



# Bringing Business Awareness to Your Operation Team

## Nagios World Conference 2013

Nicolas Brousse  
Director, Operations Engineering



# About TubeMogul...

---

- Founded in 2006
- Formerly a video distribution and analytics platform
- TubeMogul is a Brand-Focused Video Marketing Company
  - Build for Branding
  - Integrate real-time media buying, ad serving, targeting, optimization and brand measurement

TubeMogul simplifies the delivery of video ads and maximizes the impact of every dollar spent by brand marketers

<http://www.tubemogul.com>



# About TubeMogul...

---

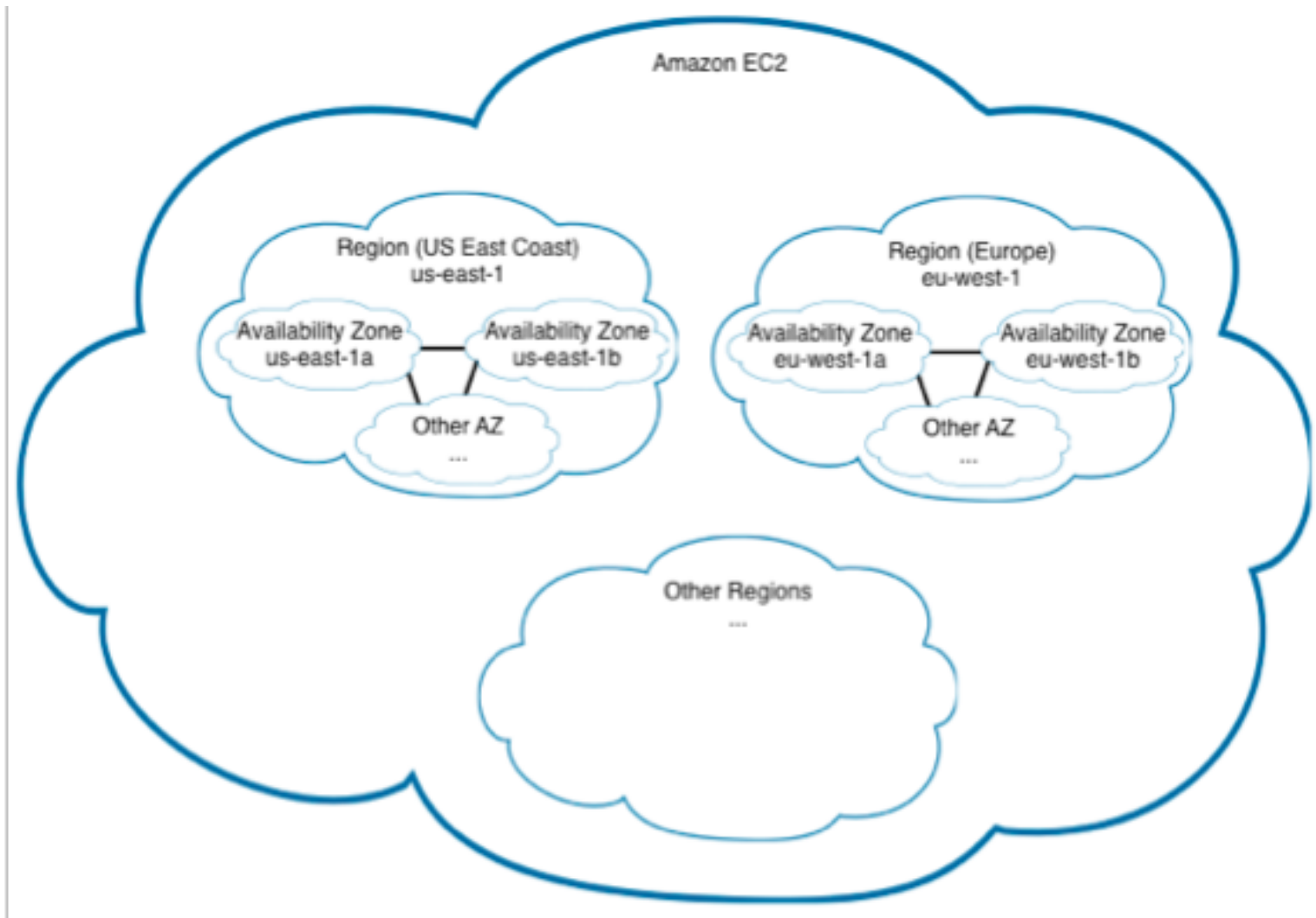
- Monitoring between 800 and 1000 servers
- Servers spread across 6 different locations
  - 4 Amazon EC2 Regions
  - 1 Hosted (Liquidweb) & 1 VPS (Linode)
- Little monitoring resources
  - Collecting over 120,000 metrics
  - Monitoring over 20,000 services with Nagios
- Multiple billions of HTTP requests a day
  - Most of it must be served in less than 100ms
  - Lost of traffic could mean lost of business opportunity
  - Or worst, over-spending...

# Our Environment

---

- Over 80 different server profiles
- Our stack:
  - Java (Embedded Jetty, Tomcat)
  - PHP, RoR
  - Hadoop: HDFS, M/R, Hbase, Hive
  - Couchbase
  - MySQL, Vertica
  - ElasticSearch
- Monitoring: Nagios, NSCA
- Graphing: Ganglia, sFlow, Graphite
- Configuration Management: Puppet

# Amazon Cloud Environment



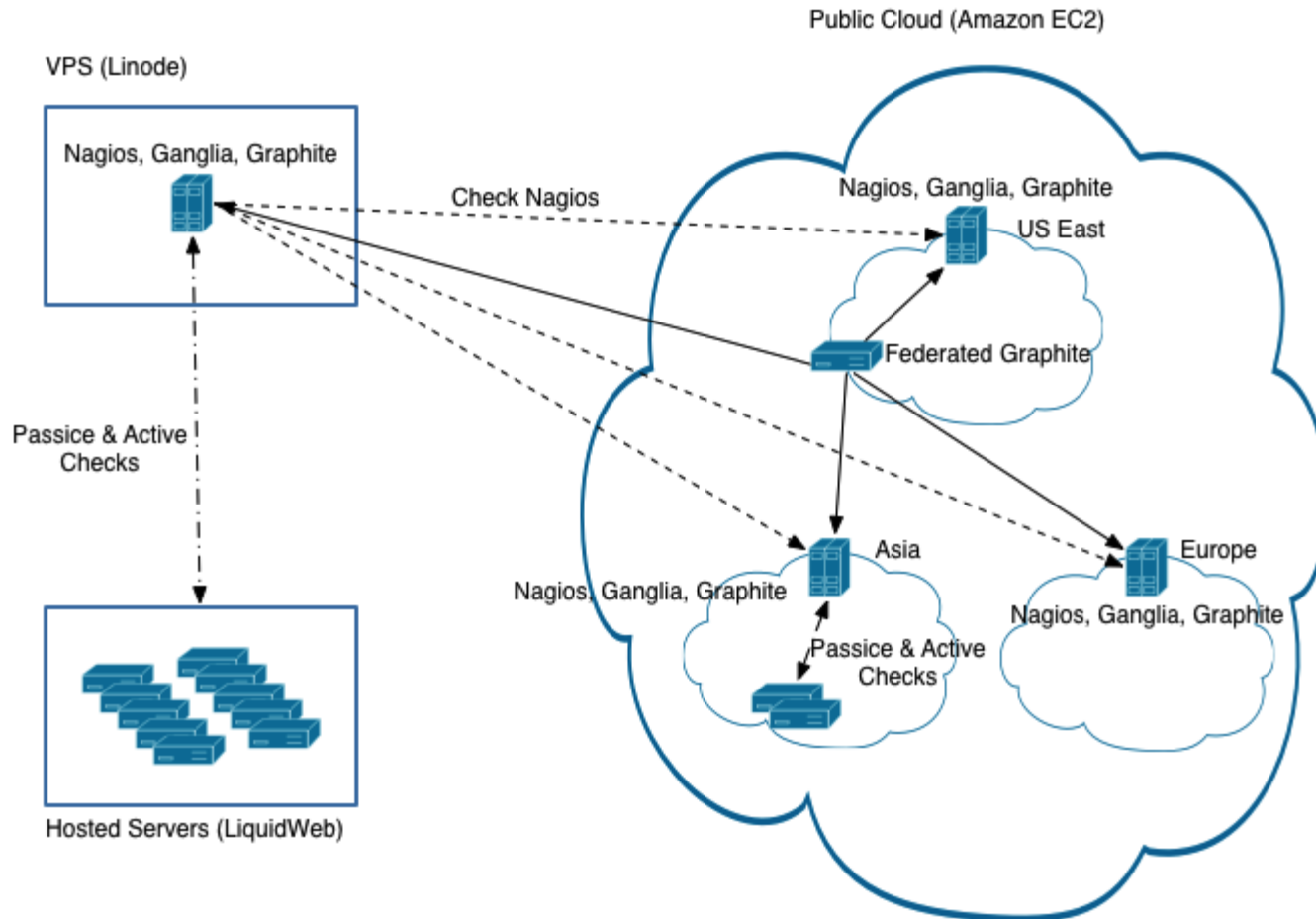
# Amazon Cloud Environment

- We use EC2, SDB, SQS, EMR, S3, etc.
- We don't use ELB
- We heavily use EC2 Tags

ec2-describe-instances -F tag:hostname=dev-build01

```
RESERVATION    r-891f48ec    290999691900    devzone
INSTANCE       i-1ed92766    ami-08f40561    ec2-██████████.compute-1.amazonaws.com    domU-██████████
t-1b    aki-427d952b    monitoring-disabled    ████████████████████████████████████████    in
BLOCKDEVICE    /dev/sde1    vol-5e736931    2012-07-02T08:51:07.000Z    false
TAG    instance    i-1ed92766    cluster    devzone
TAG    instance    i-1ed92766    hostname    dev-build01
TAG    instance    i-1ed92766    nagios_host    dev-mgmt01
TAG    instance    i-1ed92766    profile    DevBuildBox
```

# Automated Monitoring



# Automated Monitoring

---

## Process of event when starting a new host and add it to our monitoring:

1. We start a new instance using Cerveza and Cloud-init
2. Puppet configure Gmond or Host sFlow on the instance
3. Our monitoring server running Gmond and Gmetad get data from the new instance
4. A Nagios check run every minute and check for new hosts
  - Look for new hosts using EC2 API
  - Look for EC2 tag “hostname” to confirm it’s a legit host, not a zombie / fail start
  - Look for EC2 tag “nagios\_host” to see if the host belong to this monitoring instance
5. If a new host is found:
  - We build a config for the host based on a template file and doing some string replace
  - Once all config have been generated, we rebuild pre-cache objects and reload Nagios
6. If we find “Zombie” host, we generate a Warning alert
7. If the config is corrupt, we send a Critical alert



# Efficient Monitoring

We reduce noise by disabling most notifications and using our "cluster check"

```
define service{
    use                local-service
    hostgroup_name     ^[a-z0-9_-]+-cluster
    service_description PING
    servicegroups      network-status
    check_command       check_ping!100.0,20%!500.0,60%
    max_check_attempts 10
    notifications_enabled 0
}

define service{
    use                generic-service
    host_name          <%= hostname %>
    servicegroups      cluster-service
    service_description Cluster - PING
    check_command       check_cluster_service!^.+!^PING$$
    contact_groups     noc
}
```

# Efficient Monitoring

```
Usage: check_nagios_status [options]
```

```
Options:
```

```
-h, --help                show this help message and exit
-v, --verbose             Verbose logging. (default: False)
--status-file=STATUS_FILE
                          Path to the Nagios status file. (default:
                          /opt/nagios/var/status.dat)
--host-regex=HOST_REGEX  Regex used to filter host name.
--service-regex=SERVICE_REGEX
                          Regex used to filter service description. (default:
                          none)
-w WARNING, --warning=WARNING
                          Warning threshold in percent. (default: 30)
-c CRITICAL, --critical=CRITICAL
                          Critical threshold in percent. (default: 60)
-u UNKNOWN, --unknown=UNKNOWN
                          Unknown threshold in percent. (default: none)
```

# Efficient Monitoring

```
Usage: check_nagios_status_msg [options]
```

## Options:

```
-h, --help          show this help message and exit
-v, --verbose       Verbose logging. (default: False)
--status-file=STATUS_FILE
                    Path to the Nagios status file. (default:
                    /opt/nagios/var/status.dat)
--host-regex=HOST_REGEX
                    Regex used to filter host name.
--service-regex=SERVICE_REGEX
                    Regex used to filter service description. (default:
                    none)
-w WARNING, --warning=WARNING
                    Warning threshold in percent. (default: 30)
-c CRITICAL, --critical=CRITICAL
                    Critical threshold in percent. (default: 60)
--msg-filter=MSG_FILTER
                    Regex used to filter plugin output and mark it as
                    error.
```

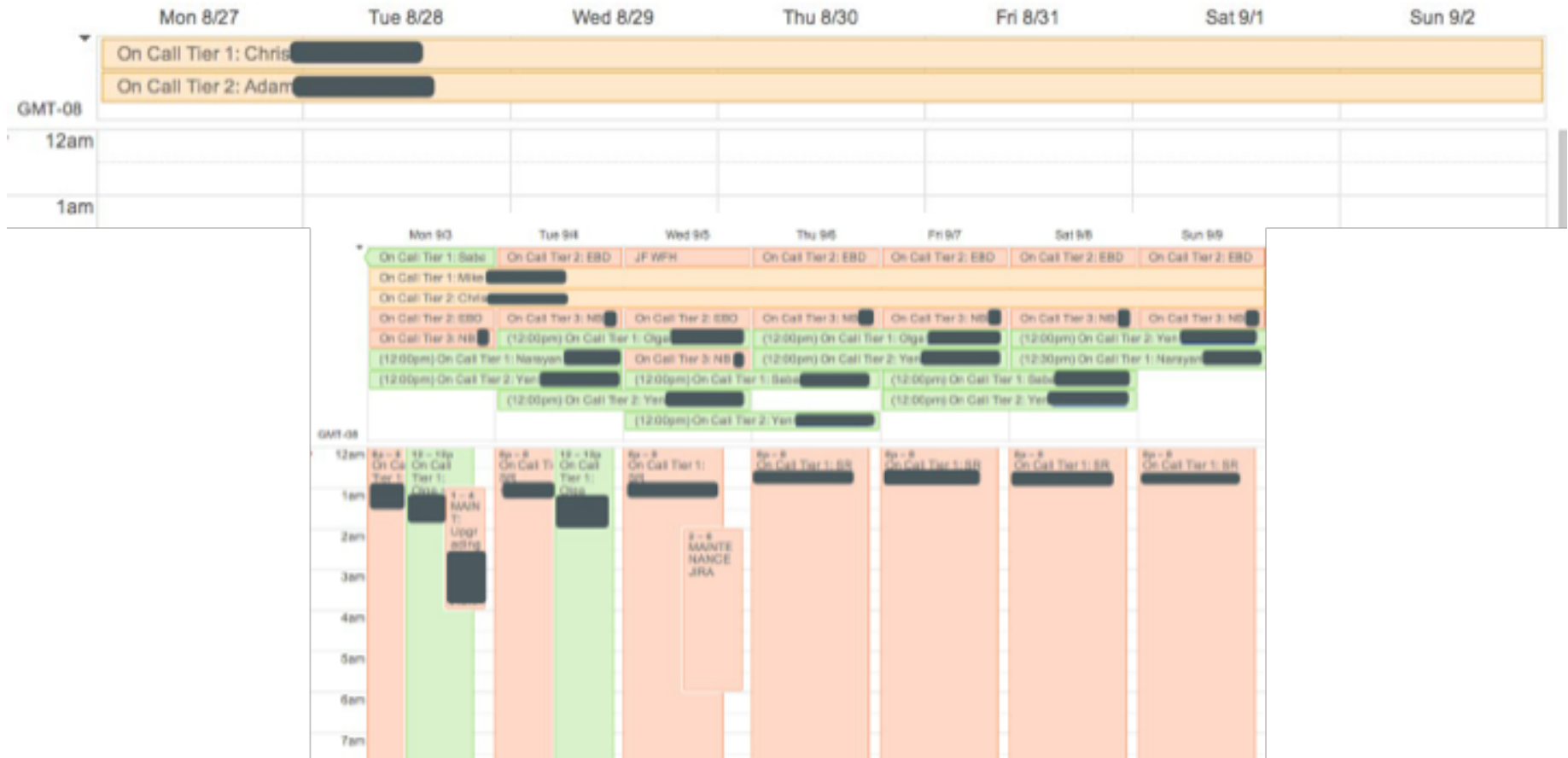
# Efficient on-call rotation

---

- Follow the sun
  - OPS team is in Ukraine, no more Tier 1 night on-call for US Team
- Timeperiod and escalation are a pain to maintain
  - Nagios notification plugged to Google Calendar
    - Using our own notification script for email and paging
    - Google Calendar make it easy for each team to manage their own on-call calendar
    - Support for multiple Tier and complex schedules
    - Caching Google Calendar info locally every hour
  - Simpler definitions and rules in Nagios contacts
  - Notify only people on-call, unless they asked for “off call” emails

# Efficient on-call rotation

## Using Google Calendar...



# Efficient on-call rotation

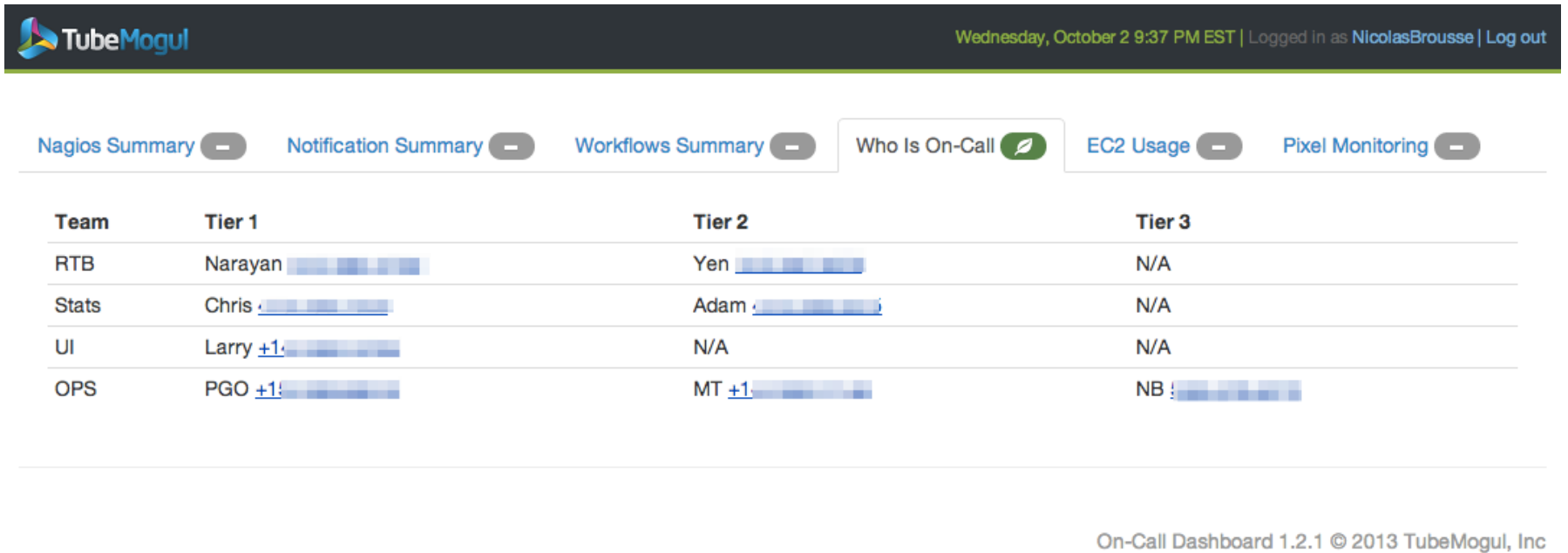
- Simple contact definitions
- Google Calendar info
- Tier Filter (Regex)
- Tier Interval (time to wait before escalating alert since last tier)
- Off call email

```
define contact{
    contact_name      nicolas
    use                pager-contact
    alias             Nicolas Brousse
    email             nicolas@tubemogul.com
    pager             [REDACTED]
    _GOOGLE_CALENDAR_ID      "tubemogul.com_[REDACTED]@group.calendar.google.com"
    _GOOGLE_CALENDAR_TIER_FILTER      "NB"
    _GOOGLE_CALENDAR_TIER_INTERVAL    "TIER3_INTERVAL=30m"
    _OFF_CALL_EMAIL                  yes
}
```



# Efficient on-call rotation

- Centralized dashboard!
  - Plugged to Google Apps
  - List all on-call contacts



Wednesday, October 2 9:37 PM EST | Logged in as NicolasBrousse | Log out

Nagios Summary Notification Summary Workflows Summary Who Is On-Call EC2 Usage Pixel Monitoring

Team	Tier 1	Tier 2	Tier 3
RTB	Narayan	Yen	N/A
Stats	Chris	Adam	N/A
UI	Larry +1:	N/A	N/A
OPS	PGO +1:	MT +1:	NB :

On-Call Dashboard 1.2.1 © 2013 TubeMogul, Inc

# Efficient on-call rotation

- Centralized View of Multiple Nagios!



Wednesday, October 2 9:40 PM EST | Logged in as NicolasBrousse | Log out

Nagios Summary

Notification Summary

Workflows Summary

Who Is On-Call

EC2 Usage

Pixel Monitoring

Contact group:  
Contact list is empty.

Nagios	Paging	Critical	Warning	Unknown	Last Update
.US-EAST-1A					Wed 02 Oct 21:36:23
.AP-SOUTHEAST-1A					Wed 02 Oct 21:36:23
.US-EAST-1B					Wed 02 Oct 21:36:02
.LINODE					Wed 02 Oct 21:36:09
.US-WEST-1A					Wed 02 Oct 21:36:16
.US-EAST-1B					Wed 02 Oct 21:36:19
DEV- .US-EAST-1B					Wed 02 Oct 21:36:07
.EU-WEST-1A					Wed 02 Oct 21:36:19

On-Call Dashboard 1.2.1 © 2013 TubeMogul, Inc





# Efficient on-call rotation

## 7 DEV-MGMT01.US-EAST-1B

Host	Service	Last Check	Duration	Attempt	Status Information
dev-rtb-adserver01	rtb-adserver-sa-presence	10-02-2013 21:37:16	56d 3h 20m 46s	3/3	no node

[Nagios Summary](#)

[Notification Summary](#)

[Workflows Summary](#)

[Who Is On-Call](#)

[EC2 Usage](#)

[Pixel Monitoring](#)

Filter by date:

Enter

Nagios	Receiver	Contact	Host	Service	Date
████████.US-EAST-1A	Tier 1: PGO	████████	mgmt01	Cluster - Bidder HTTP response per sec.	Oct. 2, 2013, 9:11 p.m.
████████.AP-SOUTHEAST-1A	Tier 1: Narayan	████████	mgmt01	Cluster - Bidder Throttler	Oct. 2, 2013, 8:56 p.m.

# Efficient on-call rotation

## 11 DEV-██████████.US-EAST-1B

Action	TT	Date	Host
Ok▼	N/A	10-02-2013 23:09:47	dev-i
Ok▼	N/A	10-02-2013 23:08:12	dev-i
Ok▼	N/A	10-02-2013 23:11:09	dev-i
Ok▼	N/A	10-02-2013 23:09:47	dev-i
Ok▼	N/A	10-02-2013 23:13:54	dev-i
Ok▼	N/A	10-02-2013 23:08:13	dev-i
Ok▼	N/A	10-02-2013 23:16:59	dev-i
Ok▼	N/A	10-02-2013 23:15:54	dev-i
Warning▼	N/A	10-02-2013 22:37:19	dev-i
Create Ticket		10-02-2013 23:13:51	dev-i
Close Ticket		10-02-2013 23:16:45	dev-i
Acknowledge			

# Business Awareness!

---

Now that I know what is  
breaking...

Which one should I fix first?

# Business Awareness!

## Service Health Dashboard!



You are Admin | [Jira Ticket](#) | Your TimeZone is [America/Los\\_Angeles](#) | [AleksyMykhalov](#) | [Logout](#)

APAC Europe US East Coast US West Coast Global

Action	Status	Service	Ticket	Details	Out of SLA
		OneLoad Harvester	N/A	Service is Operating Normally	None
		OneLoad Transcoding	N/A	Service is Operating Normally	None
		Playtime Transcoding	N/A	Service is Operating Normally	None
		Playtime Analytics	N/A	Service is Operating Normally	None
		Playtime User Segments	N/A	Service is Operating Normally	None
		Playtime User Targeting and Frequency Capping	N/A	Unknown state	None
		OneLoad Website (oneload.com)	N/A	Service is Operating Normally	None
		TubeMogul Website (tubemogul.com)	N/A	Service is Operating Normally	None
		Playtime Partners Report	N/A	Service is Operating Normally	None

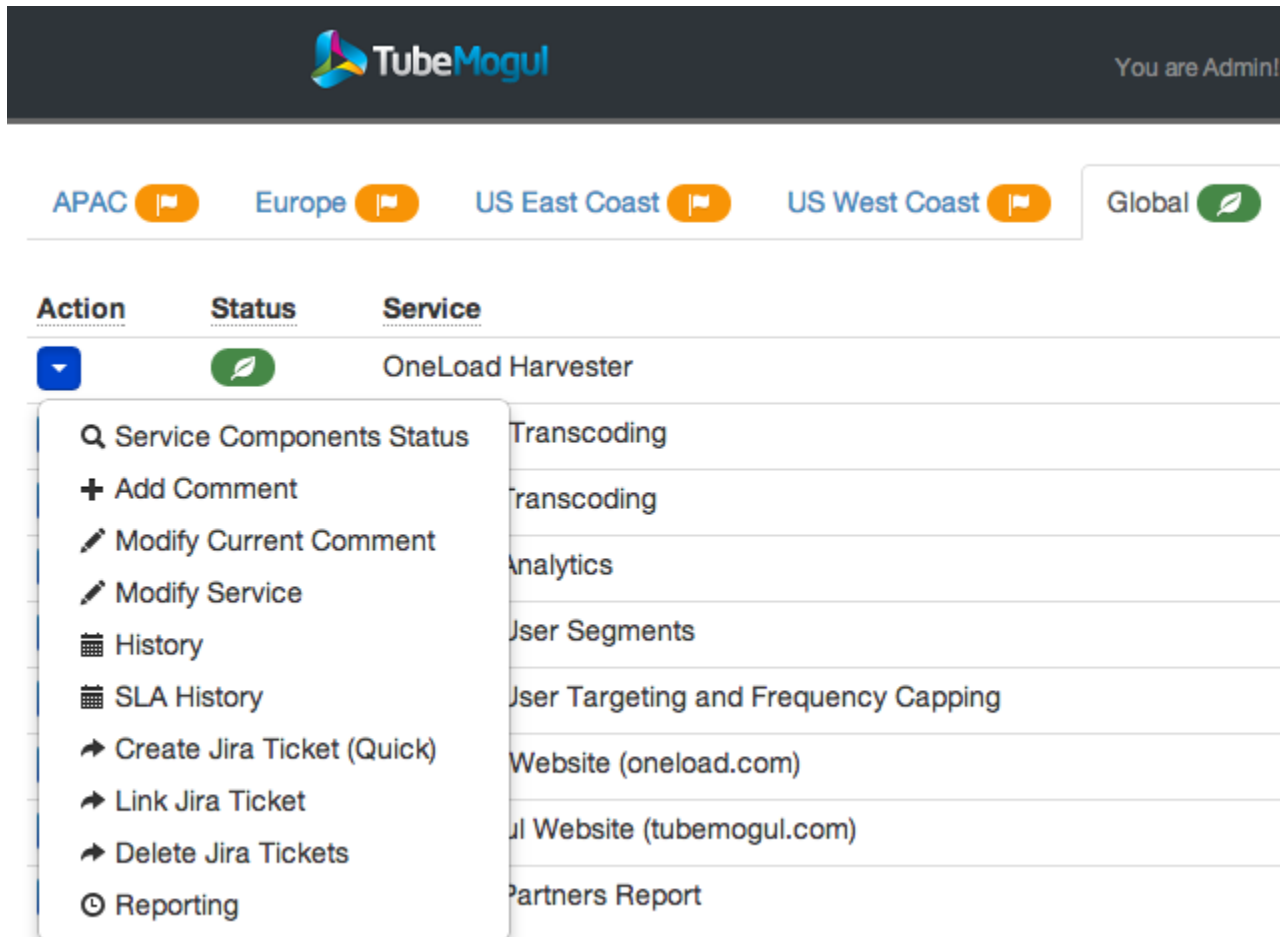
Copyright 2013 TubeMogul, Inc

# Business Awareness!

---

- A service represent a Business Critical function
- A service can be global or limited to a region
- A service is defined by multiple Service Component
  - We use Nagios Event Handler to update services component status
  - REST API dashboard allow easy update from third party monitoring, QA test, scripts, crons, Nagios
- We can define service SLA and quickly see SLA breaking (based on OLA)
- OPS Team can perform more actions, post comments, link to Jira tickets

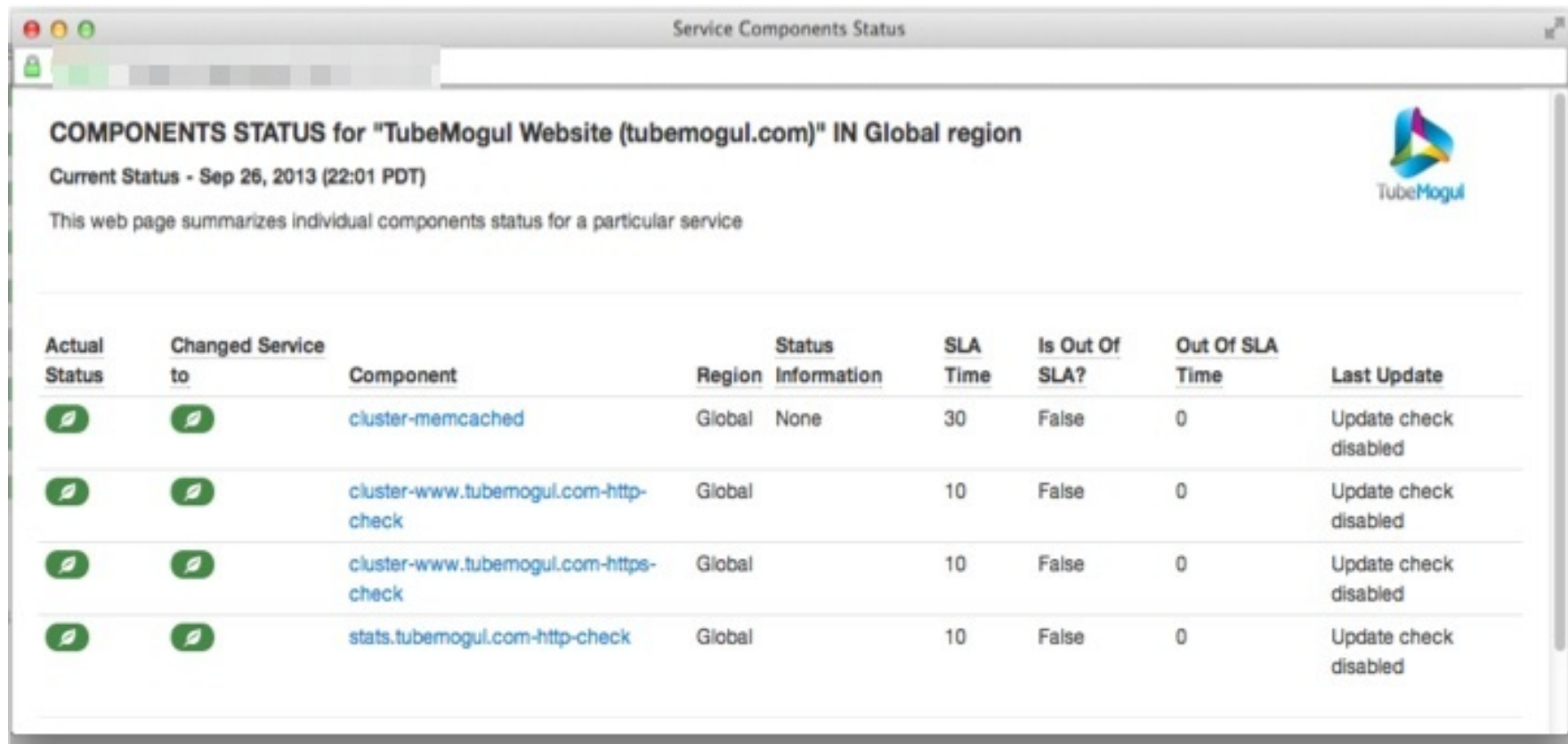
# Business Awareness!











The screenshot shows the TubeMogul admin interface. At the top, the TubeMogul logo is on the left and 'You are Admin!' is on the right. Below the header, there are navigation tabs for different regions: APAC, Europe, US East Coast, US West Coast, and Global. The Global tab is selected. Below the tabs is a table with columns for Action, Status, and Service. The first row shows a dropdown arrow in the Action column, a green leaf icon in the Status column, and 'OneLoad Harvester' in the Service column. A dropdown menu is open from the Action column, listing various actions: Service Components Status, Add Comment, Modify Current Comment, Modify Service, History, SLA History, Create Jira Ticket (Quick), Link Jira Ticket, Delete Jira Tickets, and Reporting. The background table lists several services with their respective status icons.

Action	Status	Service
▼	🌿	OneLoad Harvester
🔍 Service Components Status		Transcoding
+ Add Comment		Transcoding
✎ Modify Current Comment		Analytics
✎ Modify Service		User Segments
📅 History		User Targeting and Frequency Capping
📅 SLA History		Website (onload.com)
➔ Create Jira Ticket (Quick)		ul Website (tubemogul.com)
➔ Link Jira Ticket		
➔ Delete Jira Tickets		
🕒 Reporting		Partners Report

# Business Awareness!




The screenshot shows a web browser window titled "Service Components Status". The main heading is "COMPONENTS STATUS for 'TubeMogul Website (tubemogul.com)' IN Global region". Below this, it says "Current Status - Sep 26, 2013 (22:01 PDT)" and "This web page summarizes individual components status for a particular service". The TubeMogul logo is in the top right. A table lists the status of four components, all of which are green, indicating they are healthy.






Actual Status	Changed Service to	Component	Region	Status Information	SLA Time	Is Out Of SLA?	Out Of SLA Time	Last Update
		<a href="#">cluster-memcached</a>	Global	None	30	False	0	Update check disabled
		<a href="#">cluster-www.tubemogul.com-http-check</a>	Global		10	False	0	Update check disabled
		<a href="#">cluster-www.tubemogul.com-https-check</a>	Global		10	False	0	Update check disabled
		<a href="#">stats.tubemogul.com-http-check</a>	Global		10	False	0	Update check disabled

# Business Awareness!

History

History for Playtime User Segments in Global region



Status	Status Change Date	Affected By Component	Component's Region
	Sept. 9, 2013, 7:37 a.m.	userdb-maestro-fullUserDb	Global
	Sept. 9, 2013, 1:47 a.m.	userdb-maestro-fullUserDb	Global
	Sept. 9, 2013, 1:17 a.m.	userdb-maestro-fullUserDb	Global
	Sept. 7, 2013, 8 p.m.	userdb-maestro-externalSegment	Global
	Sept. 7, 2013, 7 p.m.	userdb-maestro-externalSegment	Global


← Previous Page 14 of 159. Next →



# Business Awareness!

SLA History

Out Of SLA History for Playtime User Segments in Global region



<u>Date</u>	<u>Affected By Component</u>	<u>Time (Min)</u>
Sept. 16, 2013, 5:16 p.m.	userdb-maestro-externalSegment (Region: Global)	50
Sept. 16, 2013, 5:09 p.m.	userdb-maestro-fullUserDb (Region: Global)	100

← Previous Page 2 of 4. Next →

Copyright 2013 TubeMogul, Inc

# Business Awareness!

The image shows a screenshot of the TubeMogul web interface. At the top, there are regional filters for APAC, Europe, and US East Coast. Below these is a table with columns for Action, Status, and Service. The 'Service' column shows 'RTB'. A dropdown menu is open under the 'Action' column, listing various options including 'Create Jira Ticket (Quick)'. Overlaid on this is a 'Submit Ticket to Jira' dialog box. The dialog has a 'CREATE TICKET' header and two main sections: 'Ticket Summary' and 'Description'. The 'Ticket Summary' field contains the text 'Issues with RTB in region US We'. The 'Description' field contains a warning message: 'Caused by component(s): Name: cluster-adpop-stats-api, Region: US West Coast, Output: Warning hosts: rtb-adpop01\nadpop-stats-api: WARNING: Freshness Threshold exceed. No recent update received by the service.\n\n'. Below the description field is a note: 'Information from this area will be available to the Service Health Dashboard end user.' and a 'Submit' button.

Action	Status	Service
		RTB

- Service Components Status
- Add Comment
- Modify Current Comment
- Modify Service
- History
- SLA History
- Create Jira Ticket (Quick)
- Link Jira Ticket
- Delete Jira Tickets
- Reporting

**CREATE TICKET**

Ticket Summary:  
Issues with RTB in region US We

Description:  
Caused by component(s):  
Name: cluster-adpop-stats-api,  
Region: US West Coast,  
Output: Warning hosts: rtb-adpop01\nadpop-stats-api: WARNING:  
Freshness Threshold exceed. No recent update received by the  
service.\n\n

Information from this area will be available to the Service Health Dashboard end user.

Submit

# Business Awareness!

The screenshot displays the TubeMogul web interface. On the left, there are navigation tabs for 'APAC', 'Europe', and 'US East'. Below these is a table with columns 'Action', 'Status', and 'Service'. The 'Status' column contains icons: a blue play button, an orange play button, and a grey minus sign. The 'Service' column lists 'RTB', 'Playtime Ad Se', and 'Playtime User T'. The main content area shows a 'Create Report for RTB' form with a dropdown menu set to 'State History', date range inputs for '2013-01-01' and '2013-10-03', an 'Export to CSV?' checkbox, and a 'Submit' button. Below the form is a 'History' table with columns for time, service name, and location.

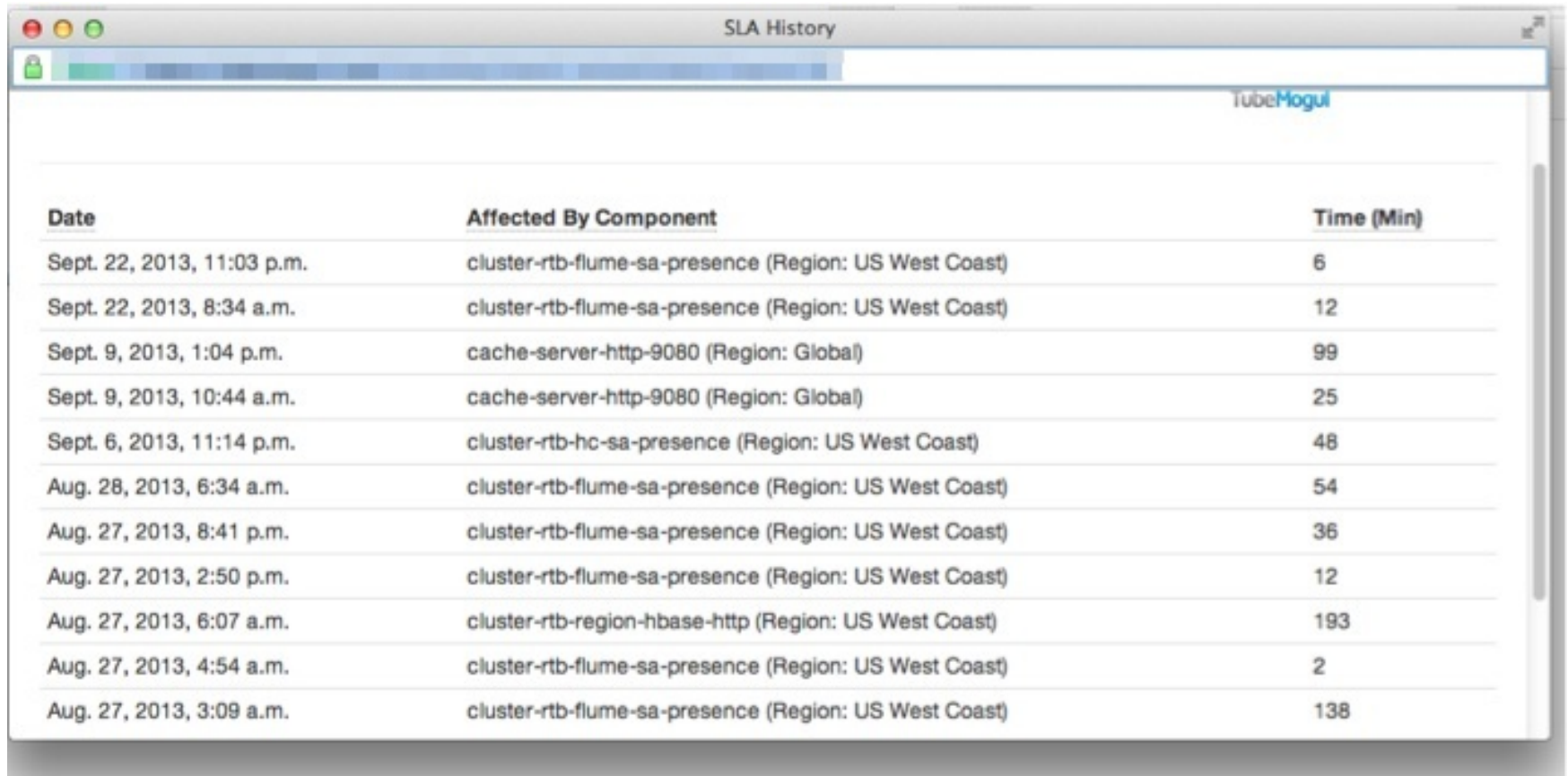
Action	Status	Service
		RTB
		Playtime Ad Se
		Playtime User T

Time	Service	Location
Sept. 9, 2013, 1:04 p.m.	cluster-adpop-stats-api	US West Coast
Sept. 9, 2013, 10:55 a.m.	cache-server-http-9080	Global
Sept. 9, 2013, 10:44 a.m.	cluster-adpop-stats-api	US West Coast
Sept. 9, 2013, 9:49 a.m.	cache-server-http-9080	Global
Sept. 9, 2013, 9:31 a.m.	cluster-adpop-stats-api	US West Coast
Sept. 9, 2013, 9:31 a.m.	cache-server-http-9080	Global
Sept. 8, 2013, 10:34 a.m.	cluster-adpop-stats-api	US West Coast
Sept. 8, 2013, 10:33 a.m.	cache-server-http-9080	Global
Sept. 7, 2013, 9:35 p.m.	cluster-adpop-stats-api	US West Coast

Copyright 2013 TubeMogul, Inc

# Business Awareness!



Date	Affected By Component	Time (Min)
Sept. 22, 2013, 11:03 p.m.	cluster-rtb-flume-sa-presence (Region: US West Coast)	6
Sept. 22, 2013, 8:34 a.m.	cluster-rtb-flume-sa-presence (Region: US West Coast)	12
Sept. 9, 2013, 1:04 p.m.	cache-server-http-9080 (Region: Global)	99
Sept. 9, 2013, 10:44 a.m.	cache-server-http-9080 (Region: Global)	25
Sept. 6, 2013, 11:14 p.m.	cluster-rtb-hc-sa-presence (Region: US West Coast)	48
Aug. 28, 2013, 6:34 a.m.	cluster-rtb-flume-sa-presence (Region: US West Coast)	54
Aug. 27, 2013, 8:41 p.m.	cluster-rtb-flume-sa-presence (Region: US West Coast)	36
Aug. 27, 2013, 2:50 p.m.	cluster-rtb-flume-sa-presence (Region: US West Coast)	12
Aug. 27, 2013, 6:07 a.m.	cluster-rtb-region-hbase-http (Region: US West Coast)	193
Aug. 27, 2013, 4:54 a.m.	cluster-rtb-flume-sa-presence (Region: US West Coast)	2
Aug. 27, 2013, 3:09 a.m.	cluster-rtb-flume-sa-presence (Region: US West Coast)	138

# Business Awareness!

---

How do we make sure we answer the business needs?

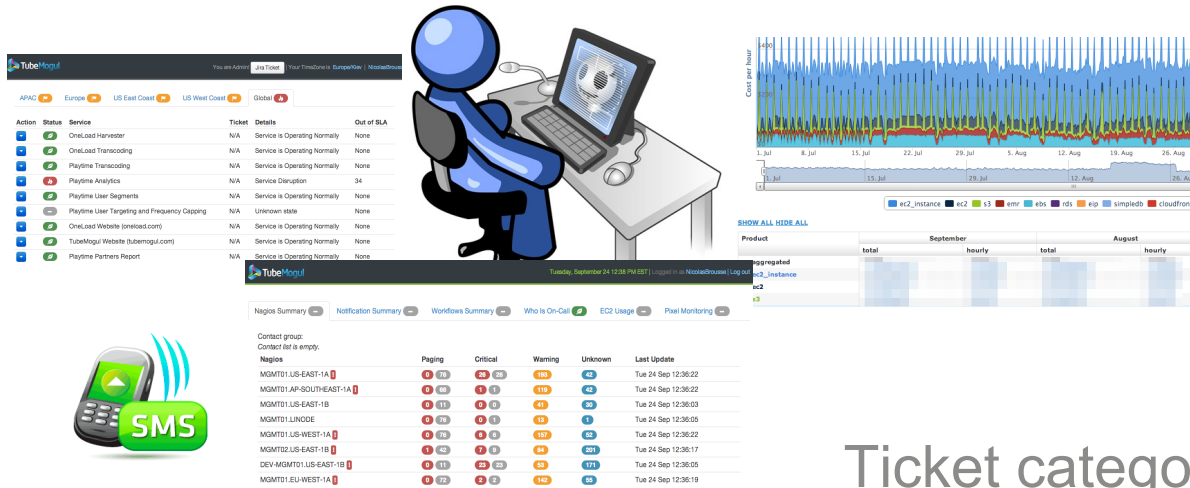
- Review SLAs and monitoring configurations monthly/quarterly
- Have a checklist when launching new product or features
  - We now have a SRE Hand-off Checklist with detailed questions
  - How is capacity planning done?
  - Who can have an impact on traffic, storage, etc?
  - Make sure to ask about expected OLAs or SLAs

# To summarize...

---

- We easily automate our deployment across multiple geos
- We control the noise
- We easily schedule our on-call rotations
- In one location we know everything
- We know what is impacting the business and how we should prioritize our actions

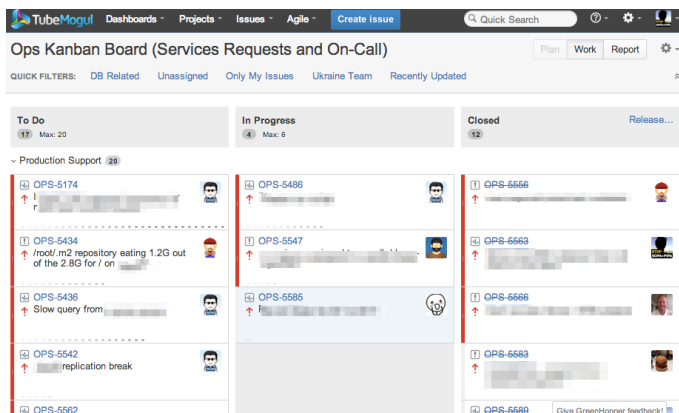
# Team Process - Daily Kanban!



Request based on Dashboards, Monitoring, Paging or Engineers

Ticket categorized in two Swimlane:

- Production Support
  - High Priority: Top to Bottom
  - On-call 24/7, OLAs, SLAs...
  - Incident are handled 1st
  - Maintenance are handled 2nd
- Developer Support
  - Best Effort: Top to Bottom
  - Long effort request moved to Infrastructure pipeline



# Team Process - Long Term Agile!

TubeMogul SRE Hand-off Checklist

## SRE Hand-off Checklist

This SRE Hand-off checklist has been developed to help developers to define their operational needs before committing to release deadlines. This checklist will also help the SRE team to properly schedule for changes and releases with a minimal back and forth. This doesn't replace our Release Process. Don't swim alone!

Please, use this google doc as a template / helper for your project. Make a copy and edit it as your project evolves. Do not use our copy of the template as we will keep updating this checklist to best reflect the engineering needs. Make sure to share your edited document with [Google Docs](#) or email us a pdf copy.

### Table of Content

- About your Project/Task
- Infrastructure Dependencies
- Accessing Resources and Security Concerns
- Capacity Planning and Resources Provisioning
- Backup and Disaster Recovery
- Quantifying Quality of Service

### Useful links

- Ops Support wiki
- How to get Ops support?
- How to request a Nagios Passive check?
- Onboarding and Training
- Permissions Requests
- Tools we use
- Security Expectations
- Release Checklists

### SRE Lead on this project

This will be filled up by the SRE Team. The SRE Lead is in charge of keeping the project moving and documentation up to date.

### SRE project wiki page

This will be filled up by the SRE Team. Main wiki page is [here](#).

### SRE Standard Operating Procedure wiki page

This will be filled up by the SRE Team. Main SOP wiki page is [here](#).

Confidential and Proprietary. This document is the property of TubeMogul, Inc. For the exclusive use of the Company. This private and personal information contained herein is proprietary business data, which can only be disclosed for limited business purposes. Any reproduction or use without permission is strictly prohibited.

© 2013 TubeMogul Incorporated. All Rights Reserved.



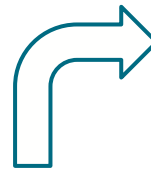
## Request moved from OPS to INF

## SRE Hand-off Checklist used to define Epic, Story and Tasks



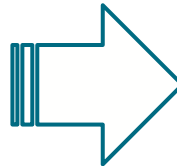
## Plan

## Check



## Act

## Do



Infrastructure Planning

QUICK FILTERS: Bugs Stories and Improvements Needs Requirements Assigned In Flight Blocked Goals UA Only My Issues Recently Updated

EPICS

All issues

Backlog (24)

Create Sprint

Plan Mode

Manage your backlog

- Create and prioritise stories in your backlog
- Estimate in story points or time

Plan your product development with epics or versions

- Use epics to track related stories across sprints and versions
- Assign stories to versions; track release progress
- Quickly drag-and-drop stories to epics or versions
- Filter by epics and versions for targeted view of your backlog

Create and launch development sprints

- Plan multiple sprints in advance
- Start a sprint, then go to Work mode to manage stories through workflow

Learn more about Plan mode and Work mode

Infrastructure Development

QUICK FILTERS: UMS OLTP Improvement Cloud Quality DB Unassigned Blocked Goals Bugs Stories and Improvements My Issues My QA UA

Recently Updated

To Do (2)

In Progress (16)

QA (11)

Awaiting Deployment (0)

Closed (2)

Release...



**All this wouldn't be possible without a strong  
SRE Operation Engineering team:**

Aleksey Mykhailov

Marylene Tanfin

Mykola Mogylenko

Nicolas Brousse

Pierre Gohon

Pierre Grandin

Stan Rudenko

# Thank You.

www.tubemogul.com | @tubemogul

We are Hiring !

<http://www.tubemogul.com/jobs>

