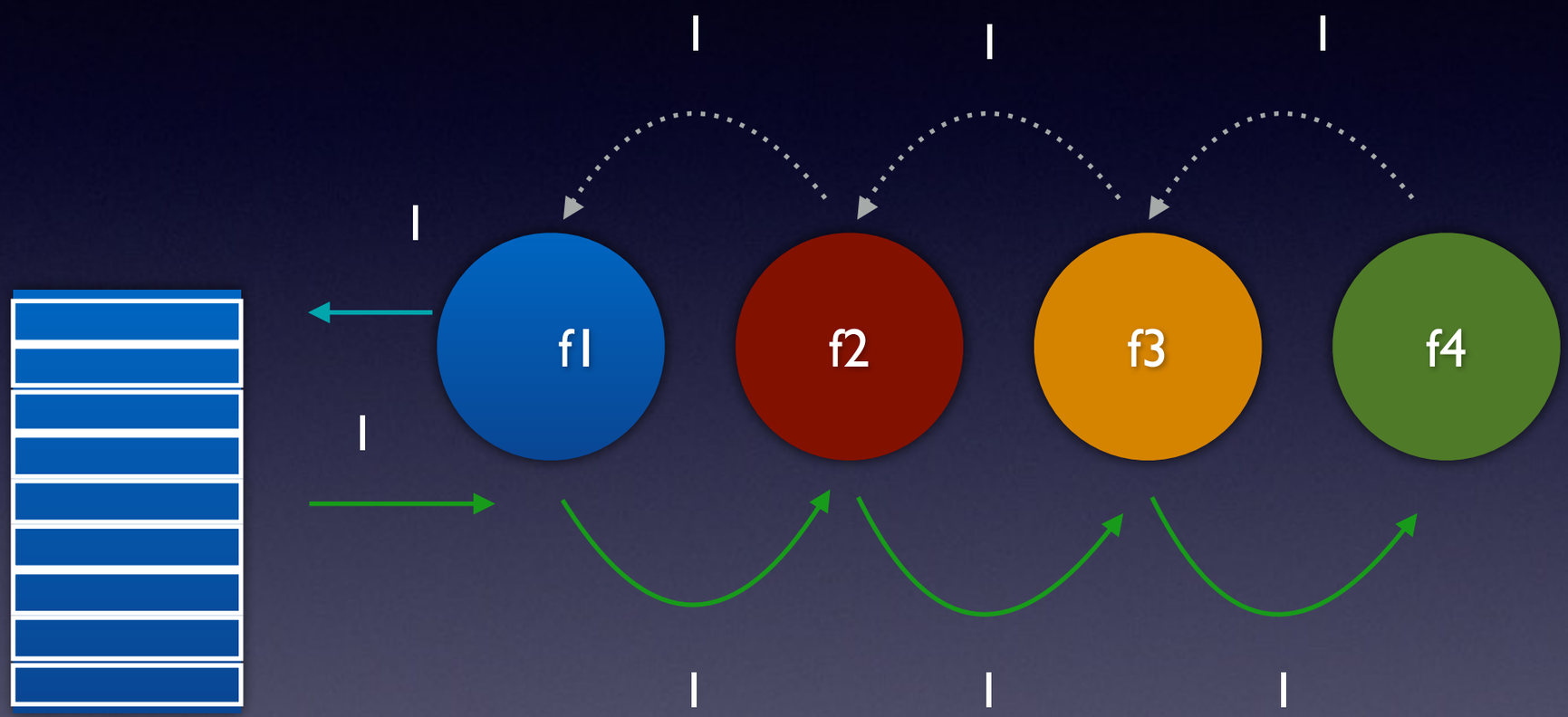
Cold Source

Push mode



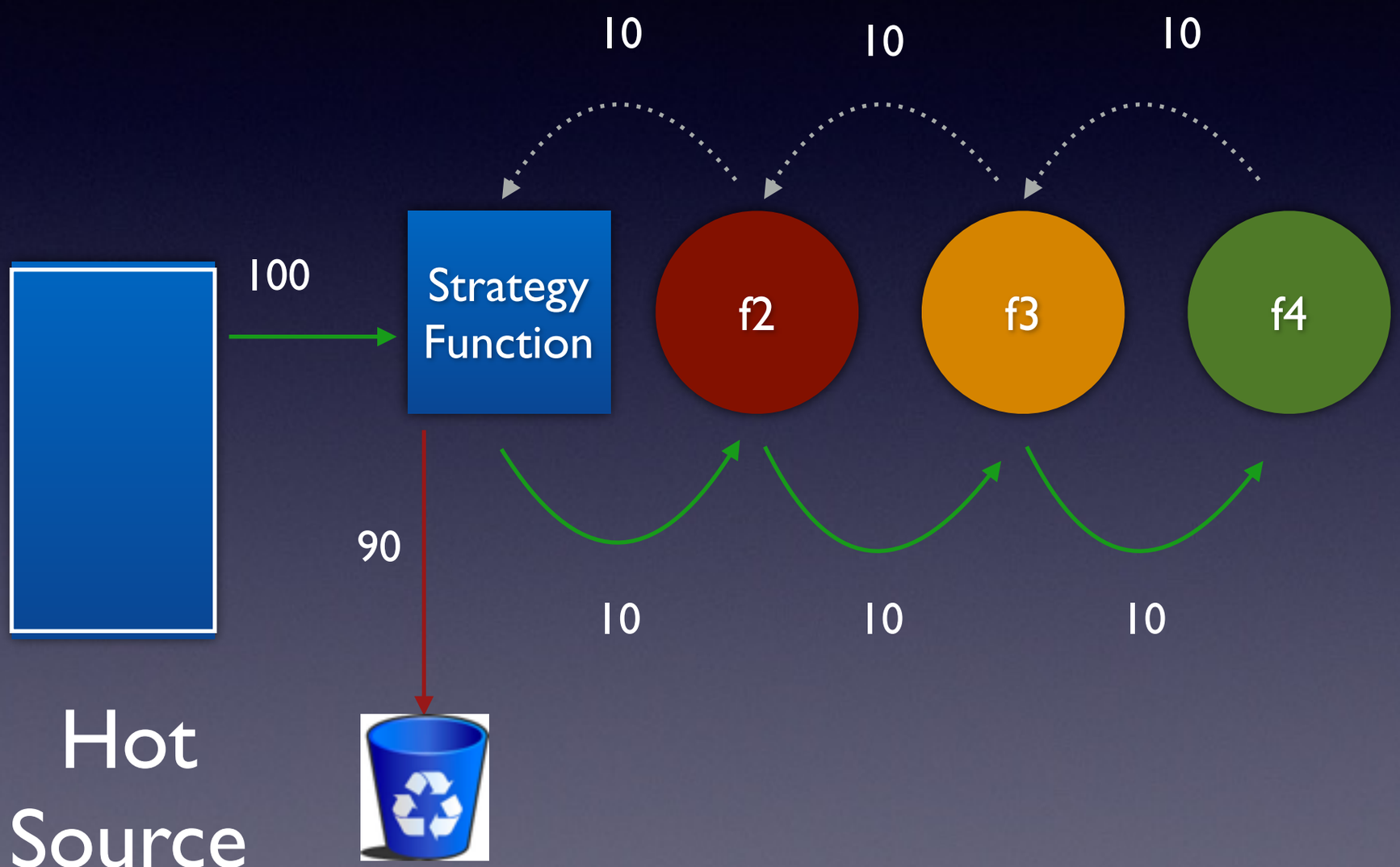
Cold Source

Pull mode



Cold Source

Backpressure Strategies (Hot Source)



Hot Source

Drop Buffer Scale-up

Questions