World of Tech WOT | 51CTO传媒

WOT 2015 互联网运维与开发者大会

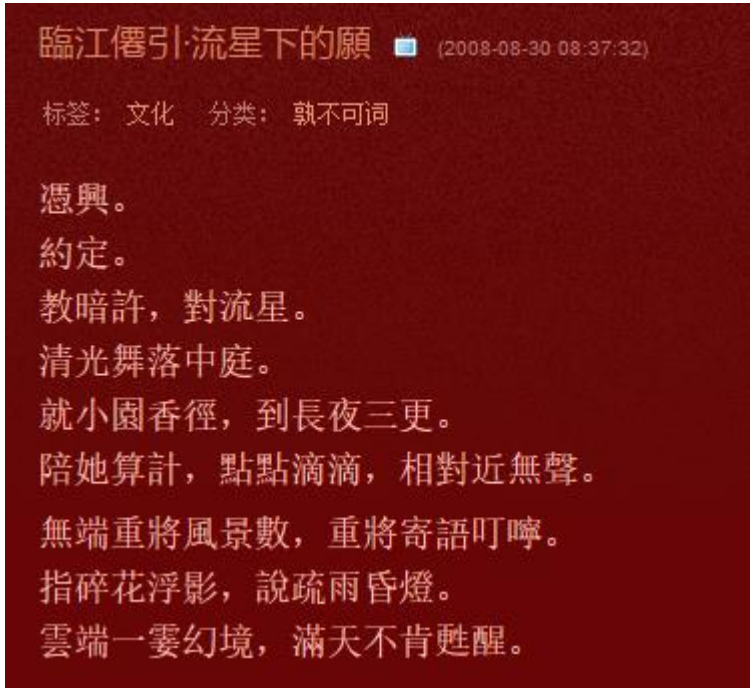
■ 2015年04月10日-11日 ■ 北京珠三角JW万豪酒店

实时数据驱动的应用运维

饶琛琳

我是？

- 给 CloudEx, China.com, RenRen, Sina 重启服务器；
- 精通 echo/say/puts/console.log(“Hello World”);
- 写过《网站运维技术与实践》；
- 翻过《Puppet 实战手册》；
- 其实，我是个网络诗人。



臨江僊引·流星下的願 (2008-08-30 08:37:32)

标签： 文化 分类： 孰不可词

憑興。
約定。
教暗許，對流星。
清光舞落中庭。
就小園香徑，到長夜三更。
陪她算計，點點滴滴，相對近無聲。
無端重將風景數，重將寄語叮嚀。
指碎花浮影，說疏雨昏燈。
雲端一霎幻境，滿天不肯甦醒。

运维是？

开发是？

“When I see ... developers – all I see is risk” Chuck Rossi, Facebook

才不是……

为什么做不好运维

- 数据不一致

为什么做不好运维

- 数据不一致
- 数据不及时

为什么做不好运维

- 数据不一致
- 数据不及时
- 数据不清晰

怎么做好运维

- 数据不一致 => Configuration Management
- 数据不及时 => Metric Monitor/Dashboard
- 数据不清晰 => Log Process/Analysis/Search

“Monitoring is the aggregation of health and performance data, events, and relationships delivered via an interface that provides an holistic view of a system's state to better understand and address failure scenarios.”

—— Ryan Frantz @Etsy

数据怎么来

- 不记你不打算用的日志；
- 次数是最感兴趣的；
- 在第一条的前提下尽可能收集所有日志；
- 没有7*24值班人员的话，IDS也就用不着实时。

——Laws of Marcus J. Ranum

大数据怎么来

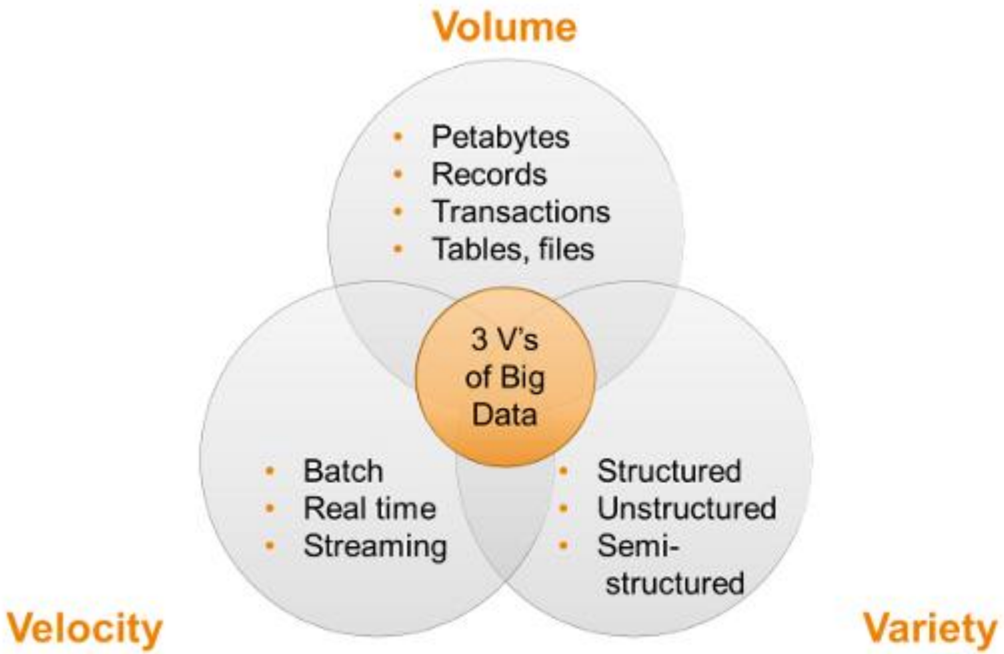
- 不记你不打算用的日志；
- 次数是最感兴趣的；
- 在第一条的前提下尽可能收集所有日志；
- 没有7*24值班人员的话，IDS也就用不着实时。

——Laws of Marcus J. Ranum

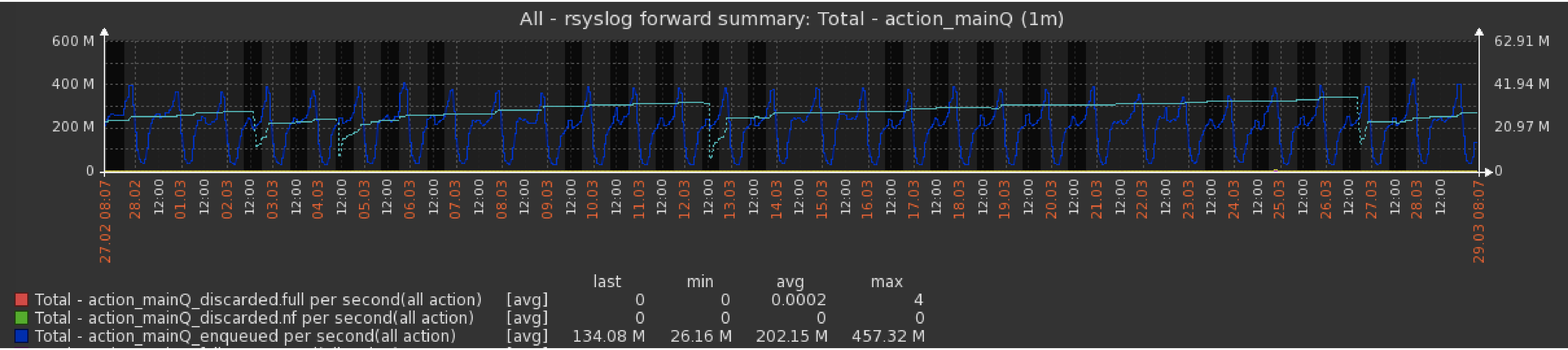
大数据是？

- Volume(越来越大)
- Velocity(越来越快)
- Variety(越来越复杂)

——Laney 2001.02



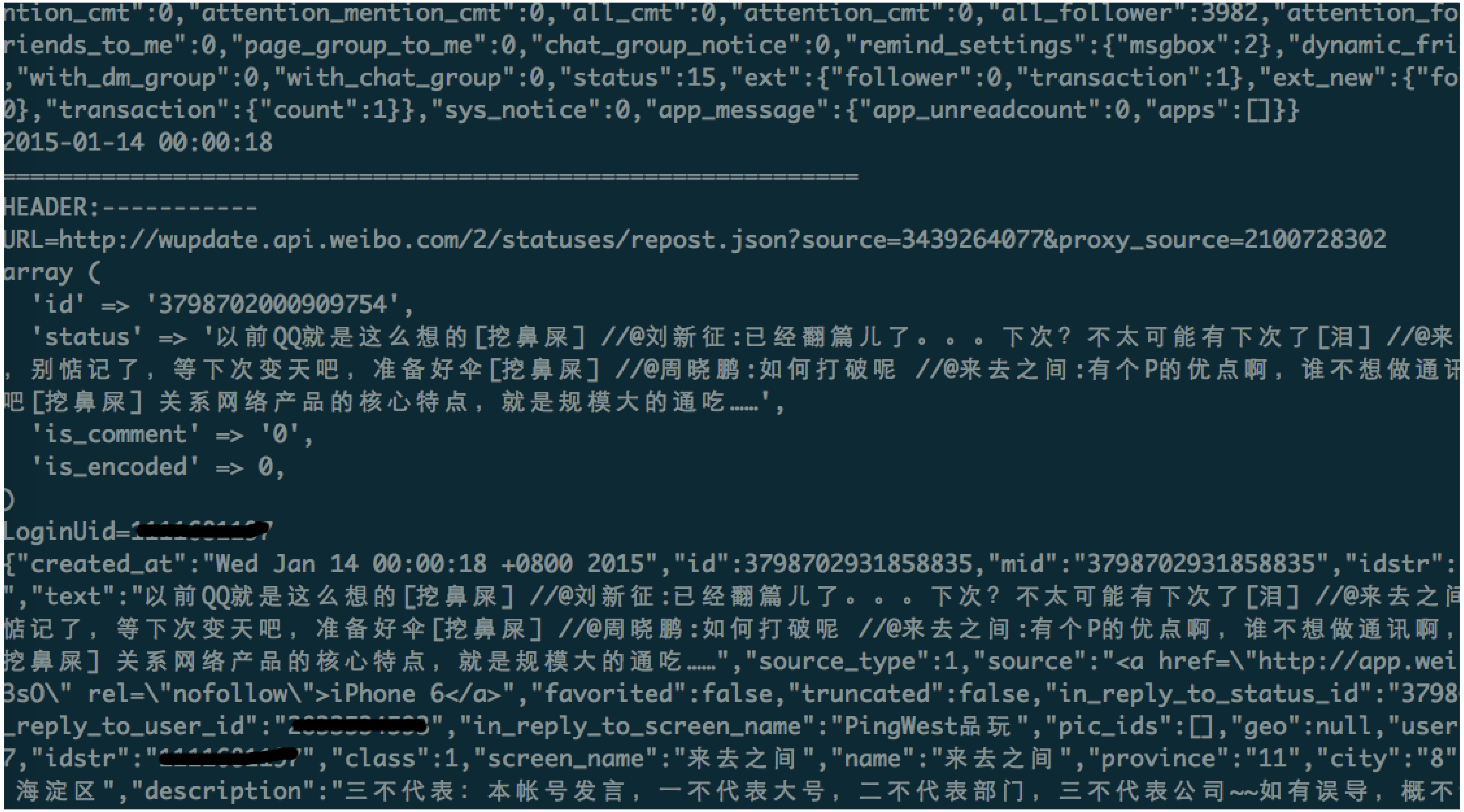
有多快?



All - rsyslog forward summary: Total - action_mainQ (1m)

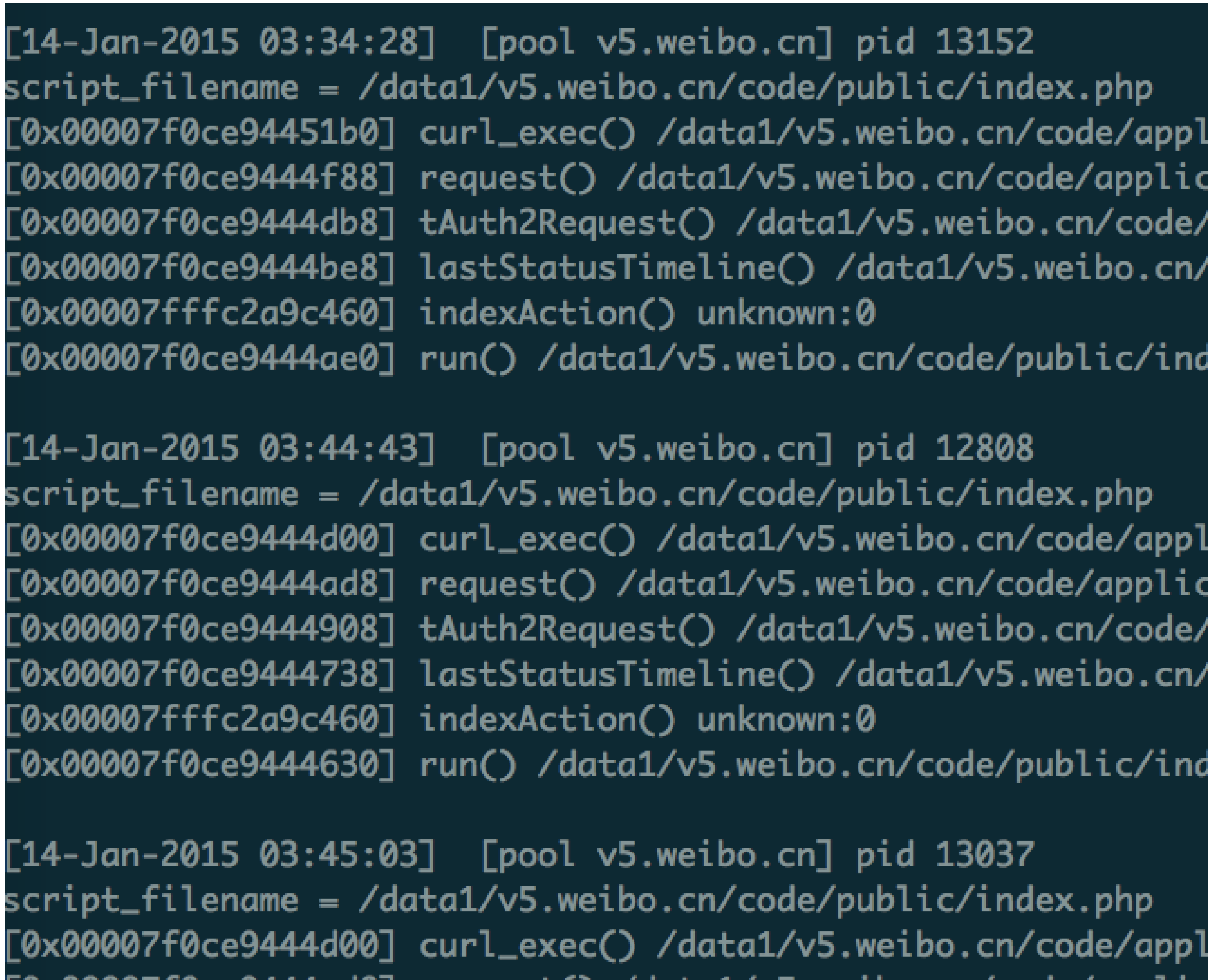
		last	min	avg	max
🟥 Total - action_mainQ_discarded.full per second(all action)	[avg]	0	0	0.0002	4
🟩 Total - action_mainQ_discarded.nf per second(all action)	[avg]	0	0	0	0
🟦 Total - action_mainQ_enqueued per second(all action)	[avg]	134.08 M	26.16 M	202.15 M	457.32 M

有多复杂？



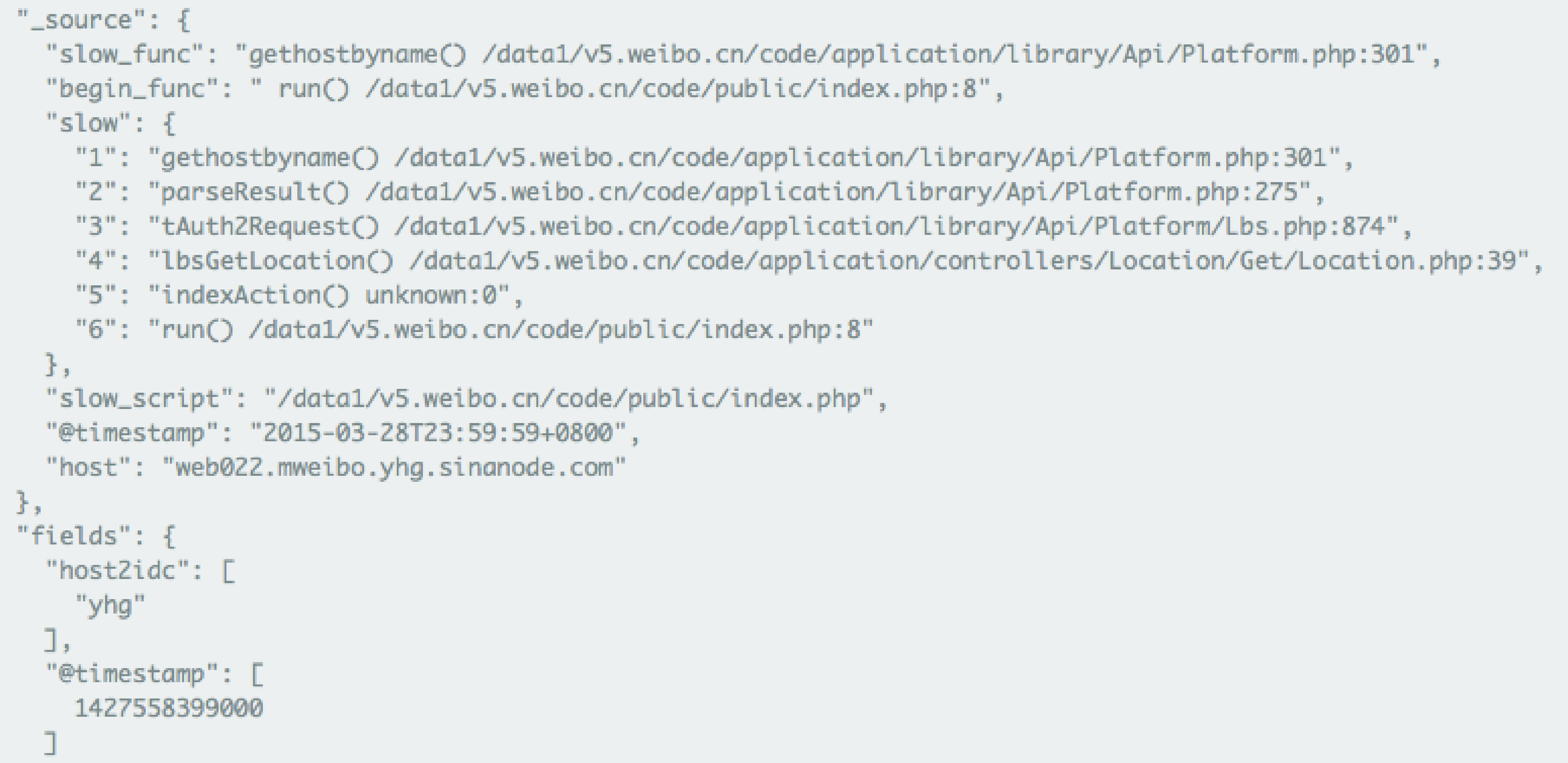
```
ntion_cmt":0,"attention_mention_cmt":0,"all_cmt":0,"attention_cmt":0,"all_follower":3982,"attention_fo
riends_to_me":0,"page_group_to_me":0,"chat_group_notice":0,"remind_settings":{"msgbox":2},"dynamic_fri
,"with_dm_group":0,"with_chat_group":0,"status":15,"ext":{"follower":0,"transaction":1},"ext_new":{"fo
0},"transaction":{"count":1}},"sys_notice":0,"app_message":{"app_unreadcount":0,"apps":[]}}
2015-01-14 00:00:18

===============================================================
HEADER:-----------
URL=http://wupdate.api.weibo.com/2/statuses/repost.json?source=3439264077&proxy_source=2100728302
array (
  'id' => '3798702000909754',
  'status' => '以前QQ就是这么想的[挖鼻屎] //@刘新征:已经翻篇儿了。。。下次？不太可能有下次了[泪] //@来
，别惦记了，等下次变天吧，准备好伞[挖鼻屎] //@周晓鹏:如何打破呢 //@来去之间:有个P的优点啊，谁不想做通
吧[挖鼻屎] 关系网络产品的核心特点，就是规模大的通吃……',
  'is_comment' => '0',
  'is_encoded' => 0,
)
LoginUid=[redacted]
{"created_at":"Wed Jan 14 00:00:18 +0800 2015","id":3798702931858835,"mid":"3798702931858835","idstr":
","text":"以前QQ就是这么想的[挖鼻屎] //@刘新征:已经翻篇儿了。。。下次？不太可能有下次了[泪] //@来去之
惦记了，等下次变天吧，准备好伞[挖鼻屎] //@周晓鹏:如何打破呢 //@来去之间:有个P的优点啊，谁不想做通讯啊，
挖鼻屎] 关系网络产品的核心特点，就是规模大的通吃……","source_type":1,"source":"<a href=\"http://app.wei
3sO\" rel=\"nofollow\">iPhone 6</a>","favorited":false,"truncated":false,"in_reply_to_status_id":"3798
_reply_to_user_id":"[redacted]","in_reply_to_screen_name":"PingWest品玩","pic_ids":[],"geo":null,"user
7,"idstr":"[redacted]","class":1,"screen_name":"来去之间","name":"来去之间","province":"11","city":"8"
 海淀区","description":"三不代表：本帐号发言，一不代表大号，二不代表部门，三不代表公司~~如有误导，概不
```

应该啥样？



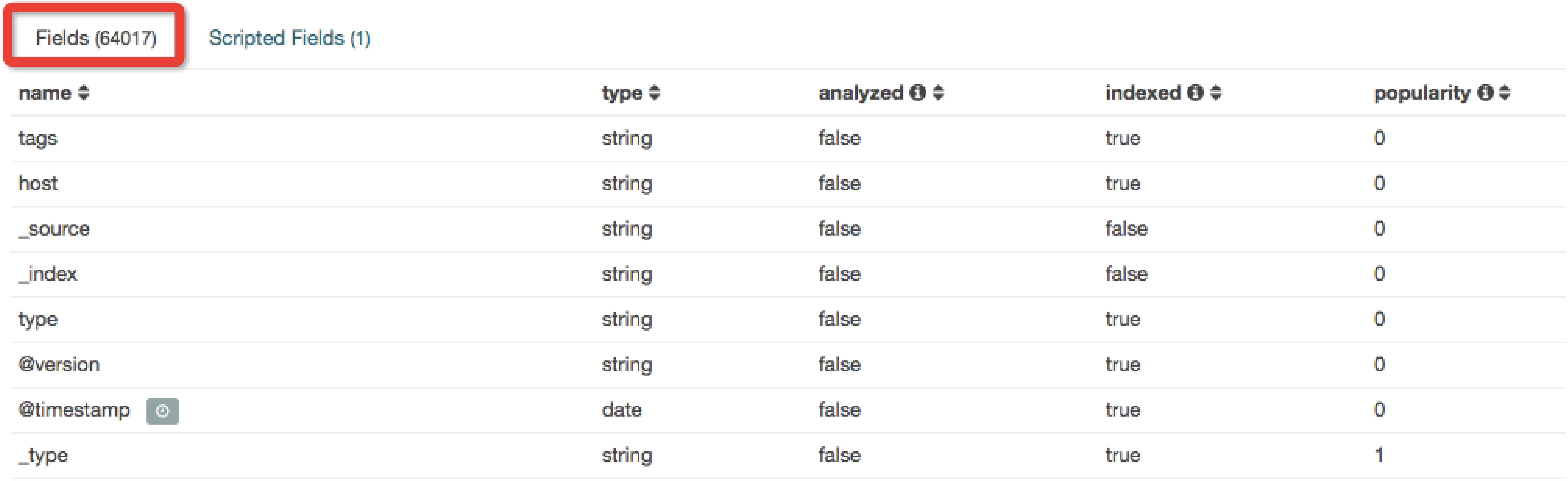
```
[14-Jan-2015 03:34:28]  [pool v5.weibo.cn] pid 13152
script_filename = /data1/v5.weibo.cn/code/public/index.php
[0x00007f0ce94451b0] curl_exec() /data1/v5.weibo.cn/code/appl
[0x00007f0ce9444f88] request() /data1/v5.weibo.cn/code/applic
[0x00007f0ce9444db8] tAuth2Request() /data1/v5.weibo.cn/code/
[0x00007f0ce9444be8] lastStatusTimeline() /data1/v5.weibo.cn/
[0x00007fffc2a9c460] indexAction() unknown:0
[0x00007f0ce9444ae0] run() /data1/v5.weibo.cn/code/public/ind

[14-Jan-2015 03:44:43]  [pool v5.weibo.cn] pid 12808
script_filename = /data1/v5.weibo.cn/code/public/index.php
[0x00007f0ce9444d00] curl_exec() /data1/v5.weibo.cn/code/appl
[0x00007f0ce9444ad8] request() /data1/v5.weibo.cn/code/applic
[0x00007f0ce9444908] tAuth2Request() /data1/v5.weibo.cn/code/
[0x00007f0ce9444738] lastStatusTimeline() /data1/v5.weibo.cn/
[0x00007fffc2a9c460] indexAction() unknown:0
[0x00007f0ce9444630] run() /data1/v5.weibo.cn/code/public/ind

[14-Jan-2015 03:45:03]  [pool v5.weibo.cn] pid 13037
script_filename = /data1/v5.weibo.cn/code/public/index.php
[0x00007f0ce9444d00] curl_exec() /data1/v5.weibo.cn/code/appl
```

应该啥样？



```
"_source": {
  "slow_func": "gethostbyname() /data1/v5.weibo.cn/code/application/library/Api/Platform.php:301",
  "begin_func": " run() /data1/v5.weibo.cn/code/public/index.php:8",
  "slow": {
    "1": "gethostbyname() /data1/v5.weibo.cn/code/application/library/Api/Platform.php:301",
    "2": "parseResult() /data1/v5.weibo.cn/code/application/library/Api/Platform.php:275",
    "3": "tAuth2Request() /data1/v5.weibo.cn/code/application/library/Api/Platform/Lbs.php:874",
    "4": "lbsGetLocation() /data1/v5.weibo.cn/code/application/controllers/Location/Get/Location.php:39",
    "5": "indexAction() unknown:0",
    "6": "run() /data1/v5.weibo.cn/code/public/index.php:8"
  },
  "slow_script": "/data1/v5.weibo.cn/code/public/index.php",
  "@timestamp": "2015-03-28T23:59:59+0800",
  "host": "web022.mweibo.yhg.sinanode.com"
},
"fields": {
  "host2idc": [
    "yhg"
  ],
  "@timestamp": [
    1427558399000
  ]
```

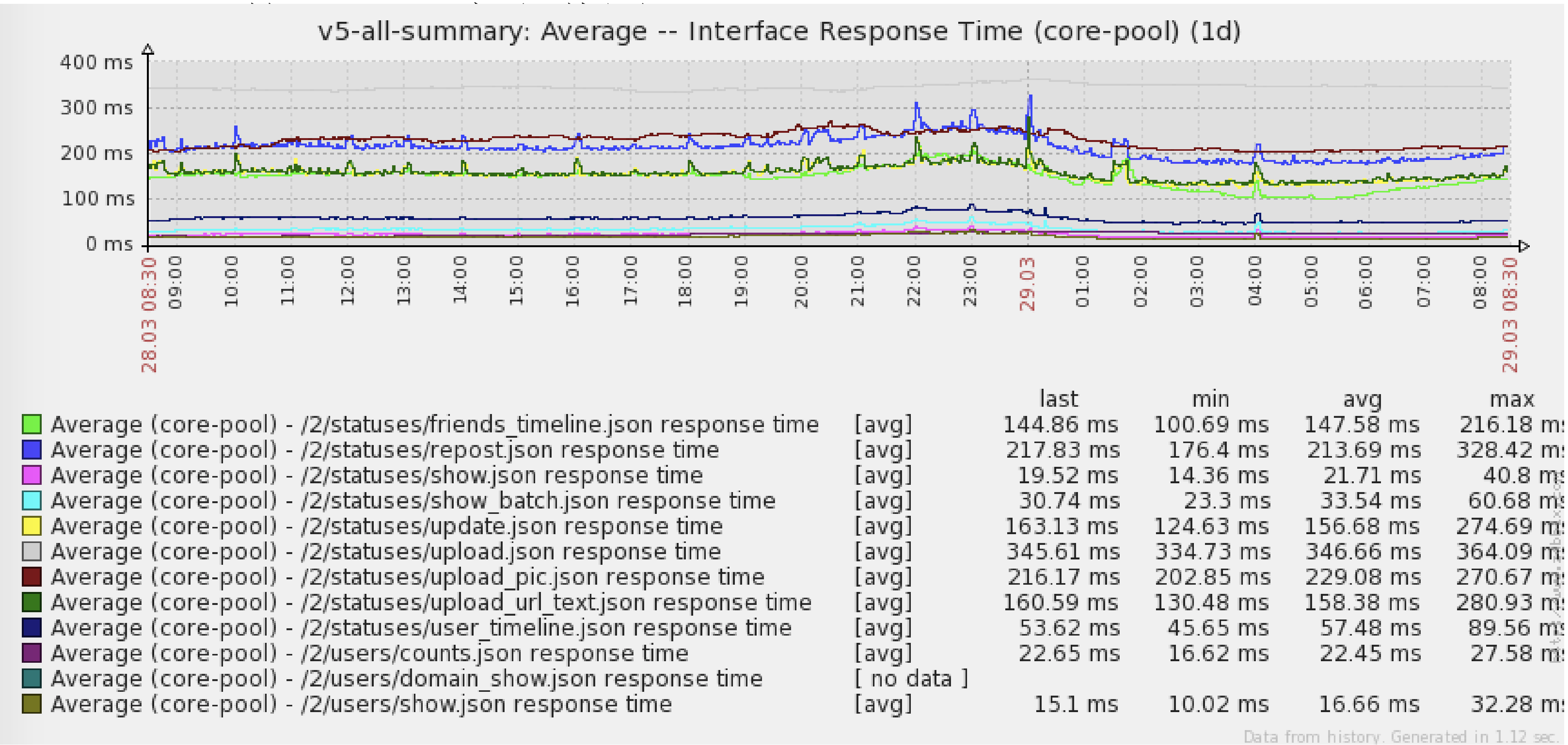
最后是这样



Fields (64017)   Scripted Fields (1)

name	type	analyzed	indexed	popularity
tags	string	false	true	0
host	string	false	true	0
_source	string	false	false	0
_index	string	false	false	0
type	string	false	true	0
@version	string	false	true	0
@timestamp	date	false	true	0
_type	string	false	true	1

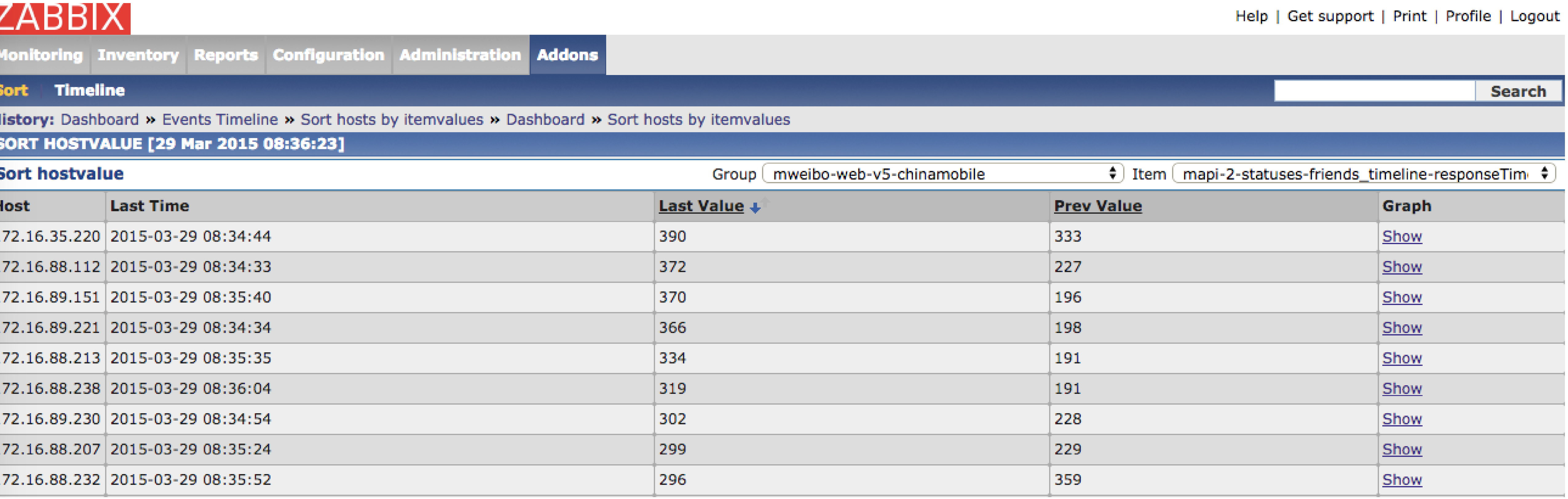
最后是这样



		last	min	avg	max
🟩 Average (core-pool) - /2/statuses/friends_timeline.json response time	[avg]	144.86 ms	100.69 ms	147.58 ms	216.18 ms
🟦 Average (core-pool) - /2/statuses/repost.json response time	[avg]	217.83 ms	176.4 ms	213.69 ms	328.42 ms
🟪 Average (core-pool) - /2/statuses/show.json response time	[avg]	19.52 ms	14.36 ms	21.71 ms	40.8 ms
🟦 Average (core-pool) - /2/statuses/show_batch.json response time	[avg]	30.74 ms	23.3 ms	33.54 ms	60.68 ms
🟨 Average (core-pool) - /2/statuses/update.json response time	[avg]	163.13 ms	124.63 ms	156.68 ms	274.69 ms
⬜ Average (core-pool) - /2/statuses/upload.json response time	[avg]	345.61 ms	334.73 ms	346.66 ms	364.09 ms
🟥 Average (core-pool) - /2/statuses/upload_pic.json response time	[avg]	216.17 ms	202.85 ms	229.08 ms	270.67 ms
🟩 Average (core-pool) - /2/statuses/upload_url_text.json response time	[avg]	160.59 ms	130.48 ms	158.38 ms	280.93 ms
🟦 Average (core-pool) - /2/statuses/user_timeline.json response time	[avg]	53.62 ms	45.65 ms	57.48 ms	89.56 ms
🟪 Average (core-pool) - /2/users/counts.json response time	[avg]	22.65 ms	16.62 ms	22.45 ms	27.58 ms
🟦 Average (core-pool) - /2/users/domain_show.json response time	[no data]				
🟫 Average (core-pool) - /2/users/show.json response time	[avg]	15.1 ms	10.02 ms	16.66 ms	32.28 ms

Data from history. Generated in 1.12 sec.

最后是这样



ZABBIX

Help | Get support | Print | Profile | Logout

onitoring | Inventory | Reports | Configuration | Administration | Addons

ort | Timeline | Search

istory: Dashboard » Events Timeline » Sort hosts by itemvalues » Dashboard » Sort hosts by itemvalues

ORT HOSTVALUE [29 Mar 2015 08:36:23]

ort hostvalue — Group: mweibo-web-v5-chinamobile — Item: mapi-2-statuses-friends_timeline-responseTim

ost	Last Time	Last Value ↓	Prev Value	Graph
72.16.35.220	2015-03-29 08:34:44	390	333	Show
72.16.88.112	2015-03-29 08:34:33	372	227	Show
72.16.89.151	2015-03-29 08:35:40	370	196	Show
72.16.89.221	2015-03-29 08:34:34	366	198	Show
72.16.88.213	2015-03-29 08:35:35	334	191	Show
72.16.88.238	2015-03-29 08:36:04	319	191	Show
72.16.89.230	2015-03-29 08:34:54	302	228	Show
72.16.88.207	2015-03-29 08:35:24	299	229	Show
72.16.88.232	2015-03-29 08:35:52	296	359	Show

这背后…

- Rsyslog/Logstash * 10
- ES masternode * 3
- ES datanode * 28 => 单日 70 亿条，4TB 索引
- Zabbix * 3 => 500k 监控项，60s 间隔

以及

在 400+ Nginx/PHP 机器上通过 Rsyslog 直接调用的 process2zbx 脚本

Elasticsearch 简介

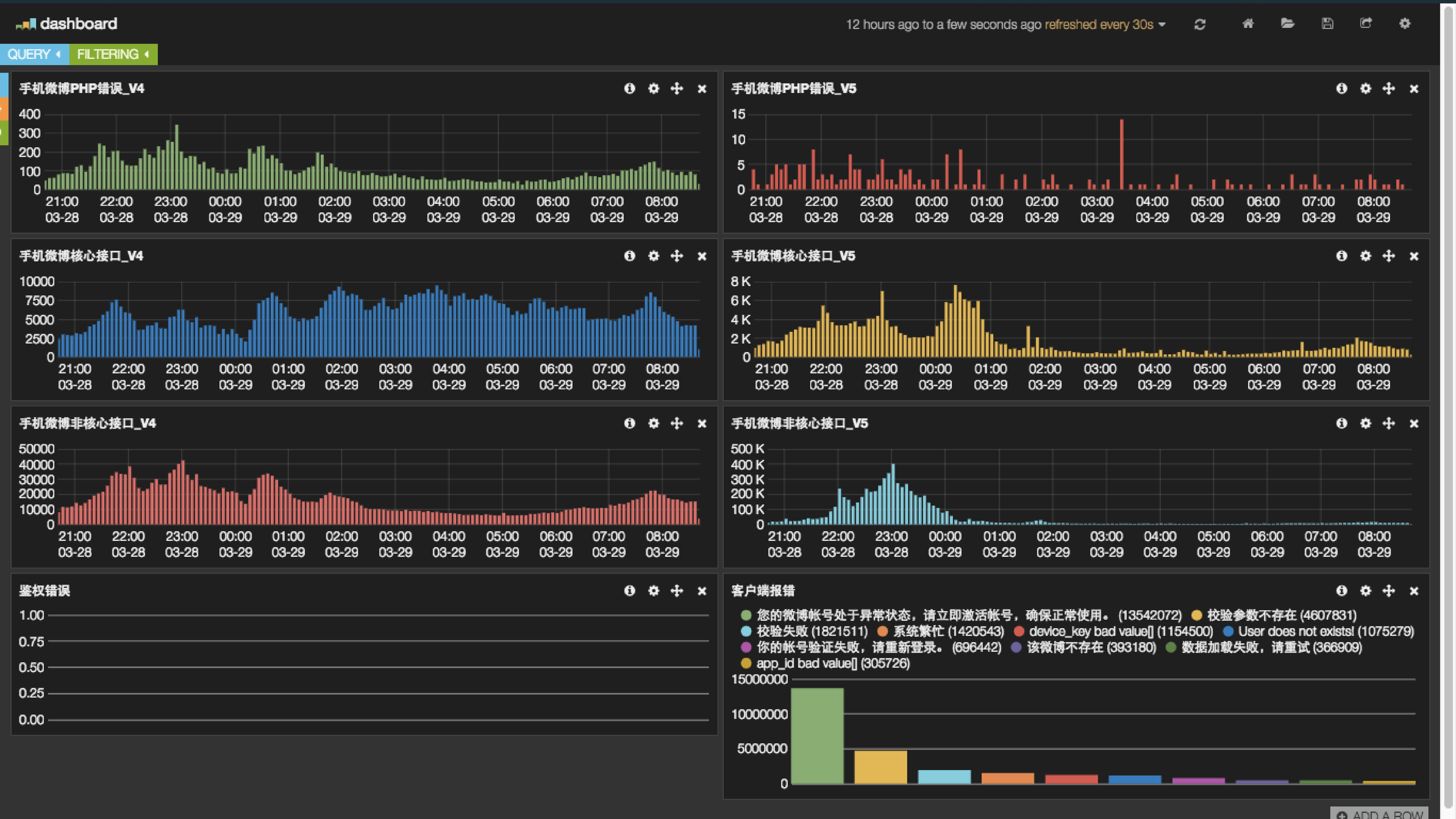
- 基于 Lucene 构建的，提供 RESTful 接口的，分布式，实时全文搜索引擎。http://www.elastic.co
- index/type/replica/shard/segment/relocation/auto-discovery…
- Master 只是维护 cluster state，数据交互可以直接发给任意节点

Kibana 简介

- 基于 Angular.JS 框架写的 Elasticsearch 数据展示前端。
- 开发者是重构狂：

V1	-> PHP
V2	-> Ruby
V3	-> Angular.js
V4beta	-> Angular.js + JRuby
V4	-> Angular.js + Node.JS

- 重复一次：count 是最重要的。



手机微博PHP错误_V4

手机微博PHP错误_V5

手机微博核心接口_V4

手机微博核心接口_V5

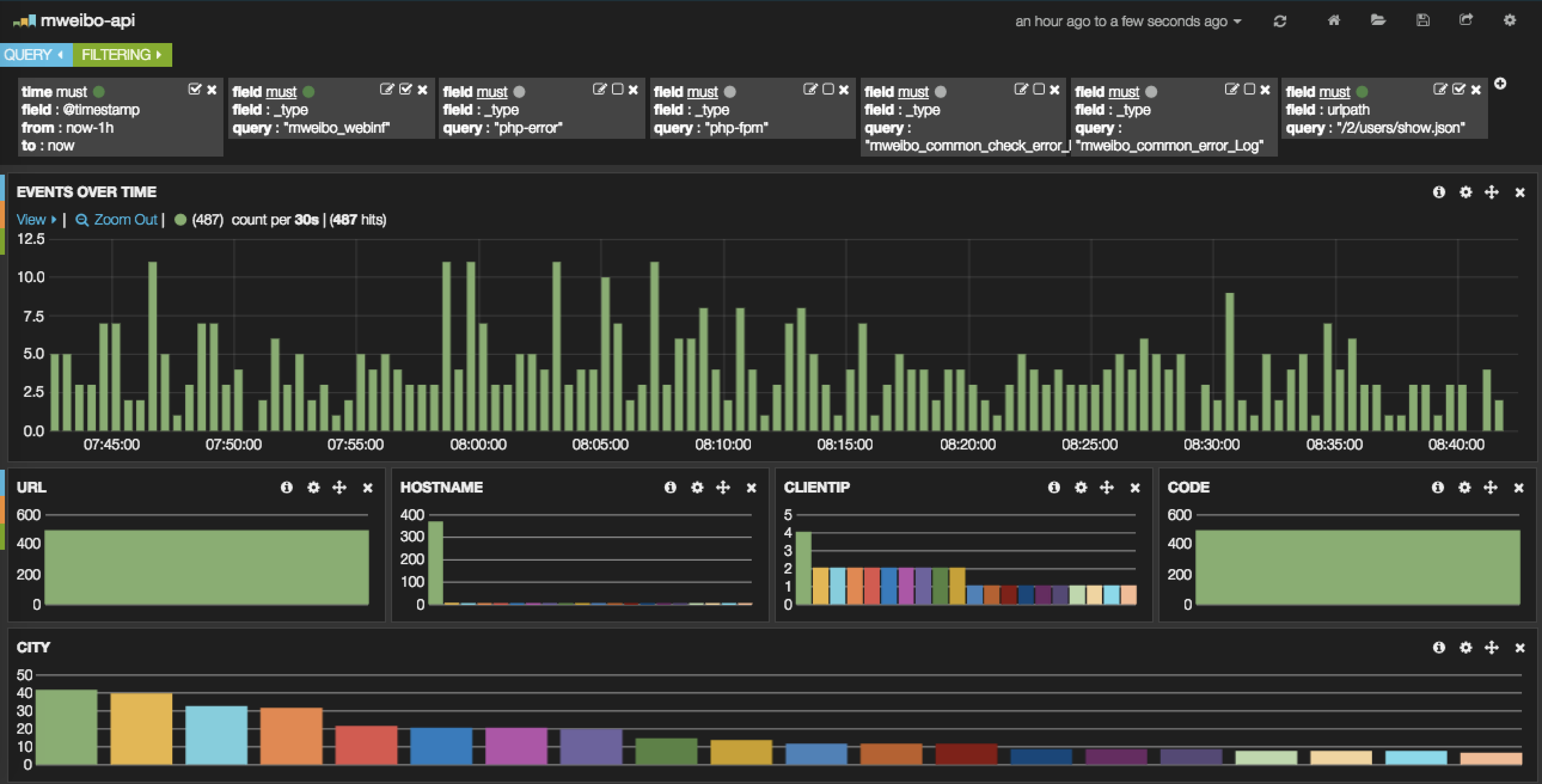
手机微博非核心接口_V4

手机微博非核心接口_V5

鉴权错误

客户端报错

● 您的微博帐号处于异常状态，请立即激活帐号，确保正常使用。 (13542072) ● 校验参数不存在 (4607831)
● 校验失败 (1821511) ● 系统繁忙 (1420543) ● device_key bad value[] (1154500) ● User does not exists! (1075279)
● 你的帐号验证失败，请重新登录。 (696442) ● 该微博不存在 (393180) ● 数据加载失败，请重试 (366909)
● app_id bad value[] (305726)



mweibo-api

an hour ago to a few seconds ago

QUERY FILTERING

time must
field : @timestamp
from : now-1h
to : now

field must
field : _type
query : "mweibo_webinf"

field must
field : _type
query : "php-error"

field must
field : _type
query : "php-fpm"

field must
field : _type
query : "mweibo_common_check_error_l

field must
field : _type
query : "mweibo_common_error_Log"

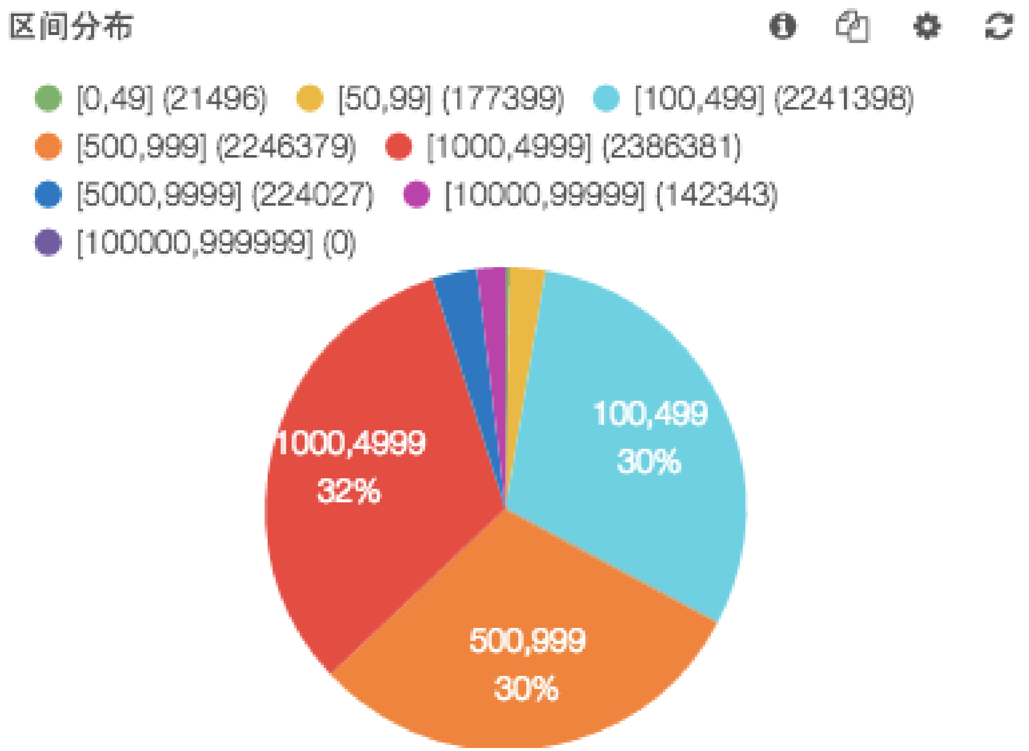
field must
field : urlpath
query : "/2/users/show.json"

EVENTS OVER TIME

View | Zoom Out | (487) count per **30s** | (**487** hits)

Elasticsearch 聚合示例

```
curl 127.0.0.1:9200/_all/_search?pretty=1 -d '{
  "query" : { "match_all" : {} },
  "aggs" : {
     "range1" : {
        "range" : {
           "field" : "resp_ms",
           "ranges" : [ { "to" : 100 },
                        { "from" : 101, "to" : 500 },
                        { "from" : 500 } ]
        }
     }
  }
}'
```



ES 自定义脚本

- config/scripts/regex.groovy 如下：

```
matcher = ( doc[fieldname].value =~ /${pattern}/ )
if (matcher.matches()) {
   matcher[0][1]
}
```

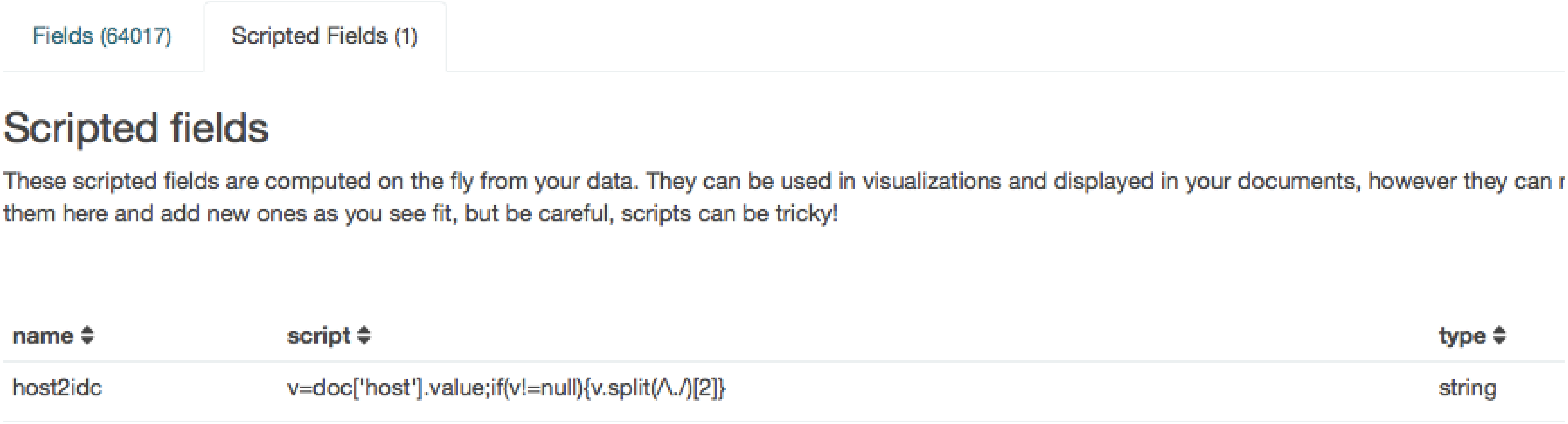
需要下发到每台 datanode 上，ES 自动探测新脚本并加载。

ES 自定义脚本运用

```
curl '127.0.0.1:9200/logstash-2014.11.27/_search?pretty&size=0' -d '{
    "aggs" : {
        "ip_class" : {
            "terms" : {
                "script" : "regex",
                "params" : {
                    "fieldname": "client_ip.raw",
                    "pattern": "^((?:\d{1,3}\.?){3})\.\d{1,3}$"
                }
            }
        }
    }
}'
```

ES 动态脚本示例

```
"aggs": {
  "2": {
```



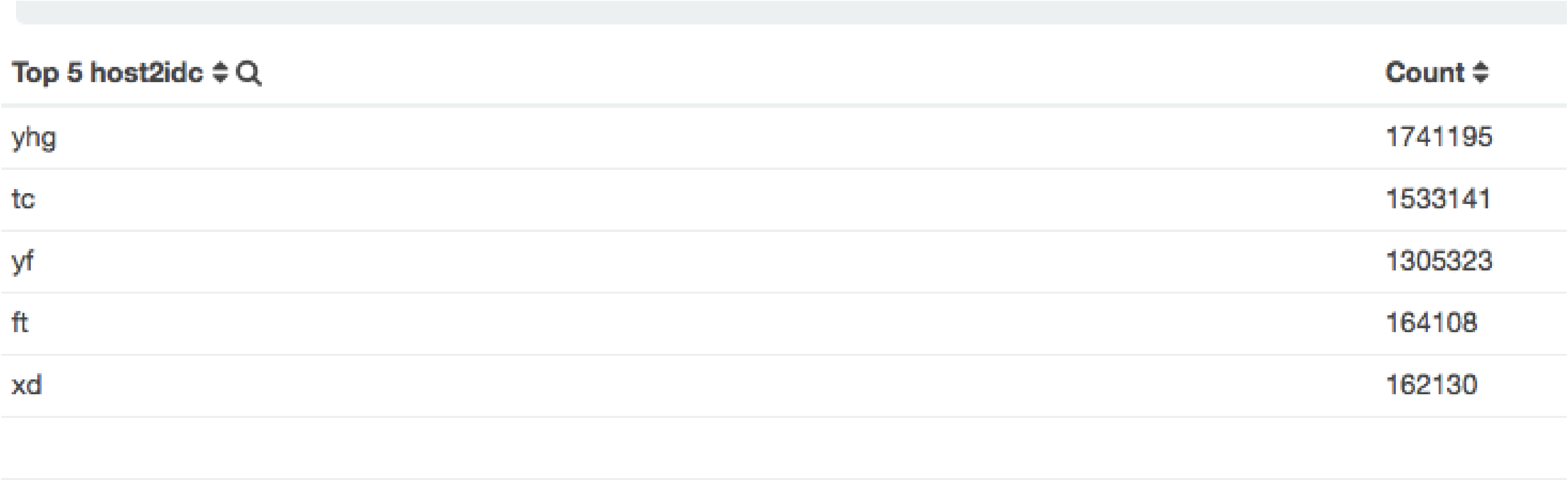
Scripted fields

These scripted fields are computed on the fly from your data. They can be used in visualizations and displayed in your documents, however they can r
them here and add new ones as you see fit, but be careful, scripts can be tricky!

name	script	type
host2idc	v=doc['host'].value;if(v!=null){v.split(/\./)[2]}	string

```
}
```

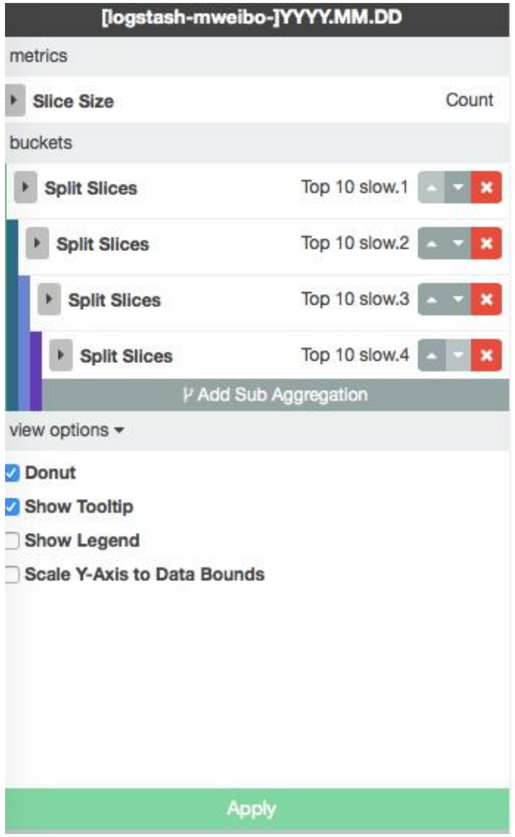
ES 动态脚本示例



Top 5 host2idc ⇕ 🔍	Count ⇕
yhg	1741195
tc	1533141
yf	1305323
ft	164108
xd	162130

ES 嵌套聚合示例

```
"aggs": {
   “1": {
   "terms": { "field": "slow.1 “ },
   "aggs": {
     “2": {
     "terms": { "field": "slow.2" },
     "aggs": {
       “3": {
       "terms": { "field": "slow.3"},
       "aggs": {
         “4": {
         "terms": { "field": "slow.4 “ }
        }
       }
      }
     }
    }
   }
```



[logstash-mweibo-]YYYY.MM.DD

metrics

Slice Size — Count

buckets

Split Slices — Top 10 slow.1

Split Slices — Top 10 slow.2

Split Slices — Top 10 slow.3

Split Slices — Top 10 slow.4

Add Sub Aggregation

view options

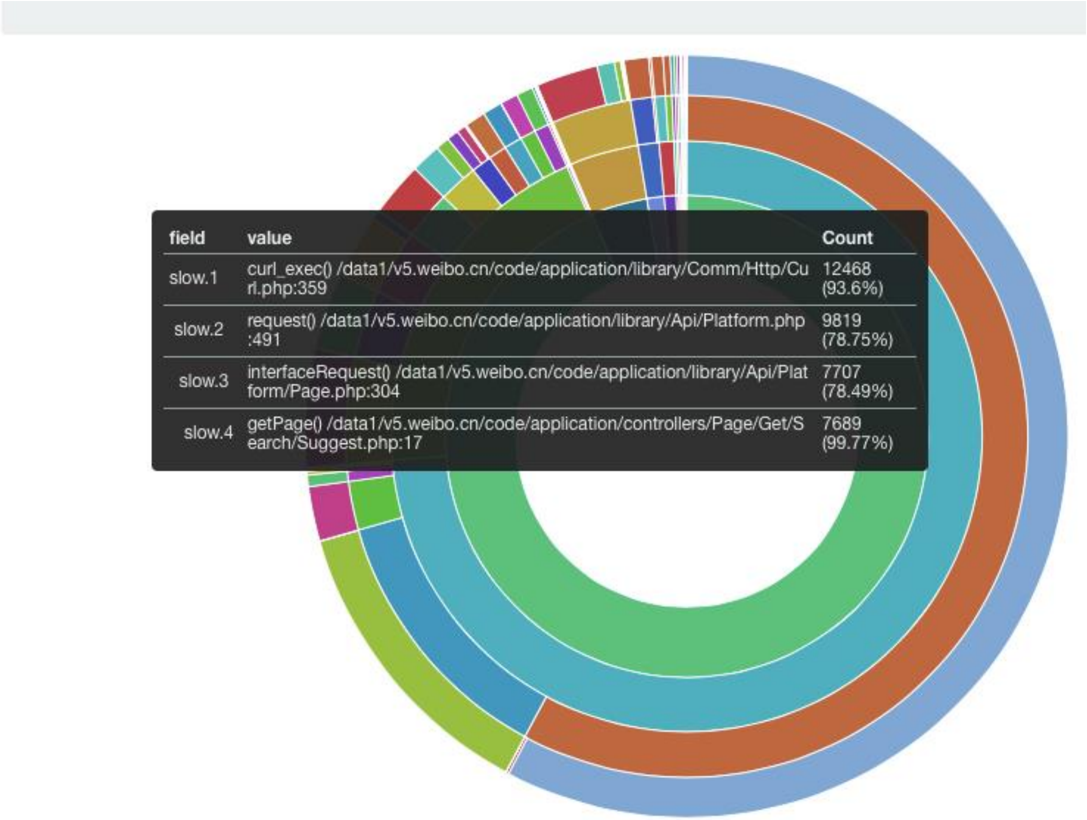
☑ Donut

☑ Show Tooltip

□ Show Legend

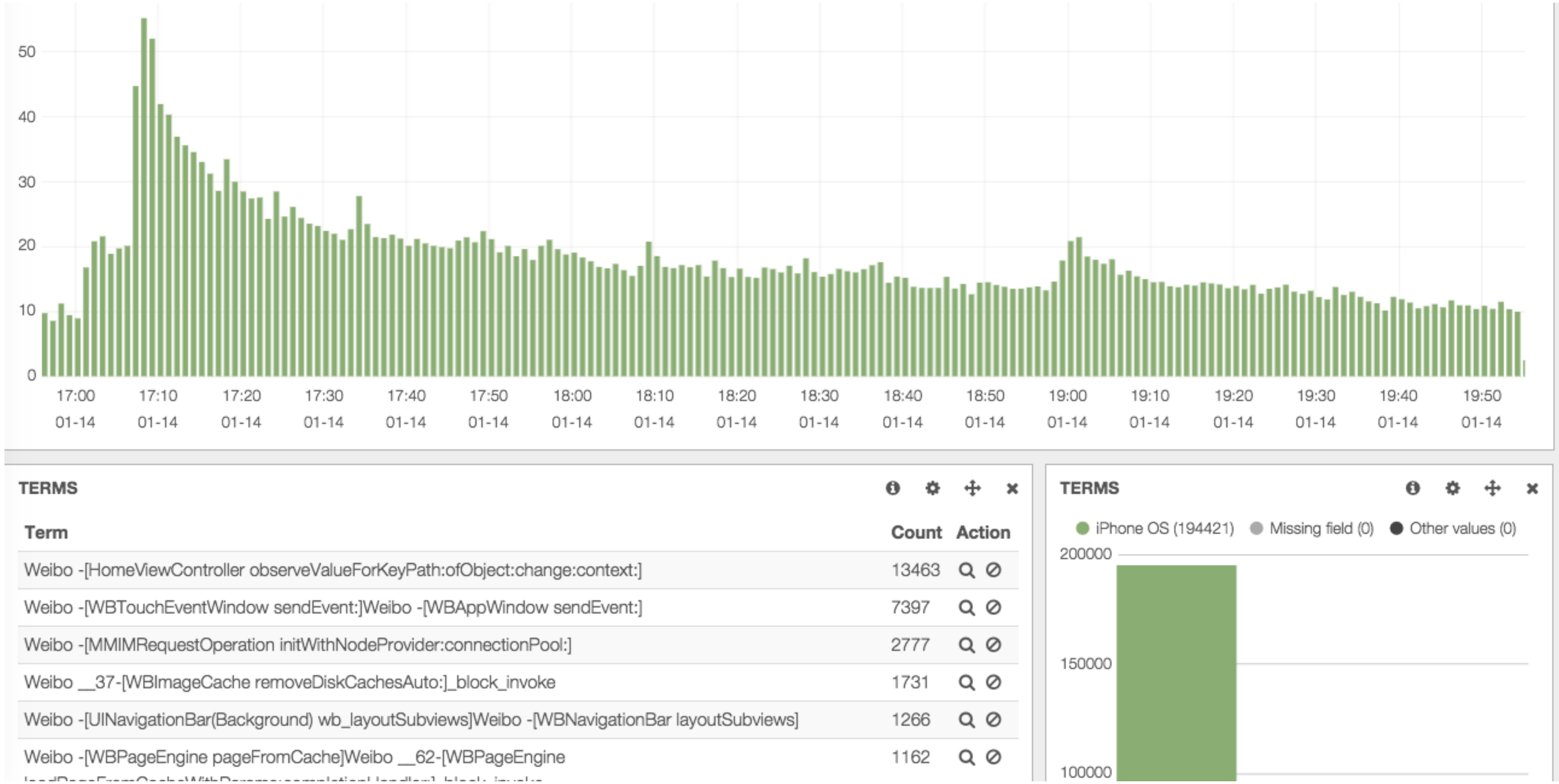
□ Scale Y-Axis to Data Bounds

Apply



field	value	Count
slow.1	curl_exec() /data1/v5.weibo.cn/code/application/library/Comm/Http/Curl.php:359	12468 (93.6%)
slow.2	request() /data1/v5.weibo.cn/code/application/library/Api/Platform.php:491	9819 (78.75%)
slow.3	interfaceRequest() /data1/v5.weibo.cn/code/application/library/Api/Platform/Page.php:304	7707 (78.49%)
slow.4	getPage() /data1/v5.weibo.cn/code/application/controllers/Page/Get/Search/Suggest.php:17	7689 (99.77%)

场景示例(crash func)

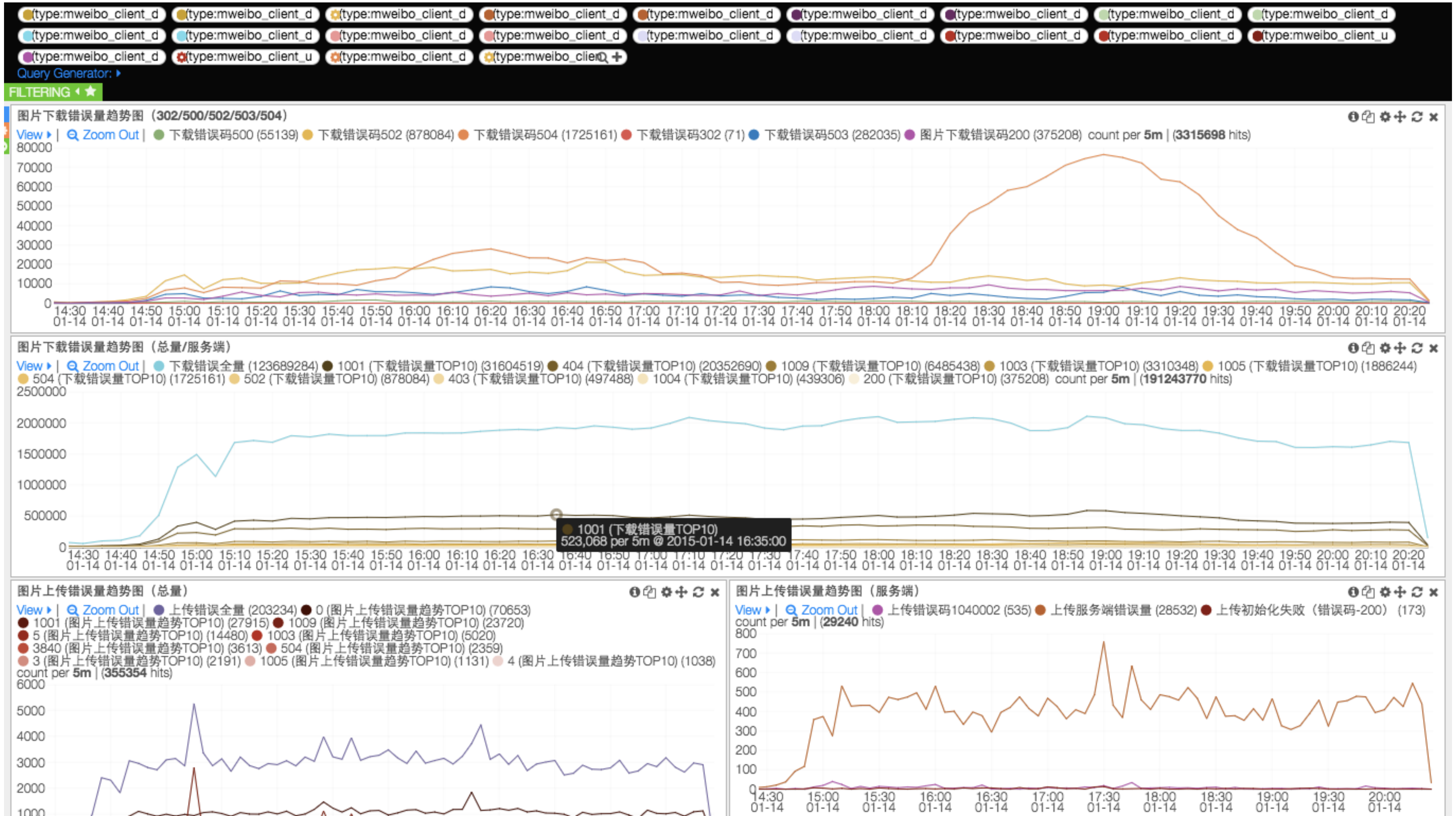


TERMS

Term	Count	Action
Weibo -[HomeViewController observeValueForKeyPath:ofObject:change:context:]	13463	🔍 ⊘
Weibo -[WBTouchEventWindow sendEvent:]Weibo -[WBAppWindow sendEvent:]	7397	🔍 ⊘
Weibo -[MMIMRequestOperation initWithNodeProvider:connectionPool:]	2777	🔍 ⊘
Weibo __37-[WBImageCache removeDiskCachesAuto:]_block_invoke	1731	🔍 ⊘
Weibo -[UINavigationBar(Background) wb_layoutSubviews]Weibo -[WBNavigationBar layoutSubviews]	1266	🔍 ⊘
Weibo -[WBPageEngine pageFromCache]Weibo __62-[WBPageEngine loadPageFromCacheWithParams:completionHandler:]_block_invoke	1162	🔍 ⊘

TERMS

场景示例(error trend)



图片下载错误量趋势图（302/500/502/503/504）

View | Zoom Out | ● 下载错误码500 (55139) ● 下载错误码502 (878084) ● 下载错误码504 (1725161) ● 下载错误码302 (71) ● 下载错误码503 (282035) ● 图片下载错误码200 (375208) count per **5m** | (**3315698** hits)

图片下载错误量趋势图（总量/服务端）

View | Zoom Out | ● 下载错误全量 (123689284) ● 1001 (下载错误量TOP10) (31604519) ● 404 (下载错误量TOP10) (20352690) ● 1009 (下载错误量TOP10) (6485438) ● 1003 (下载错误量TOP10) (3310348) ● 1005 (下载错误量TOP10) (1886244) ● 504 (下载错误量TOP10) (1725161) ● 502 (下载错误量TOP10) (878084) ● 403 (下载错误量TOP10) (497488) ● 1004 (下载错误量TOP10) (439306) ● 200 (下载错误量TOP10) (375208) count per **5m** | (**191243770** hits)

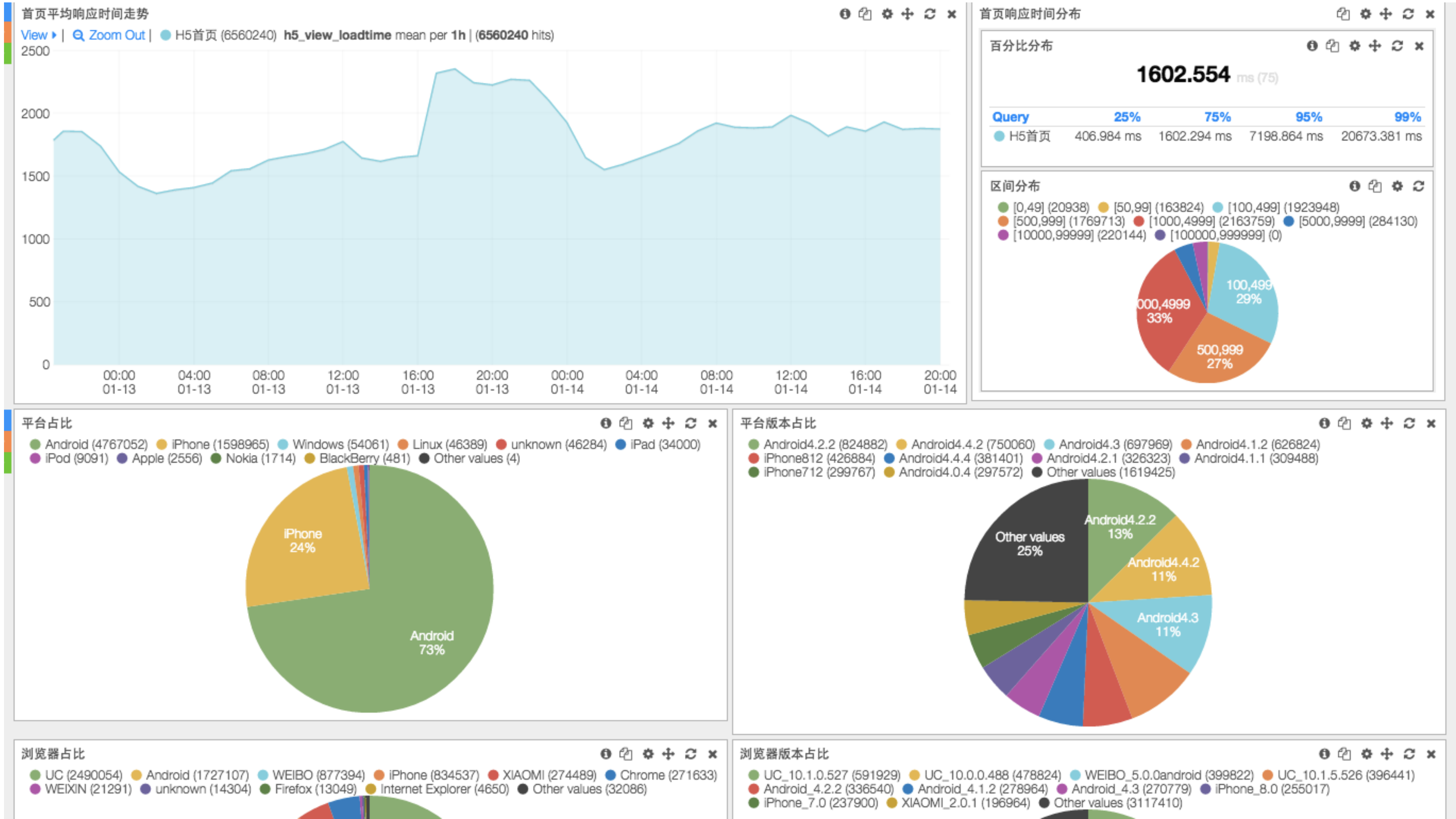
图片上传错误量趋势图（总量）

View | Zoom Out | ● 上传错误全量 (203234) ● 0 (图片上传错误量趋势TOP10) (70653) ● 1001 (图片上传错误量趋势TOP10) (27915) ● 1009 (图片上传错误量趋势TOP10) (23720) ● 5 (图片上传错误量趋势TOP10) (14480) ● 1003 (图片上传错误量趋势TOP10) (5020) ● 3840 (图片上传错误量趋势TOP10) (3613) ● 504 (图片上传错误量趋势TOP10) (2359) ● 3 (图片上传错误量趋势TOP10) (2191) ● 1005 (图片上传错误量趋势TOP10) (1131) ● 4 (图片上传错误量趋势TOP10) (1038) count per **5m** | (**355354** hits)

图片上传错误量趋势图（服务端）

View | Zoom Out | ● 上传错误码1040002 (535) ● 上传服务端错误量 (28532) ● 上传初始化失败（错误码-200） (173) count per **5m** | (**29240** hits)

场景示例(SLA)



首页平均响应时间走势

View ▸ | 🔍 Zoom Out | ● H5首页 (6560240) **h5_view_loadtime** mean per **1h** | (**6560240** hits)

首页响应时间分布

百分比分布

1602.554 ms (75)

Query	25%	75%	95%	99%
● H5首页	406.984 ms	1602.294 ms	7198.864 ms	20673.381 ms

区间分布

● [0,49] (20938) ● [50,99] (163824) ● [100,499] (1923948)
● [500,999] (1769713) ● [1000,4999] (2163759) ● [5000,9999] (284130)
● [10000,99999] (220144) ● [100000,999999] (0)

平台占比

● Android (4767052) ● iPhone (1598965) ● Windows (54061) ● Linux (46389) ● unknown (46284) ● iPad (34000)
● iPod (9091) ● Apple (2556) ● Nokia (1714) ● BlackBerry (481) ● Other values (4)

平台版本占比

● Android4.2.2 (824882) ● Android4.4.2 (750060) ● Android4.3 (697969) ● Android4.1.2 (626824)
● iPhone812 (426884) ● Android4.4.4 (381401) ● Android4.2.1 (326323) ● Android4.1.1 (309488)
● iPhone712 (299767) ● Android4.0.4 (297572) ● Other values (1619425)

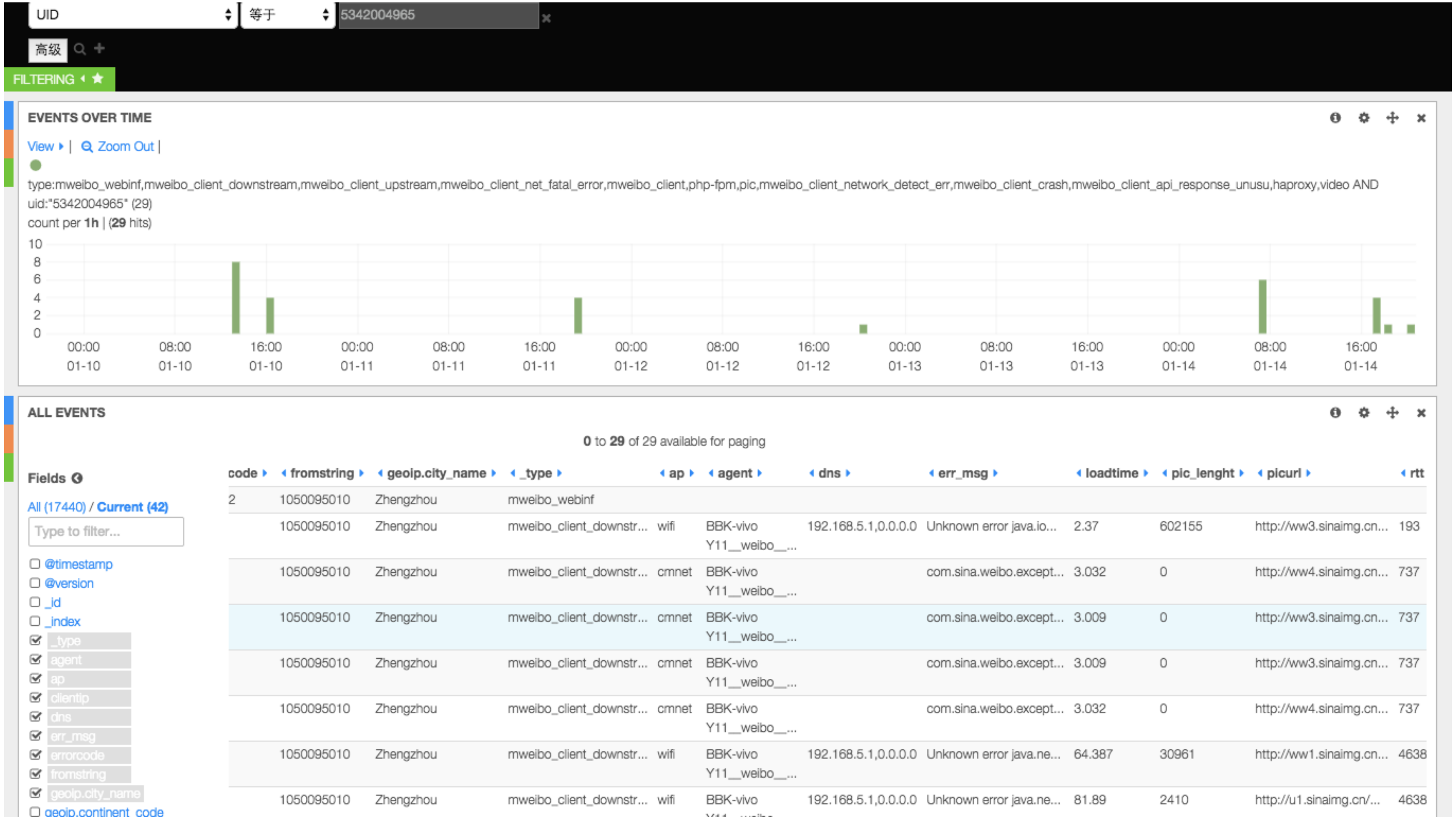
浏览器占比

● UC (2490054) ● Android (1727107) ● WEIBO (877394) ● iPhone (834537) ● XIAOMI (274489) ● Chrome (271633)
● WEIXIN (21291) ● unknown (14304) ● Firefox (13049) ● Internet Explorer (4650) ● Other values (32086)

浏览器版本占比

● UC_10.1.0.527 (591929) ● UC_10.0.0.488 (478824) ● WEIBO_5.0.0android (399822) ● UC_10.1.5.526 (396441)
● Android_4.2.2 (336540) ● Android_4.1.2 (278964) ● Android_4.3 (270779) ● iPhone_8.0 (255017)
● iPhone_7.0 (237900) ● XIAOMI_2.0.1 (196964) ● Other values (3117410)

场景示例(user track)



type:mweibo_webinf,mweibo_client_downstream,mweibo_client_upstream,mweibo_client_net_fatal_error,mweibo_client,php-fpm,pic,mweibo_client_network_detect_err,mweibo_client_crash,mweibo_client_api_response_unusu,haproxy,video AND uid:"5342004965" (29)

count per **1h** | (**29** hits)

0 to **29** of 29 available for paging

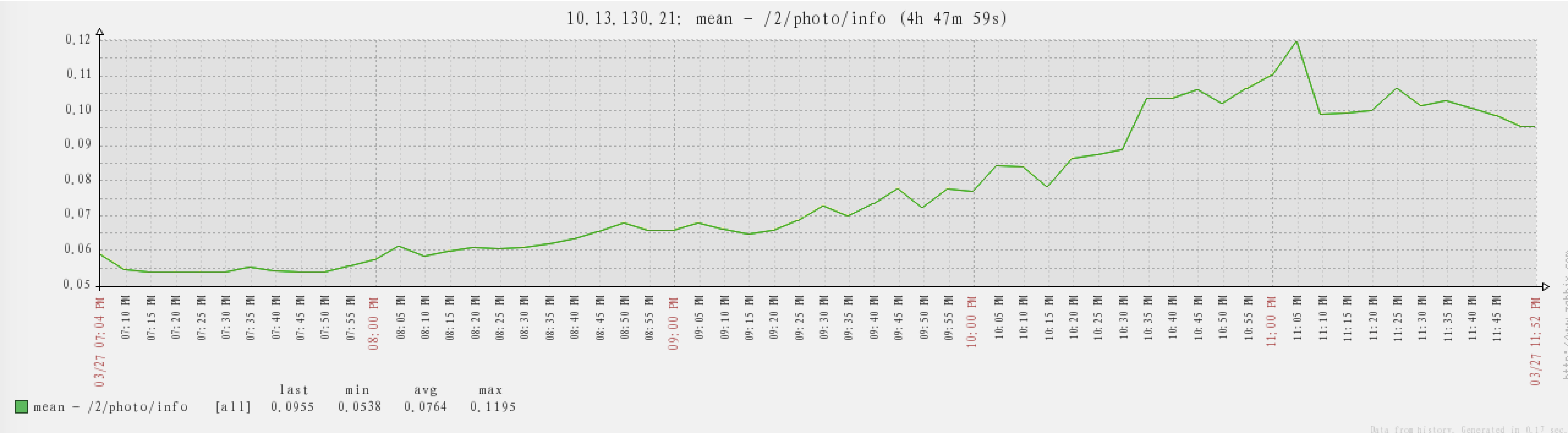
□ @timestamp
□ @version
□ _id
□ _index
☑ _type
☑ agent
☑ ap
☑ clientip
☑ dns
☑ err_msg
☑ errorcode
☑ fromstring
☑ geoip.city_name
□ geoip.continent_code

code	fromstring	geoip.city_name	_type	ap	agent	dns	err_msg	loadtime	pic_lenght	picurl	rtt
2	1050095010	Zhengzhou	mweibo_webinf								
	1050095010	Zhengzhou	mweibo_client_downstr...	wifi	BBK-vivo Y11__weibo__...	192.168.5.1,0.0.0.0	Unknown error java.io...	2.37	602155	http://ww3.sinaimg.cn...	193
	1050095010	Zhengzhou	mweibo_client_downstr...	cmnet	BBK-vivo Y11__weibo__...		com.sina.weibo.except...	3.032	0	http://ww4.sinaimg.cn...	737
	1050095010	Zhengzhou	mweibo_client_downstr...	cmnet	BBK-vivo Y11__weibo__...		com.sina.weibo.except...	3.009	0	http://ww3.sinaimg.cn...	737
	1050095010	Zhengzhou	mweibo_client_downstr...	cmnet	BBK-vivo Y11__weibo__...		com.sina.weibo.except...	3.009	0	http://ww3.sinaimg.cn...	737
	1050095010	Zhengzhou	mweibo_client_downstr...	cmnet	BBK-vivo Y11__weibo__...		com.sina.weibo.except...	3.032	0	http://ww4.sinaimg.cn...	737
	1050095010	Zhengzhou	mweibo_client_downstr...	wifi	BBK-vivo Y11__weibo__...	192.168.5.1,0.0.0.0	Unknown error java.ne...	64.387	30961	http://ww1.sinaimg.cn...	4638
	1050095010	Zhengzhou	mweibo_client_downstr...	wifi	BBK-vivo	192.168.5.1,0.0.0.0	Unknown error java.ne...	81.89	2410	http://u1.sinaimg.cn/...	4638

ES 与 zabbix

```
rm -f /tmp/es2zbx.res
time=`date +"%Y.%m.%d"`
for url in `cat /path/to/url.list`;do
  key=`echo $url | sed 's!/!-!g'`
  curl -o /tmp/$key.res -s -XGET "http://esdomain:9200/logstash-mweibo-nginx-${time}/_search?size=0" -d '{
    "query": { "filtered": { "filter": { "bool": { "must": [
              { "range": {"@timestamp": { "from": "now-5m", "to": "now" } } },
              {"term": { "urlpath": "'"$url"'" } }
     ] } } } },
    "aggs": { "req_avg": { "avg": { "field": "request_time“ } } }
  }'
  count=`cat /tmp/$key.res | grep -oP '"hits":{"total":\d+' | awk -F: '{print $NF}'`
  avg=`cat /tmp/$key.res | grep -oP '"value":\d+\.\d+' | awk -F: '{print $NF}'`
  [[ -z "$avg" ]] && avg=0
  echo “zabbixserver hits$key $count" >> /tmp/es2zbx.res
  echo “zabbixserver mean$key $avg" >> /tmp/es2zbx.res
done
zabbix_sender -i /tmp/es2zbx.res -z zabbixserver
```

ES 与 zabbix


10.13.130.21: mean - /2/photo/info (4h 47m 59s)

		last	min	avg	max
🟩 mean - /2/photo/info	[all]	0.0955	0.0538	0.0764	0.1195

Data from history. Generated in 0.17 sec.

ES 与 spark streaming

近似：

a. 两者都是 near realtime，ES 有 refresh interval，spark streaming 有 batch window；

b. 两者都有一定的聚合统计能力，ES 有 nested aggs，spark 有 Hive SQL；

差别：

a. ES 内已有数据和 schema，可以直接请求，spark 需要以 pull 方式取数据然后再处理成 schemaRDD(而且我们当前的中转系统 rsyslog 不支持让别人 pull 的方式)；

b. spark SQL 直接能通过 group by 语句可以列出日志中全部的 urlpath 的统计情况，ES 更擅长在数据中通过 query 快速缩小聚合范围，或者 terms 方式计算 topN 结果，要监控全部 urlpath，还是要另外单独提供列表才高效。

谢谢大家