



2013中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2013

大数据 数据库架构与优化 数据治理与分析

SequeMedia
盛拓传媒

IT168.com

ITPUB

ChinaUnix

携程集中式日志及其周
边生态系统

Ctrip R&D Framework

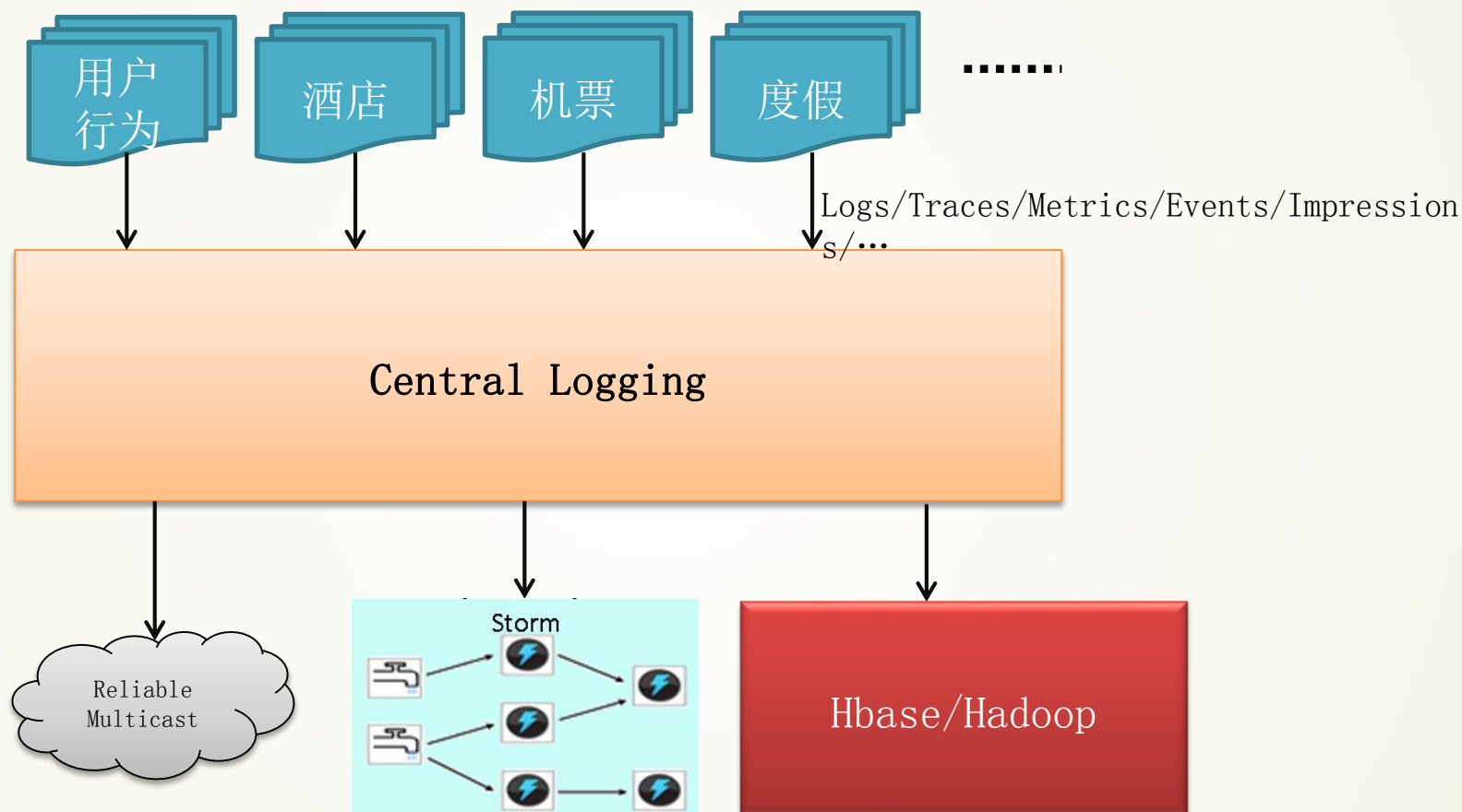
2013-04

Database
BDaaS
flowingdata
DB2
NoSQL MySQL
Oracle Big Data

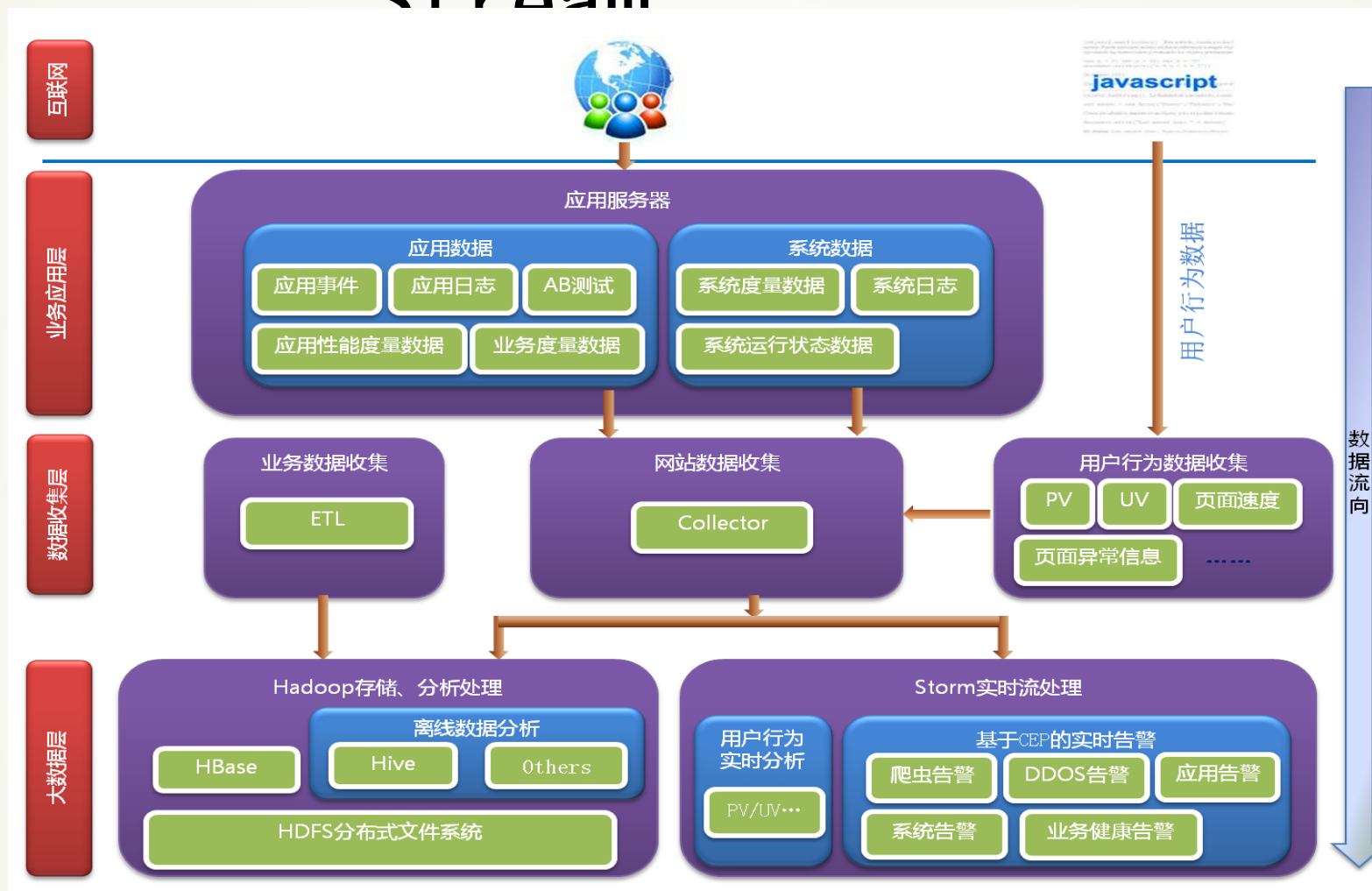
Agenda

1. Brief Introduction
2. Central Logging
3. Log View
4. Dashboard
5. User Behavior Tracking
6. Alerting
7. Data Farm

Central Logging



Central Logging - Data Stream



Central Logging Goals

■ Dev:

- 诊断URL/WS/DB/APP异常，帮助调试
- 发现性能瓶颈 (profiling)
- 关联前、后端的调用

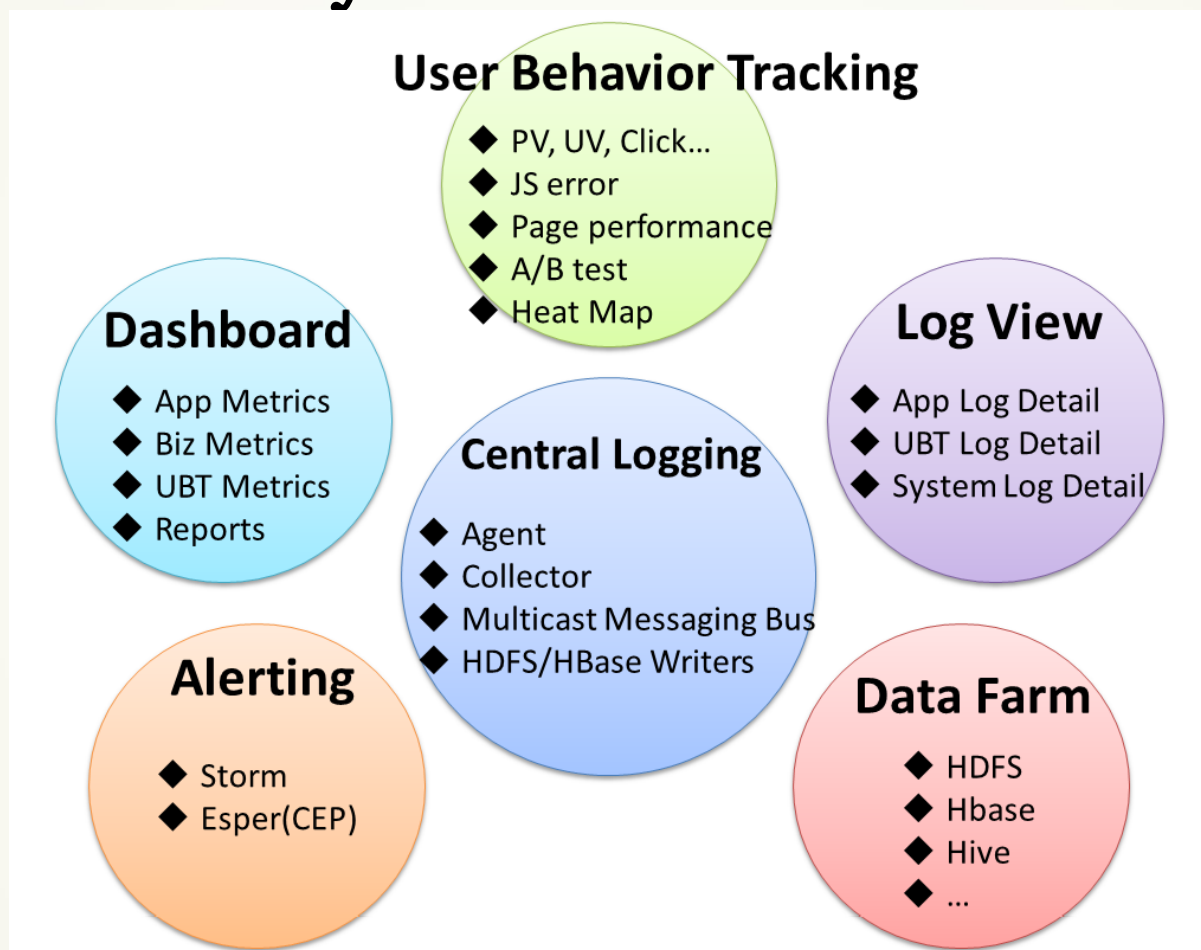
■ PM/BU:

- 通过用户查询理解用户需求 (query understanding)
- 基于用户行为调整搜索结果和推荐 (impressions/clicks/orders)
- 实时收集、分析A/B测试结果，以提高转换率

■ Ops:

- 网站Capacity分析，更合理地规划硬件资源
- 系统、应用、业务的实时告警
- 了解不同应用之间的依赖耦合关系

Central Logging & Eco-system

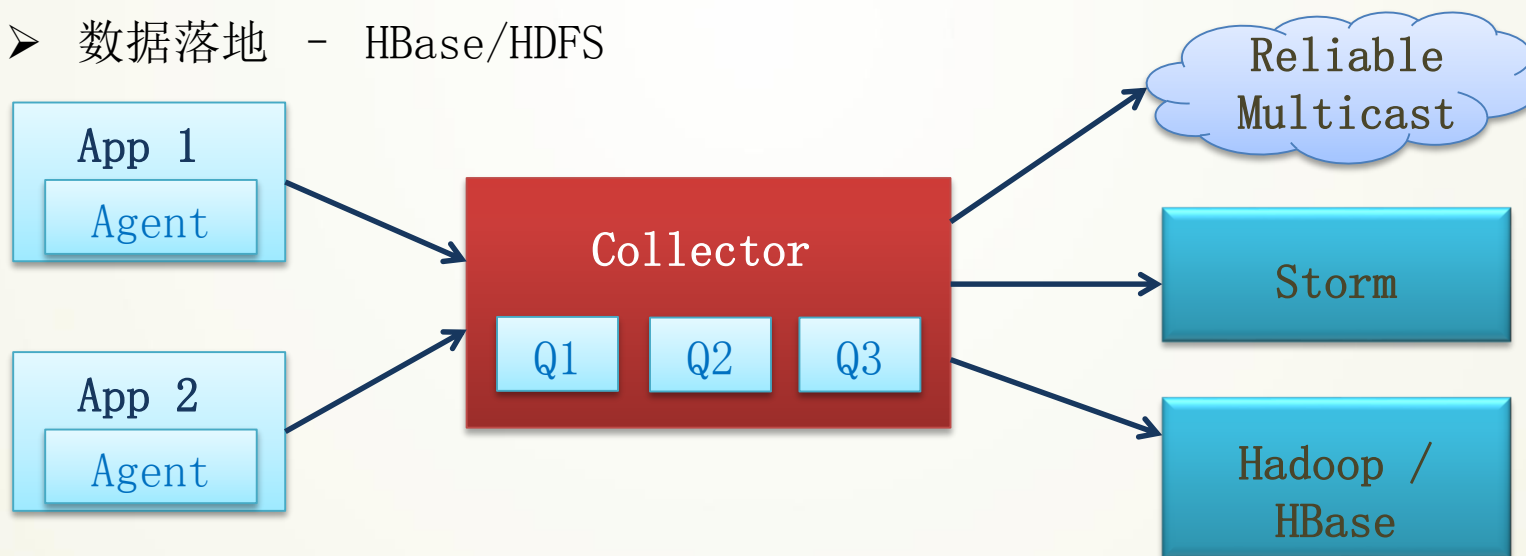


Agenda

1. Brief Introduction
2. Central Logging
3. Log View
4. Dashboard
5. User Behavior Tracking
6. Alerting
7. Data Farm

Central Logging

- 客户端 - Logging Agent
- 数据汇聚层 - Collector
- 数据消费者:
 - 实时消息处理 - Reliable Multicast
 - 实时流式处理 - Storm
 - 数据落地 - HBase/HDFS



Central Logging - Agent API

■ Logging API

- 提供类logback/log4j/log4net接口
- 基于tags扩展，便于搜索查询

■ Trace API

- 度量单次URL/WS/DB调用
- 通过TraceID关联前、后端调用，厘清依赖关系

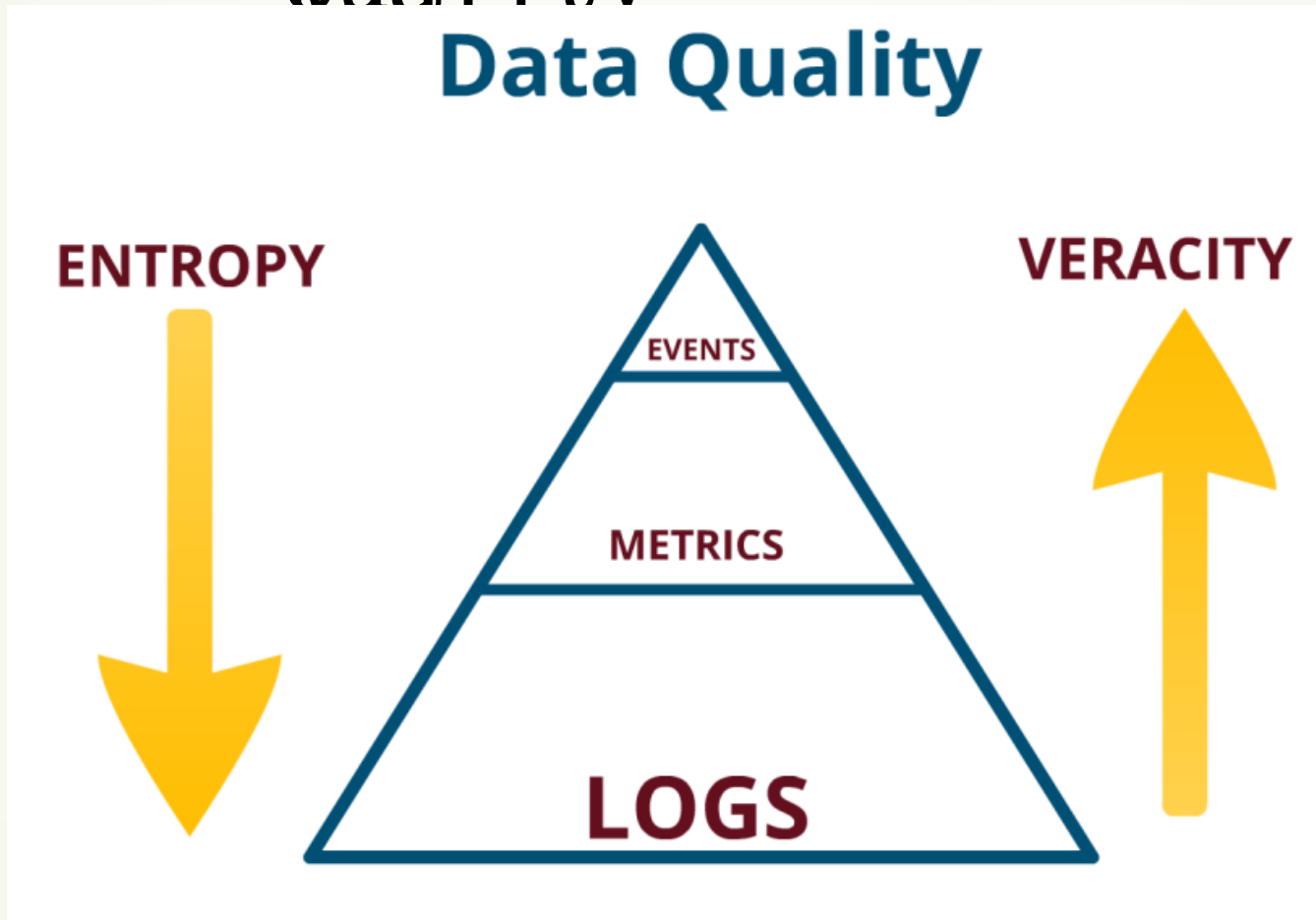
■ Metrics API

- 分钟级聚合统计
- 支持avg, sum, min, max等聚合操作
- 支持用户自定义tags(维度)

■ Event API

- 指明event类型、发生时间和其他相关属性
- 作为CEP实时处理的数据源

Central Logging - Data Quality



Reference: <http://www.infoq.com/presentations/Big-Data-Monitoring-eBay>

Central Logging - Agent Data Collection

■ SOA框架

- 应用启动、关闭、心跳事件
- 错误、异常日志度量统计
- URL/WS/DB/Cache等Trace调用
 - 度量统计
 - TraceID生成和传递
 - SQL hash、statement采样

■ 应用

- 应用日志
- 业务、应用度量数据
- 应用自定义事件

Central Logging - Logging Collector

■ 类似消息队列的broker

- At-most-once (not reliable)
- 日志无排序需求
- 支持Topic语义
- 内存 + mmap队列，屏蔽后端故障

■ 吞吐量优先

- 支持长、短链接通信
- 支持以chunk为单位打包、压缩传输

■ 可水平扩展

■ 高可用

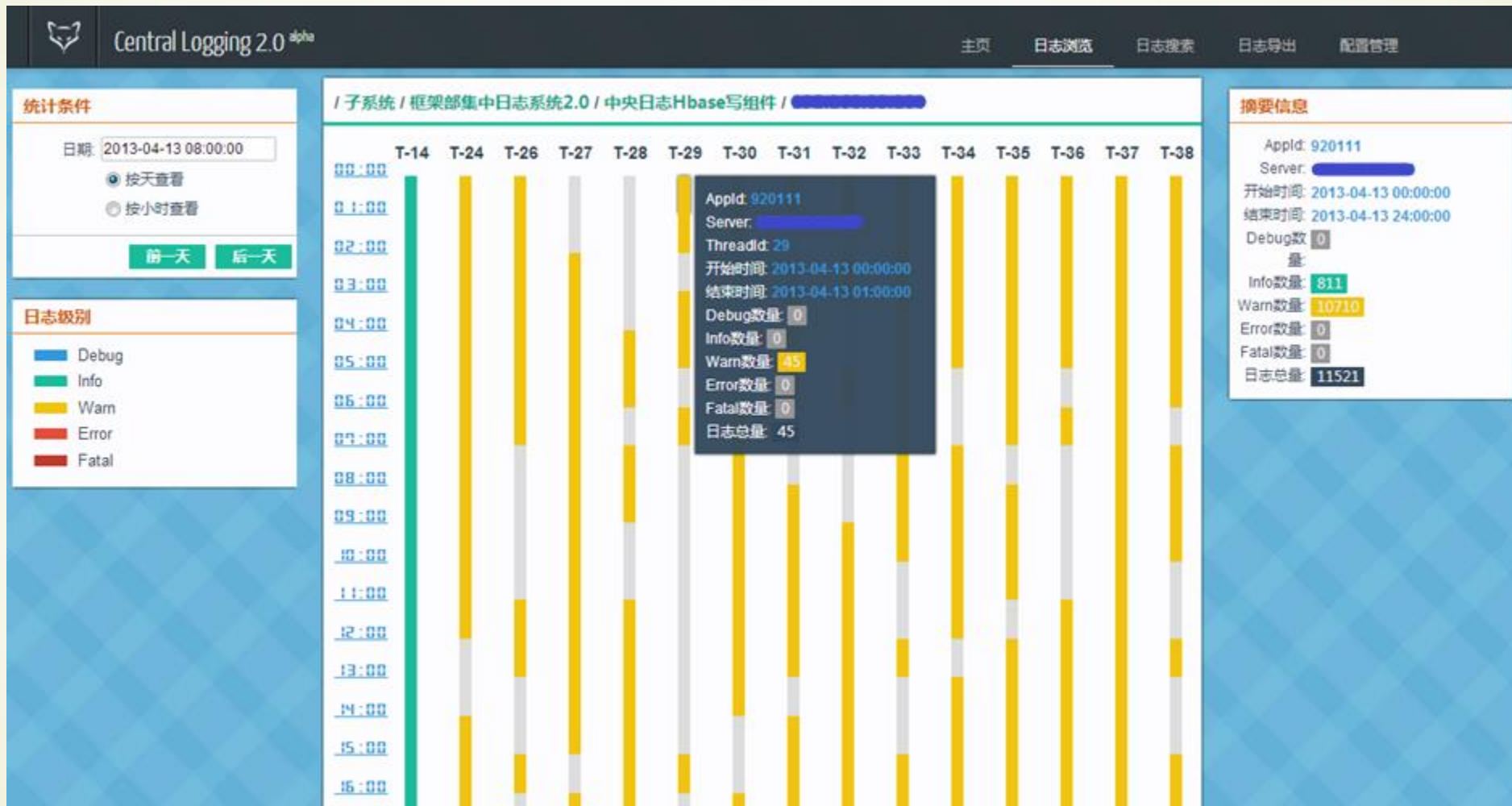
Agenda

1. Brief Introduction
2. Central Logging
3. Log View
4. Dashboard
5. User Behavior Tracking
6. Alerting
7. Data Farm

Log View

- 查看Log、Trace信息
 - 支持浏览或搜索方式
- 关联前、后端调用链日志
- 性能分析、Profiling
- 关键技术：Hbase Data Schema Design

Log View - Log Navigation Demo



Log View - Log Navigation Demo (Cont.)



Log View - Log Search Demo (Old)

Log Entries

Quickfilter

AppID(*):

Date From:

Date To:

Host Name or IP:

Log Title:

Message:

Log Source:

Log Level: or higher

Log Type:

Add additional field Run filter

Quickfilter hit 97 messages, last searched timestamp is 18 March 2013 04:45:01.

Date	Level	Title	Message
18 March 2013 16:46:00	ERROR	反序列化错误	There is an error in XML document (5, 21). at System.Xml.Serialization.XmlSerializer.Deserialize(XmlReader xmlReader, String ...
18 March 2013 16:46:00	ERROR	反序列化错误	There is an error in XML document (5, 21). at System.Xml.Serialization.XmlSerializer.Deserialize(XmlReader xmlReader, String ...
18 March 2013 16:46:00	ERROR	反序列化错误	There is an error in XML document (5, 21). at System.Xml.Serialization.XmlSerializer.Deserialize(XmlReader xmlReader, String ...
18 March 2013 16:46:00	ERROR	反序列化错误	There is an error in XML document (5, 21).

Message 452301130

反序列化错误

From:

AppId: 900201

Date: 18 March 2013 16:46:00

Severity: ERROR

Type: APP

Source:

TraceId: 0

ThreadId: 11

Full message:

There is an error in XML document (5, 21).
at System.Xml.Serialization.XmlSerializer.Deserialize(XmlReader xmlReader, String encodingStyle, XmlDeserializationEvents events)
at System.Xml.Serialization.XmlSerializer.Deserialize(XmlReader xmlReader, String encodingStyle)
at Ctrip.SOA.Comm.XmlSerializer.Deserialize(String xml, Type type, Encoding encode, Boolean needException)

Additional info:

OrderID: 0

ErrorID: 900201000000

ClientIP:

ModuleID: 900201000

Remark: error:xml version="1.0"
Response

 Header ServerIP="fe80::d2a:877c:6a3d:6a6%1"

 Should Record PerformanceTim



2013中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2013
大数据 数据库架构与优化 数据治理与分析


SequeMedia
盛拓传媒

IT168

ITPUB

ChinaUnix

Log View - Log Search Demo (New)

 Central Logging 2.0 alpha

主页 日志浏览 日志搜索 日志导出 配置管理

搜索条件

应用号: 999912
开始时间: 2013-04-12 00:00:00
结束时间: 2013-04-12 24:00:00
IP: 172.16.154.158
HostName:
线程:
标题:
信息:
来源:
日志级别: ☒ 以上
日志类型:

搜索概要

本次搜索到16条日志信息
最后搜索时间:2013-04-12 15:34:50

搜索结果:

NA	test message	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
NA	this is the end of the world!	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
hello	hello world	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
test xml	<a>this is a test xmlhello world	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
warn	something bad will happen	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
just throw this exception	java.lang.Exception: just throw this exception at com.ctrip.freeway.sample.demo.LogDemo.doSomething(LogDemo.java:72) at com.ctrip.freeway.sample.demo.LogDemo.main(LogDemo.java:48) at sun.reflect...	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
NA	这是一个测试日志!!!!!!副书记呵	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
NA	This is a debug log	来自:172.16.154.158	线程:1	APP	2013-04-12 15:35:59
NA	test message	来自:172.16.154.158	线程:1	APP	2013-04-12 15:34:51
hello	hello world	来自:172.16.154.158	线程:1	APP	2013-04-12 15:34:51
test xml	<a>this is a test xmlhello world	来自:172.16.154.158	线程:1	APP	2013-04-12 15:34:51

日志级别

Debug

Info

Warn

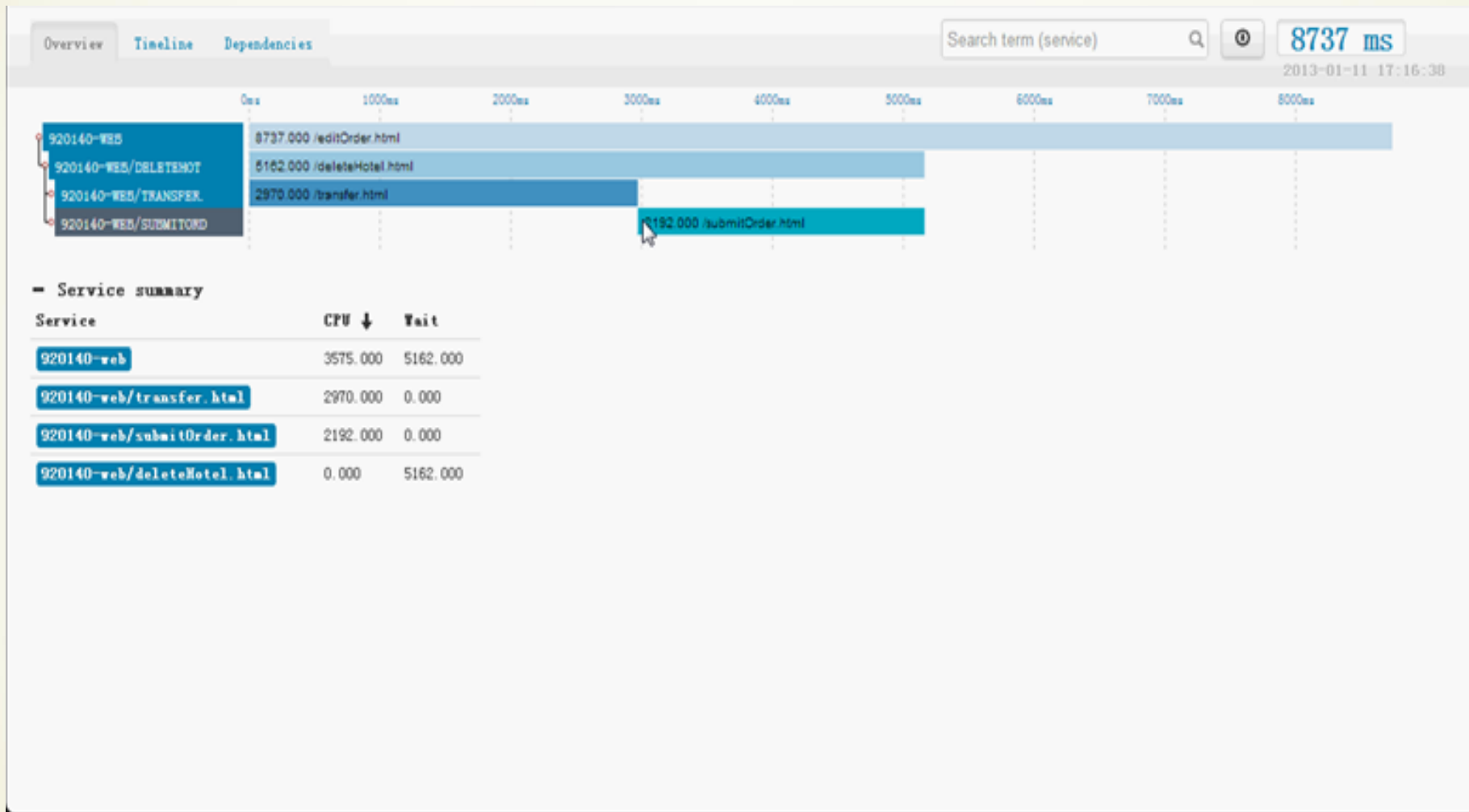
Error

Fatal

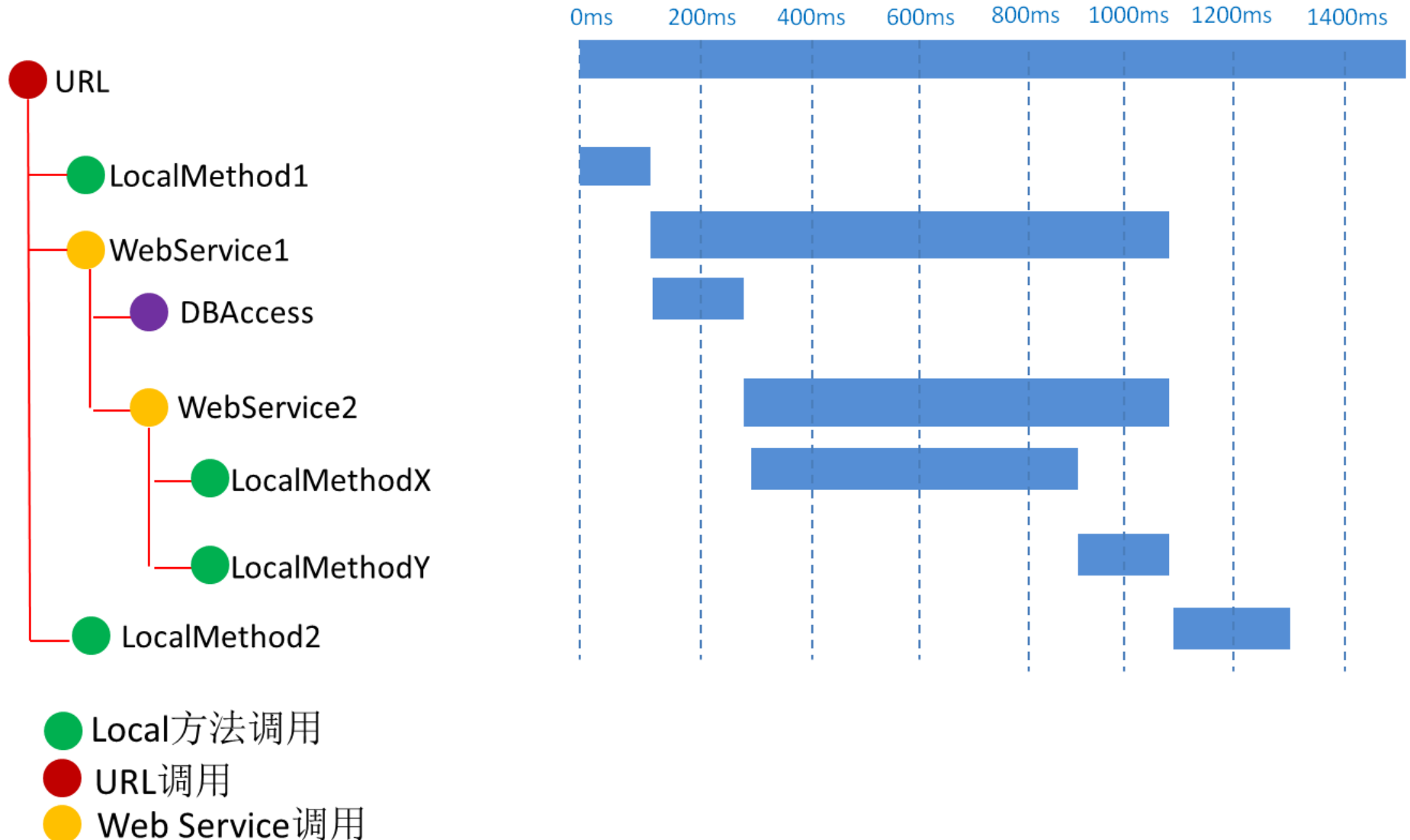
概要信息

应用号: 999912
创建时间: 2013-04-12 15:35:59
IP: 172.16.154.158
HostName: DST52542.cn1.global.ctrip.co
线程: 1
级别: ERROR
类型: APP
来源: com.ctrip.freeway.sample.der

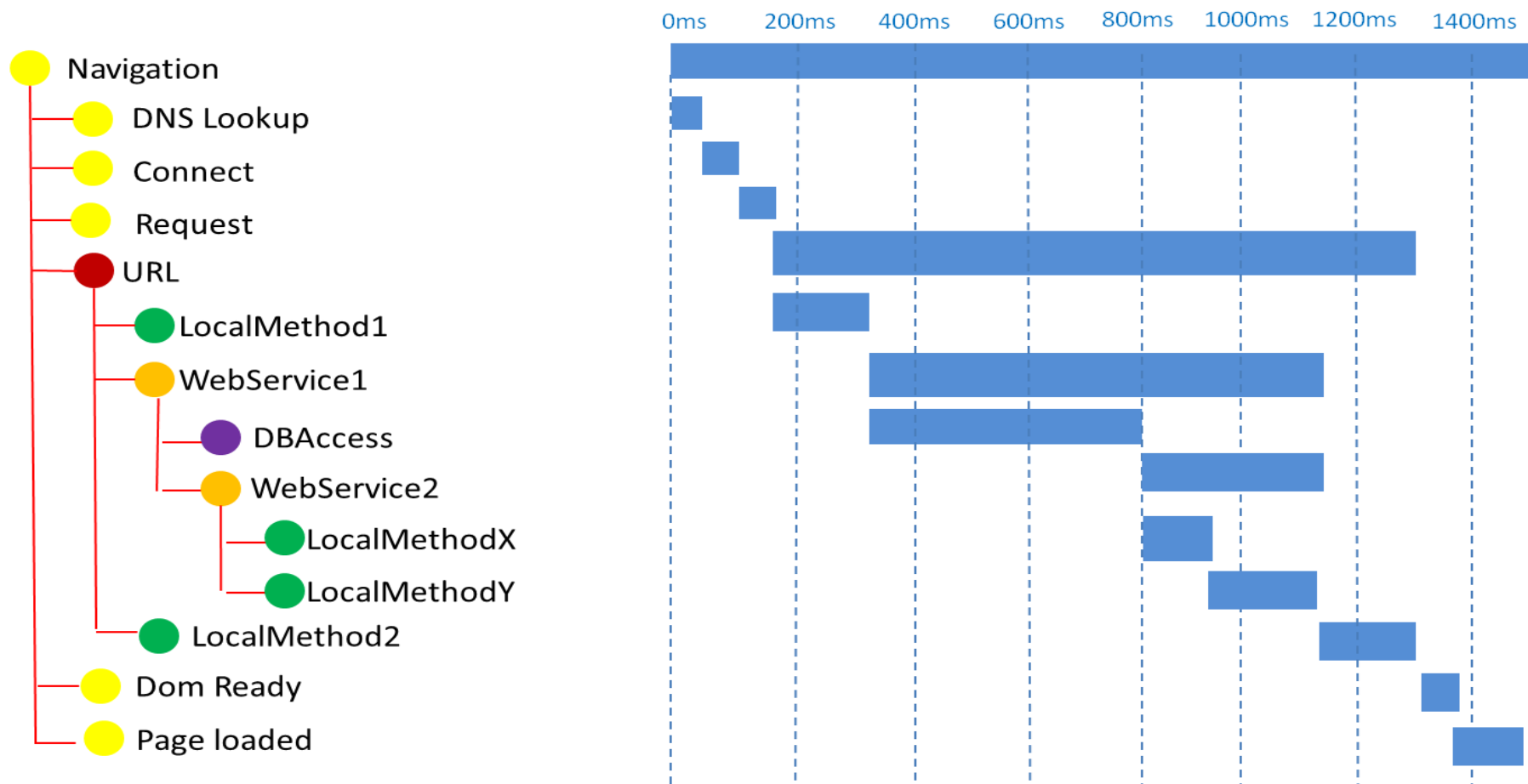
Log View - Trace Demo



Log View - Trace Today



Log View - Trace Tomorrow



- Local方法调用
- URL调用
- Web Service调用
- DB调用
- Browser开销

Log View - FE & BE Correlation (Coming soon)

flights.ctrip.com/booking/SHA-BJS-day-1.htm?showdiag=1

首页 旅游度假 国内酒店 海外酒店 惠选酒店 国内机票 国际机票 团购特卖 礼品卡 目的地探索 火车票 订餐 企业商旅

上海到北京机票 (单程)

机票查询 机+酒

单程 出发城市* 上海 到达城市* 北京 出发日期* 2013-03-14 重新搜索 高级搜索

03-13 周三 点击查询 03-14 周四 ¥510 03-15 周五 ¥380 03-16 周六 ¥400 03-17 周日 ¥400 03-18 周一 ¥400 03-19 周二 ¥340 查看90天最低价

航空公司 全部 ☐春秋航空 ☐东方航空 ☐海南航空 ☐吉祥航空 ☐南方航空 ☐上海航空 ☐中国国航
计划机型 全部 ☐大型机 ☐中型机 起飞机场 全部 ☐虹桥 ☐浦东 降落机场 全部 ☐首都 ☐南苑

起飞时间从早到晚 时间 全天 价格

时间↓	机场	航空公司/舱位	价格/折扣	舱
07:00	虹桥国际机场	东方航空 MU5137	¥1020	经济舱
09:20	首都国际机场	计划机型: 333 (大)	9.0折 退改签	更多舱位

最高返现¥50 查看更多舱位

业务开发对单次调用输出详细调试信息

Showdiag=1可以展示详细debug信息, 只适用内网

特价往返信息

往返特价 海南航空
往返程均乘坐海南航空航班享受特价
¥565+¥565起 (不含税)

航班时刻表

上海航班时刻表
上海到北京时刻表
上海实时起降
上海到北京实时起降

预订即送 惊喜不断

Connected to DB: FlightDBHost(10.2.1.3)

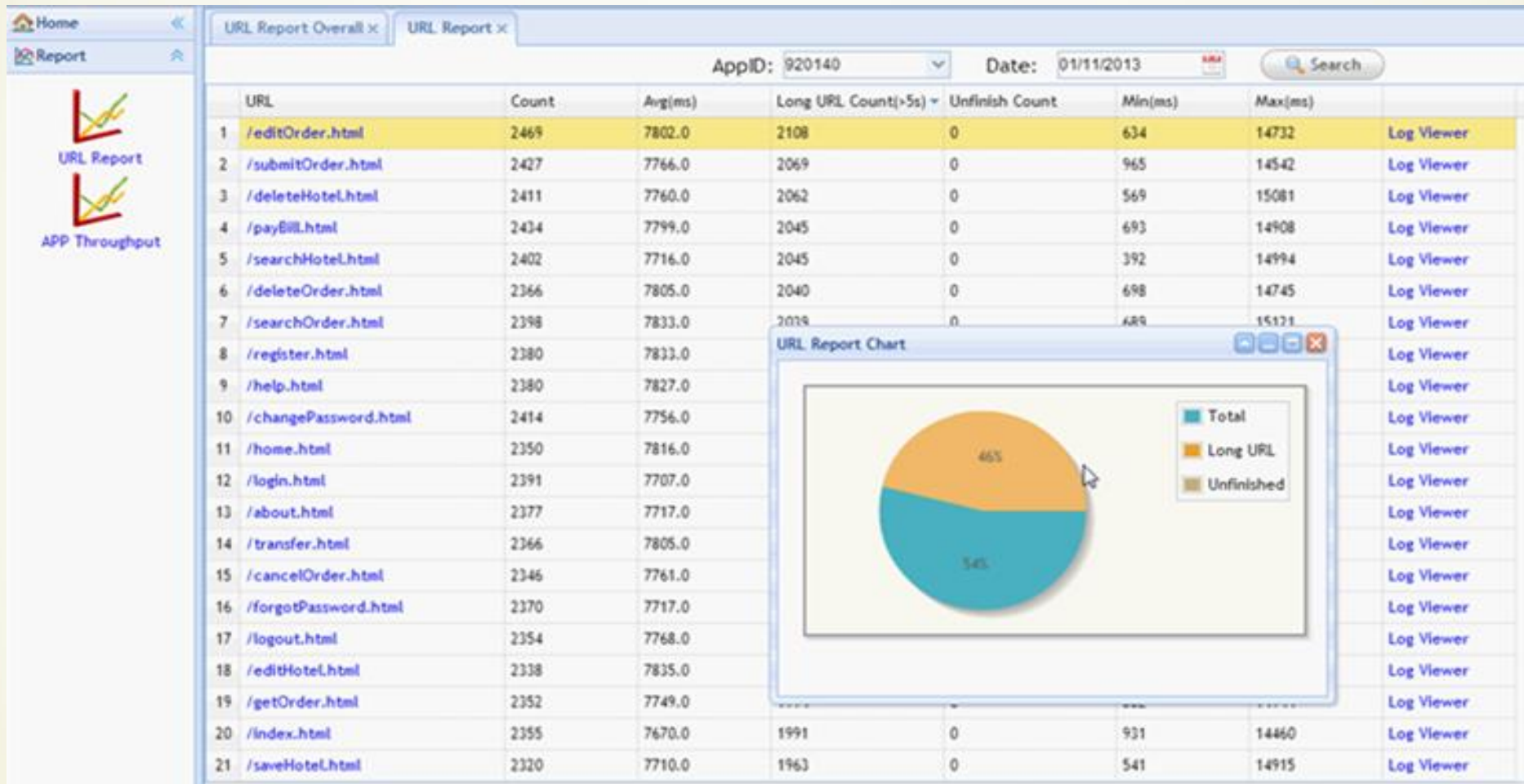
Connected to FlightEngineWS: <http://flightengine.ctriptravel.com>, 参数departcity=上海, ...

Results from FlightEngineWS: <xml> ... </xml>

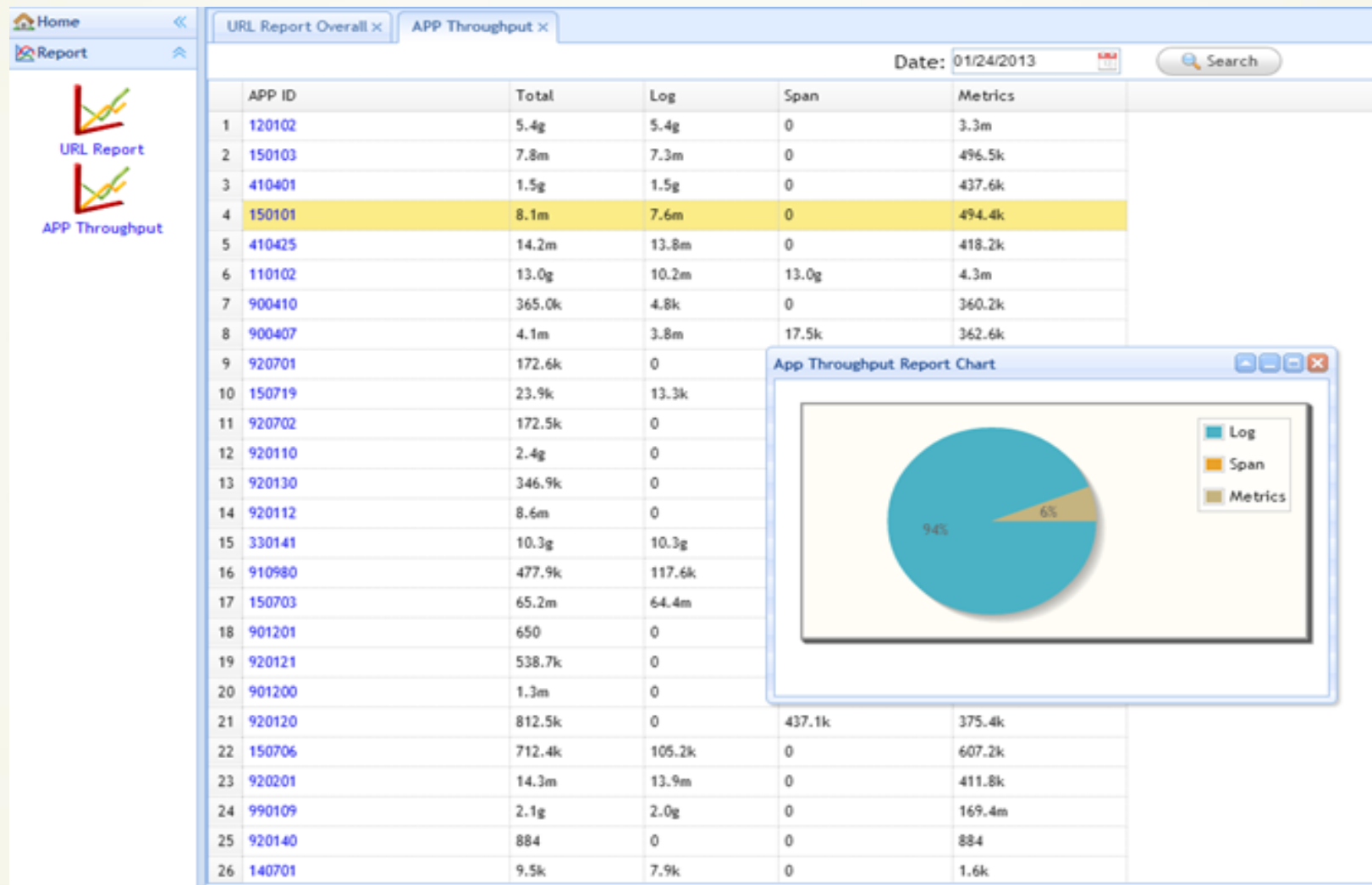
Central logging: <http://traceview.logging.sh.ctriptravel.com/appid=102021>

Other debug information

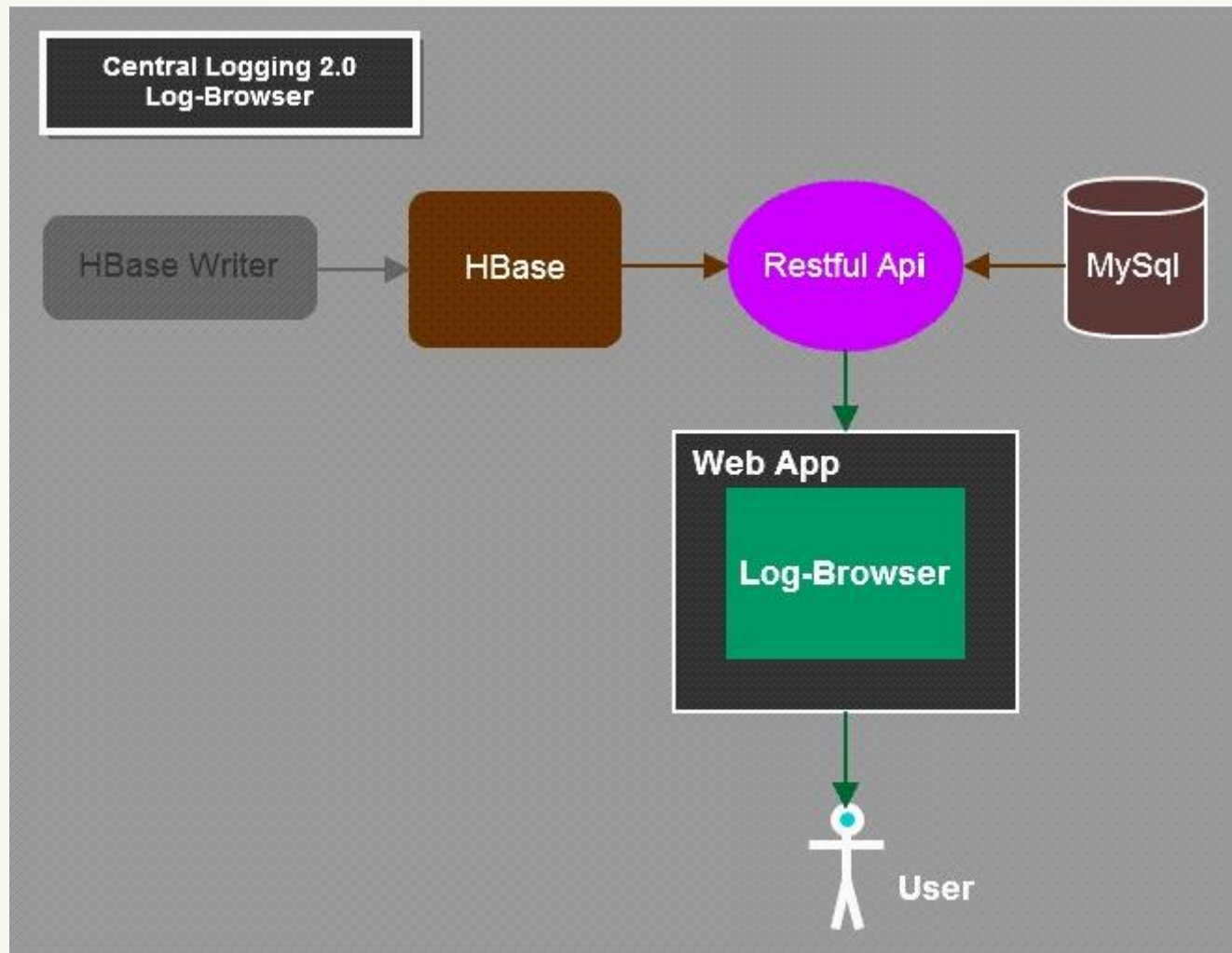
Log View - Long URL Requests Demo



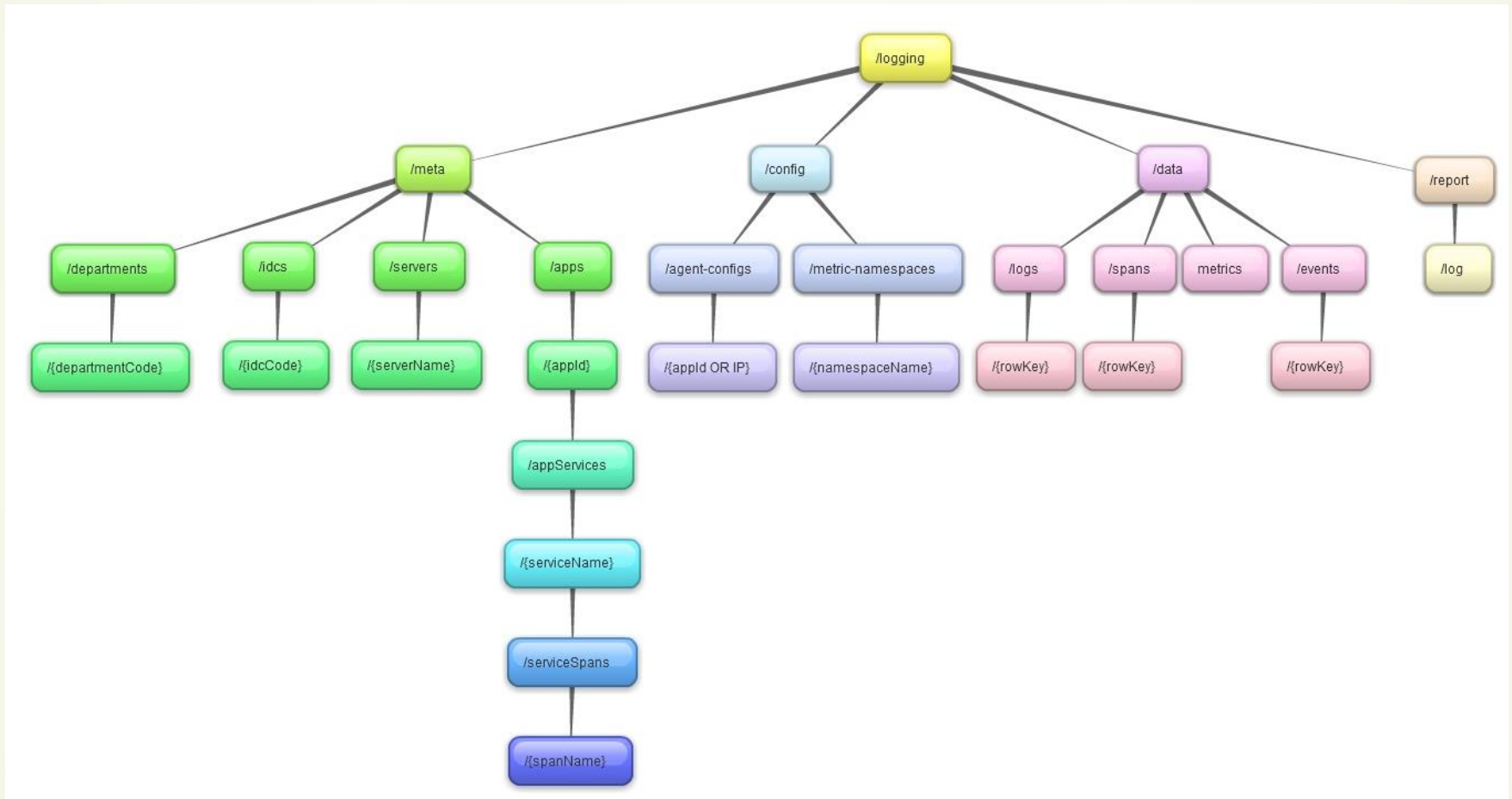
Log View - Application Log Throughput Demo



Log View - Architecture



Log View - RESTful API Design



Log View - Optimizations

- 索引表Row Key设计
- 两段式查询优化
- 基于Atomic Append的列更新、插入 (HBASE-4102)

Log View - Index Row Key Design

■ 分片前缀 (sharding prefix)

- 使得数据平衡分布

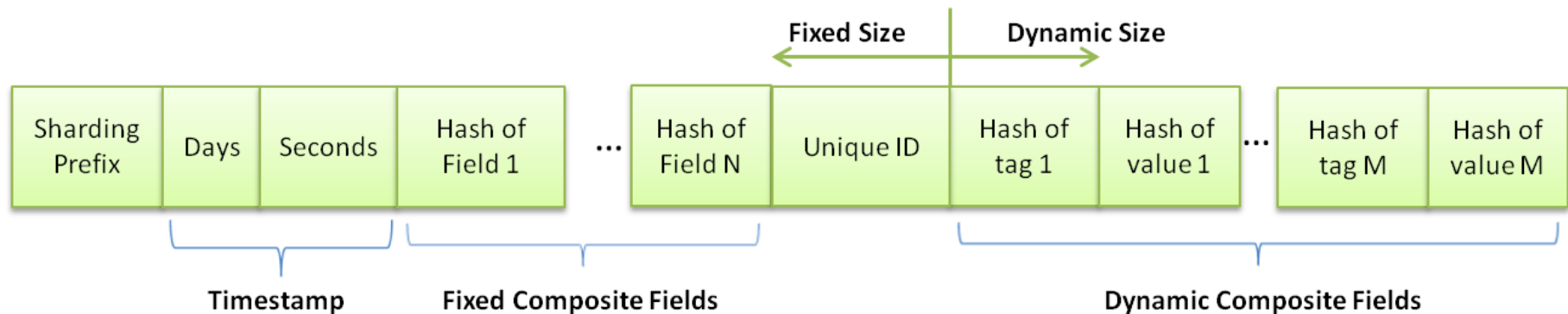
■ 时间戳 (Timestamp)

- Region重用
- 按时间顺序倒排

■ 组合索引优化

- 固定字段索引优化
- 用户自定义标签索引优化

Log View - Index Row Key Design (Cont.)



■ 分片前缀 (Sharding prefix)

- 数值类型, e.g. application id
- Bit/byte reversal

■ 时间戳

- 天: $(\text{System.currentTimeMillis()} / (86400 * 1000)) \% 30$, (TTL=30天)
- 秒: $86400 - ((\text{System.currentTimeMillis()} / 1000) \% 86400)$

■ 索引字段

- 字段按值分布进行hash压缩 (1-2字节)
- 支持静态、动态索引字段

Log View - Two-Path Query Design

■ 两段式查询优化

1. 对索引表scan查询

- 目前采用 基于Key filter非回溯正则
- 将来考虑使用FuzzyRowFilter (fast-forwarding优化)

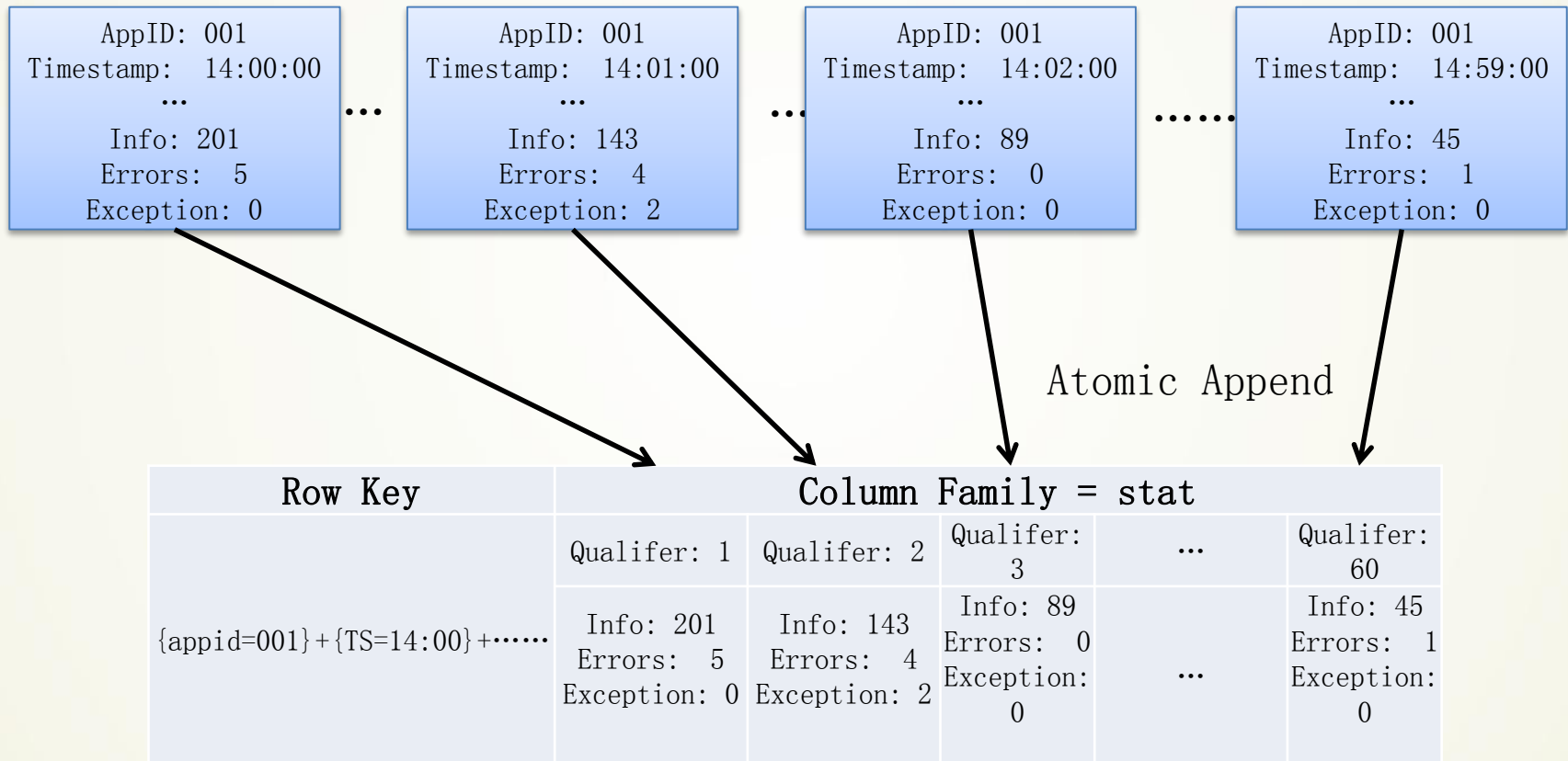
2. 对原表multiget查询 (+column filters)

■ 解决hash冲突: Skew factor

■ 性能提升 10X-100X

Log View - Atomic Append Optimization (HBASE-4102)

应用日志统计消息 - 时间线



Agenda

1. Brief Introduction
2. Central Logging
3. Log View
4. Dashboard
5. User Behavior Tracking
6. Alerting
7. Data Farm

Dashboard

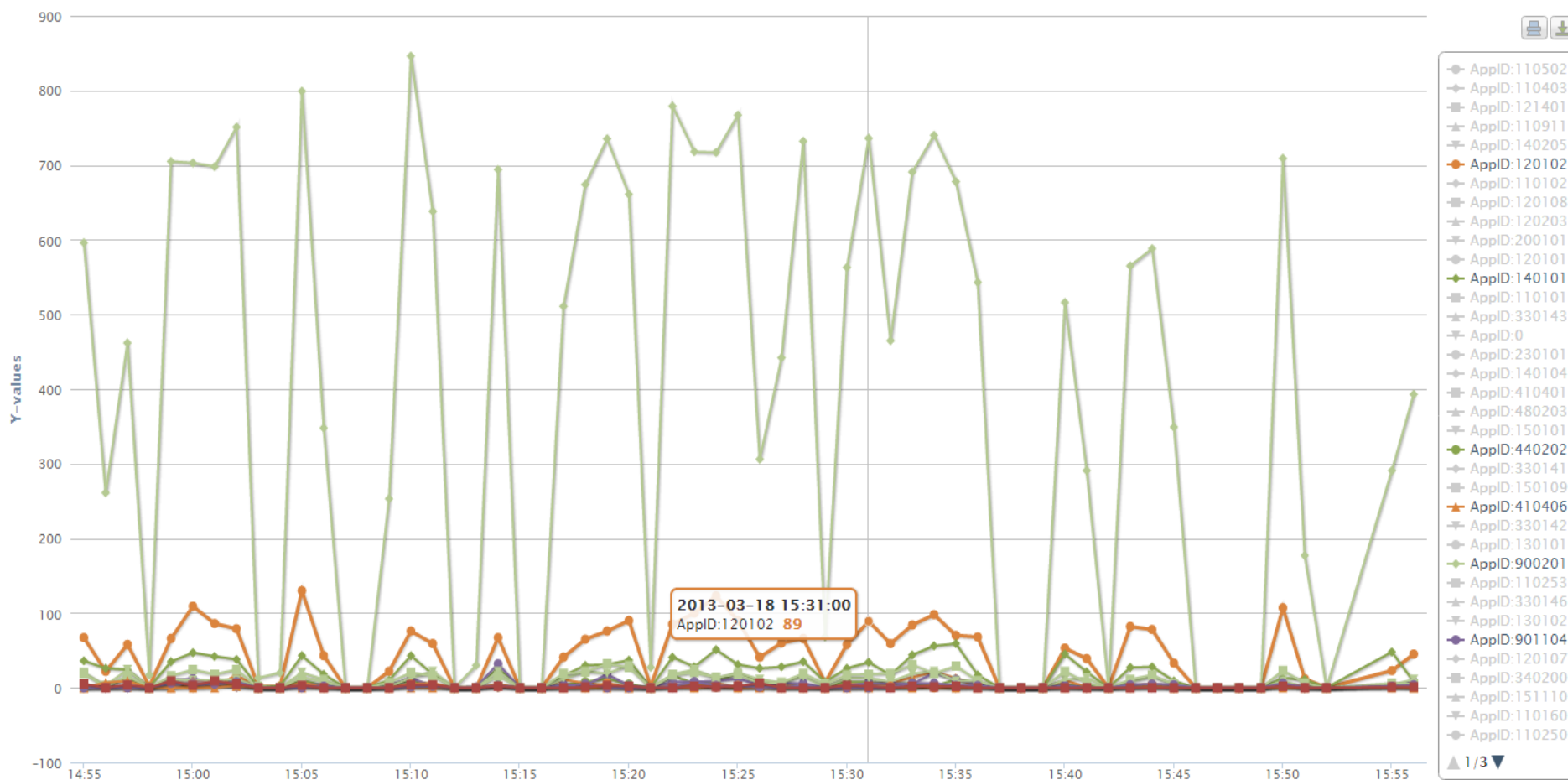
- 度量数据的展示
 - 应用、业务、系统度量数据
 - 多维度实时钻取
- 可水平扩展的度量查询引擎
- JS框架无缝整合
- 后端使用定制的OpenTSDB

Dashboard - Trace Log Demo

AppID

Dashboard

Dashboard of freeway.application.tracelog 实时

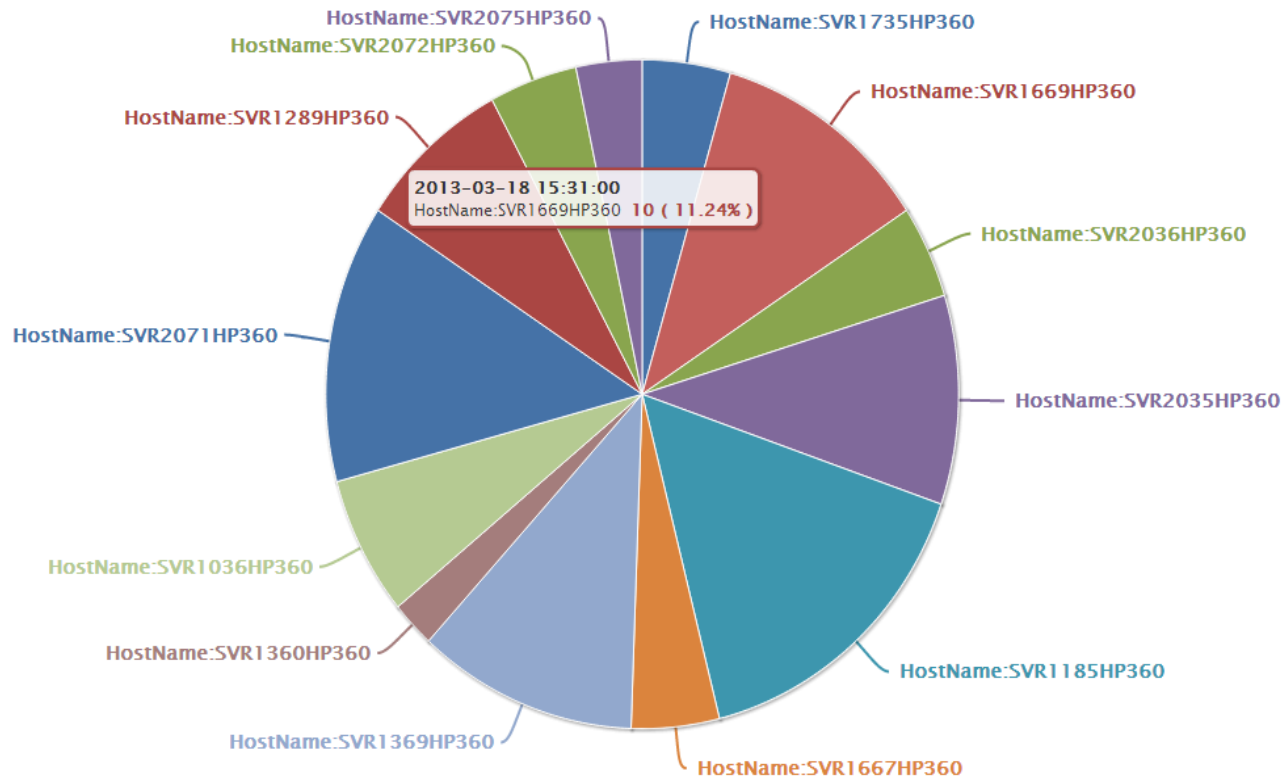


Dashboard - Trace Log Demo (Cont.)

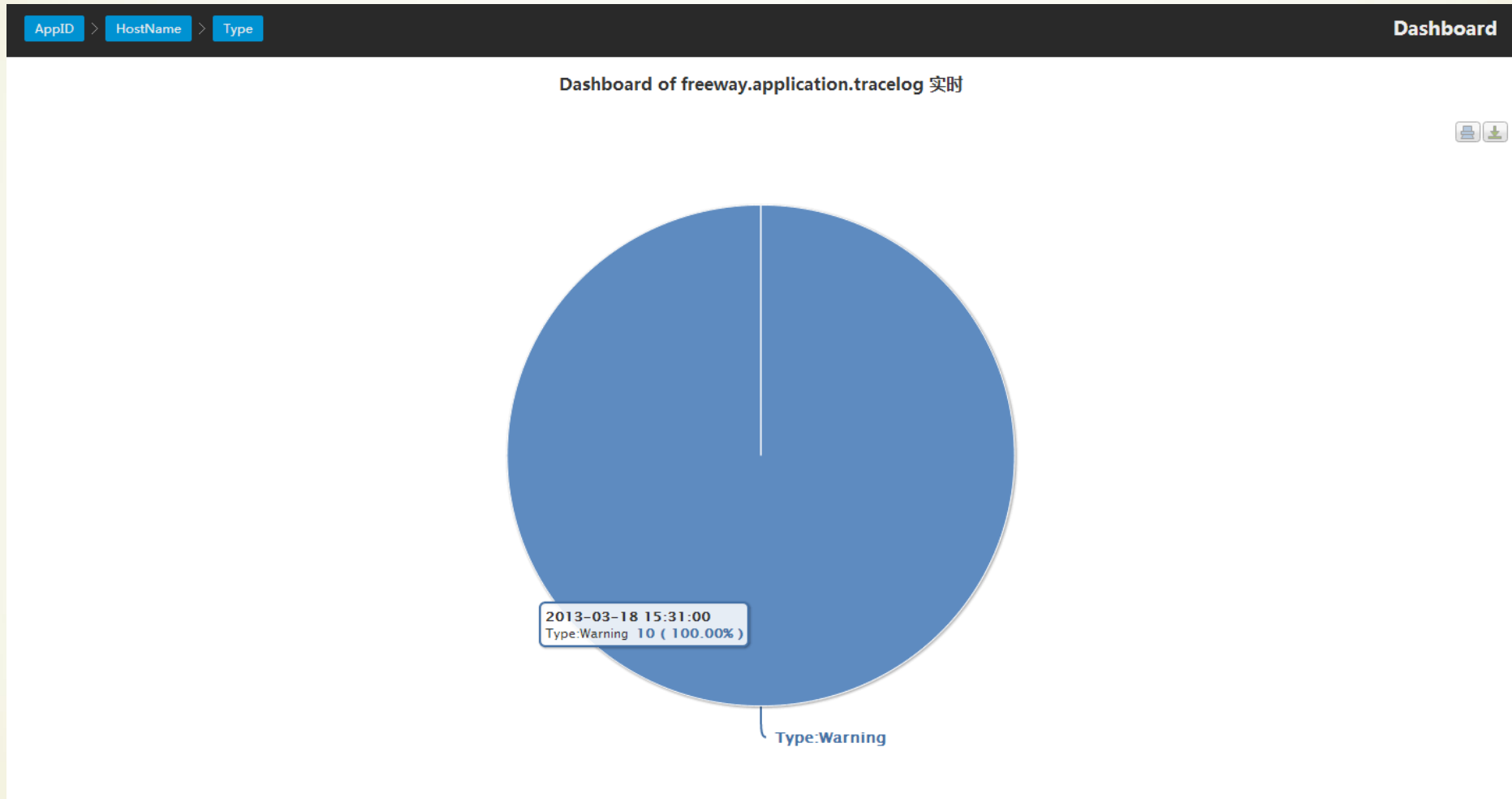
AppID > HostName

Dashboard

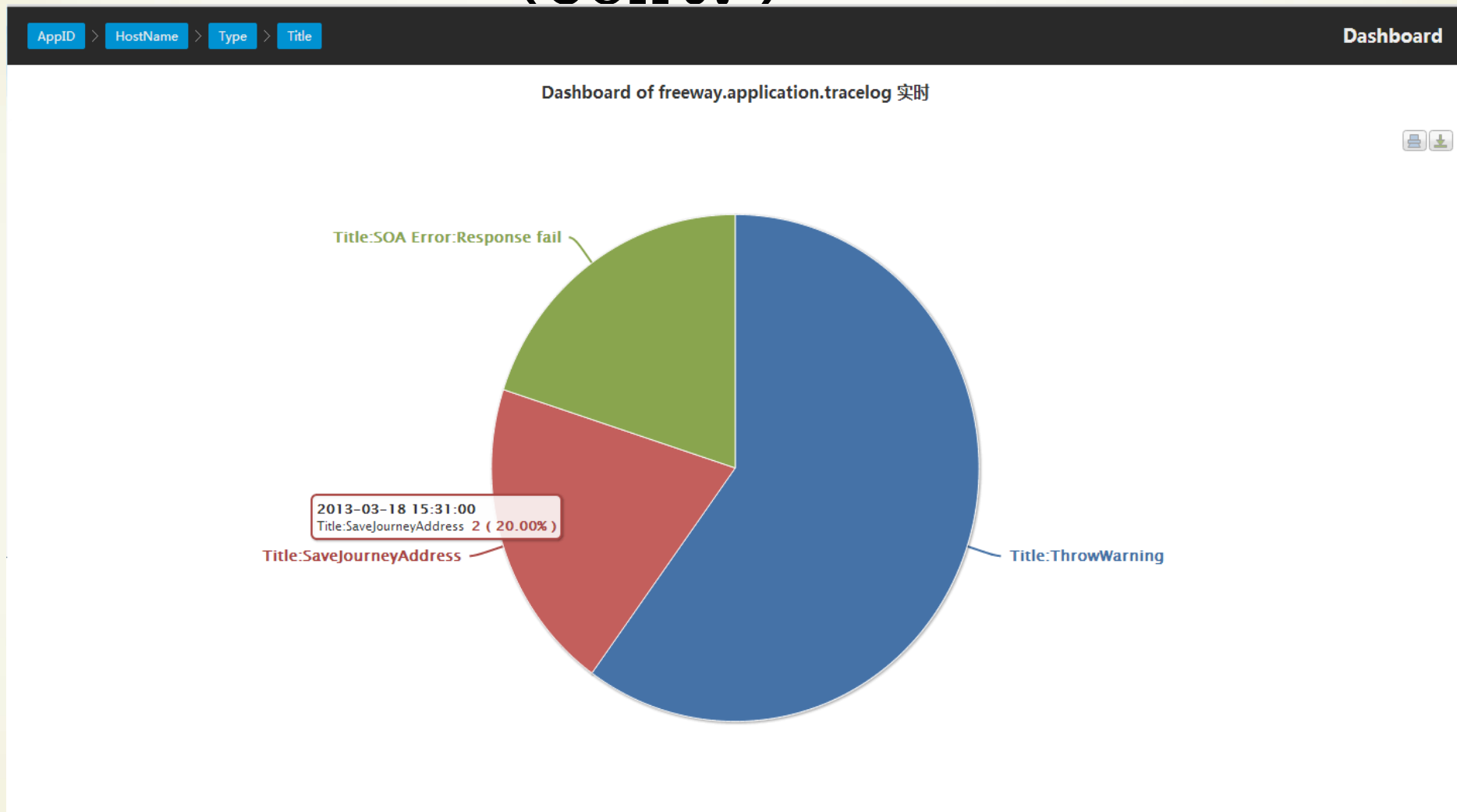
Dashboard of freeway.application.tracelog 实时




Dashboard - Trace Log Demo (Cont.)



Dashboard - Trace Log Demo (Cont.)



Dashboard – Trace Log Demo (Cont.)



Log Entries

Quickfilter

Quickfilter hit 2 messages, last searched timestamp is 18 March 2013 03:30:01.

Date	Level	Title	Message
18 March 2013 15:30:20	WARN	SaveJourneyAddress	Ctrip.Flight.Booking.Bussiness.User.UserImplement.User.SaveJourneyAddress: 配送页
18 March 2013 15:30:20	WARN	SaveJourneyAddress	Ctrip.Flight.Booking.Bussiness.User.UserImplement.User.SaveJourneyAddress: 配送页

Message 451697728

SaveJourneyAddress

From: [REDACTED]

AppId: -120102

Date: 18 March 2013 15:30:20

Severity: WARN

Type: APP

Source: Ctrip.UI.TraceService.CentralLoggingWriter

TraceId: 0

ThreadId: 11

Full message:
Ctrip.Flight.Booking.Bussiness.User.UserImplement.User.SaveJourneyAddress: 配送页

Additional info:

OrderID: 0

ErrorID: 0

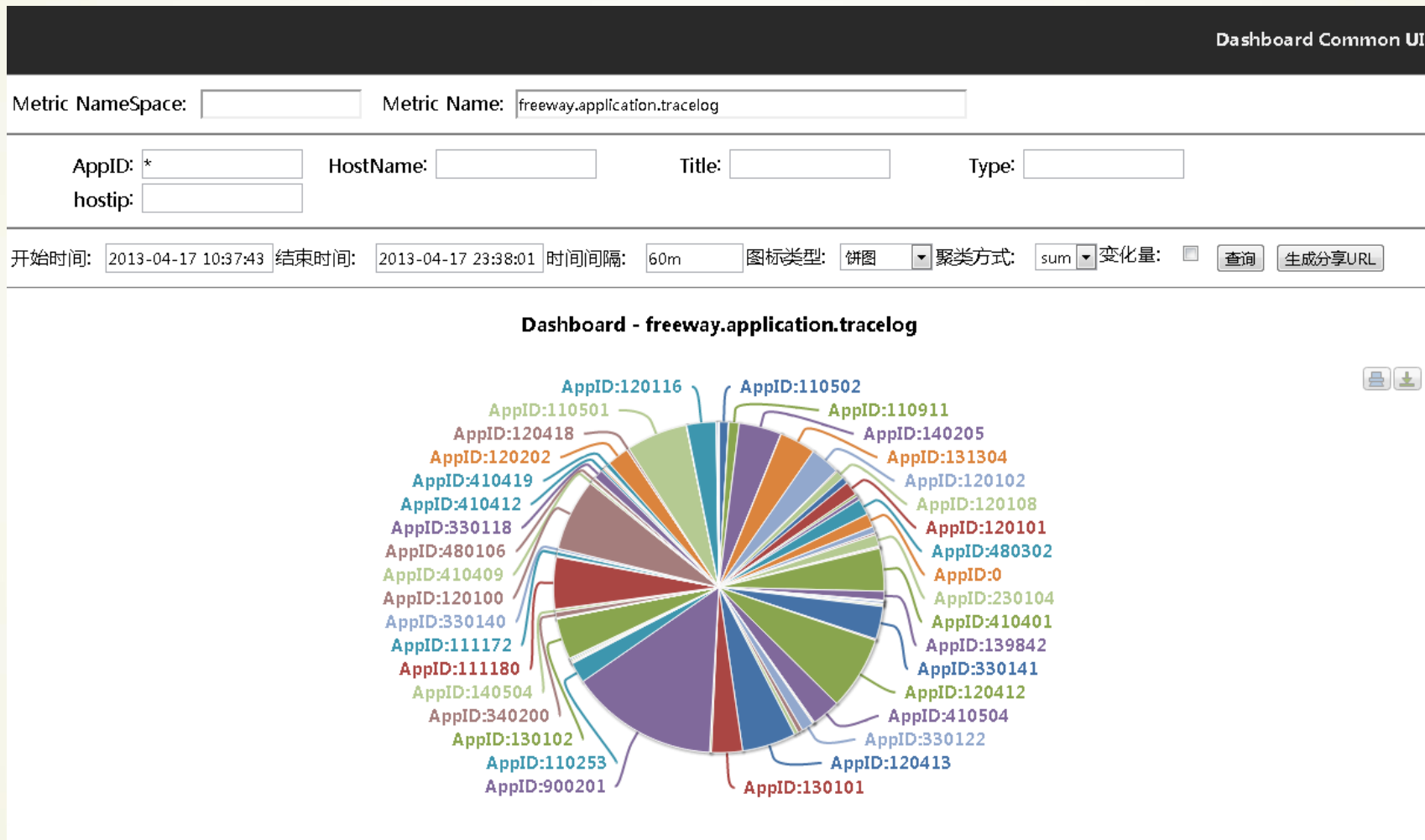
ClientIP:

ModuleID: 0

Remark: 保存用户邮寄地址接口异常

[Link to the log](#)

Dashboard - Common UI Demo



Dashboard - Dashboard JS Designer



Dashboard - Dashboard JS Designer (Cont.)

The screenshot displays the Dashboard JS Designer interface. On the left, a sidebar contains a 'Sitemap' and 'Page Notes' section, with a 'Show Links and Options' button and links to 'Dashboard Widget' and 'General Dashboard'. The main area is titled 'General Dashboard View' with the subtitle 'Design By Framework Team'. Below this, there is a '选择matrix' (Select matrix) dropdown set to 'hotel'. A '酒店查询服务器性能' (Hotel Query Server Performance) chart is visible, showing a line graph with a date range of '本周' (This Week). A '代码编辑器' (Code Editor) window is open, displaying configuration for the 'Hotel Performance Dashboard'. The configuration includes:

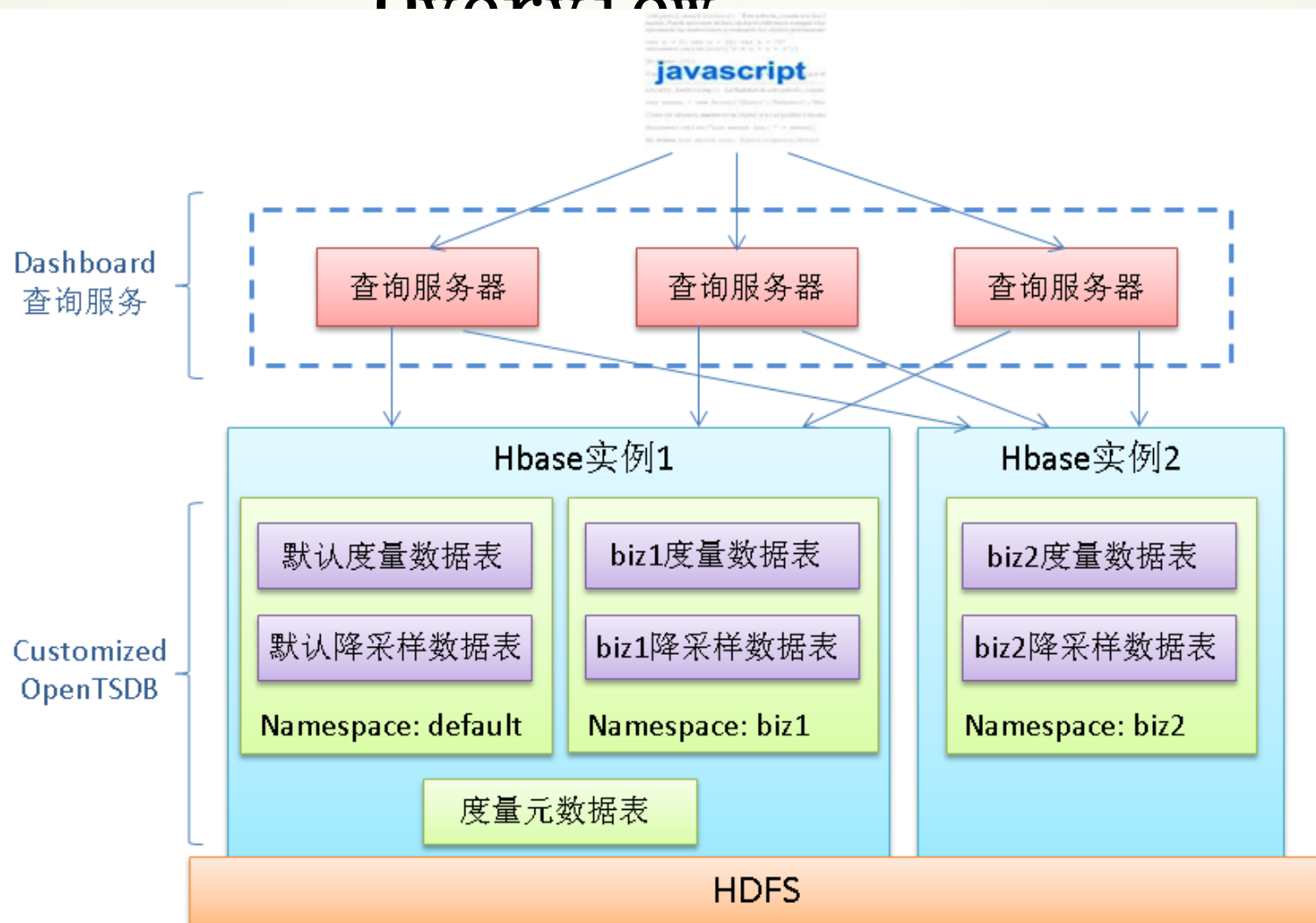
- title:** Hotel Performance Dashboard
- bases:**

```
{ '%chart%': 'stock', '%startTime%': '2012-11-01 00:00:00', '%endTime%': '2012-12-01 00:00:00', '%interval%': '1d', '%matric%': 'hotel_overview_performance', }
```
- filters:**

```
{ '@server@': { available: ['db', 'app', 'storage'], default: 'db' }, '%chart%': ['stock', 'line'], '#date': { range#: { available: [ 'last week': { '@original@': '@online' } ] } }
```
- groups:** ['@server@', '@original@']
- statistics:** 求和 sum
- theme:** default
- template:** [Empty text area]

The code editor has '预览' (Preview) and '确定' (Confirm) buttons. On the right, there are tabs for '工具' (Tools) and '代码生成' (Code Generation), and a '图标类型' (Icon Type) dropdown set to '股市图' (Stock Chart). A line chart is visible in the background, showing data from July 2010 to November 2010.

Dashboard - Architecture Overview



Dashboard - Customized OpenTSDB

- 新增basic time series (BTS) 元数据 (申请专利)
- 查询服务的cache优化 (申请专利)
- 基于不同业务类型 (namespace) 分片
- 查询服务可水平扩展
- 定时对原数据进行降采样 (map/reduce)
- 放宽原来最多8个tags的限制 (至32个)
- 支持中文encoding
- 扩展复合数值类型, 支持avg/dev降采样聚合

Dashboard - JS Framework

- 整合到携程JS框架
- 支持各类图表展示（基于highcharts）
- 所见即所得的JS设计
- 支持OpenTSDB的各类聚合方式
- 支持自定义查询条件、group by方式
- 按查询时间范围自适应调整时间间隔

Agenda

1. Brief Introduction
2. Central Logging
3. Log View
4. Dashboard
5. User Behavior Tracking
6. Alerting
7. Data Farm

User Behavior Tracker (UBT)

■ BI分析重要数据源

➤ 用户行为数据

- PV, UV, Click, JS error, page performance...

➤ 前端业务自定义数据回传

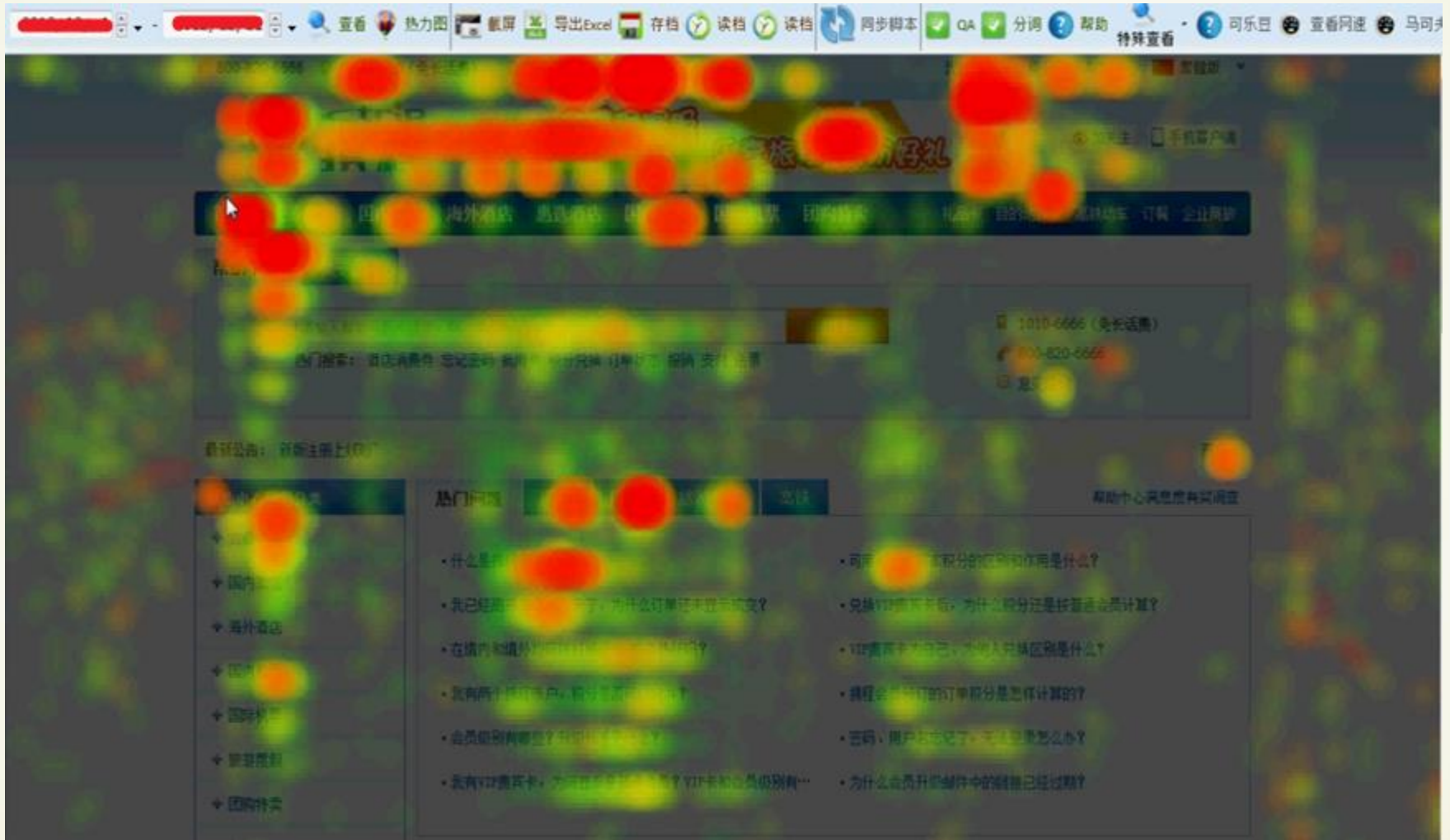
➤ AB测试数据收集

➤ 页面、流程转化率

➤ 用户体验分析

- Query understanding
- Impression/Click-Through
- 订单完成费力度

UBT - Heat Map Demo



UBT - Page UBT Visualization Demo



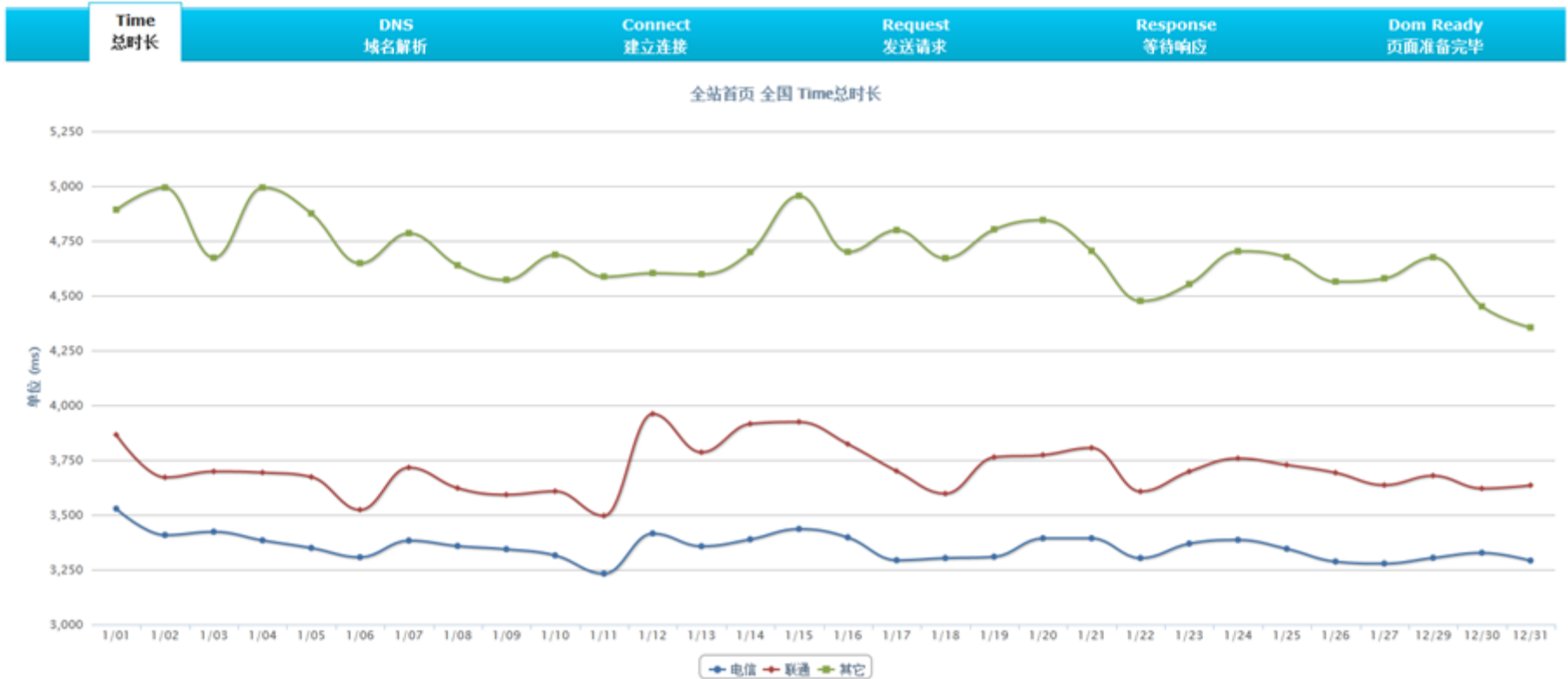
UBT - Page Speed Monitoring Demo

Page Speed Monitor beta
Big Footprints 旗下子系统

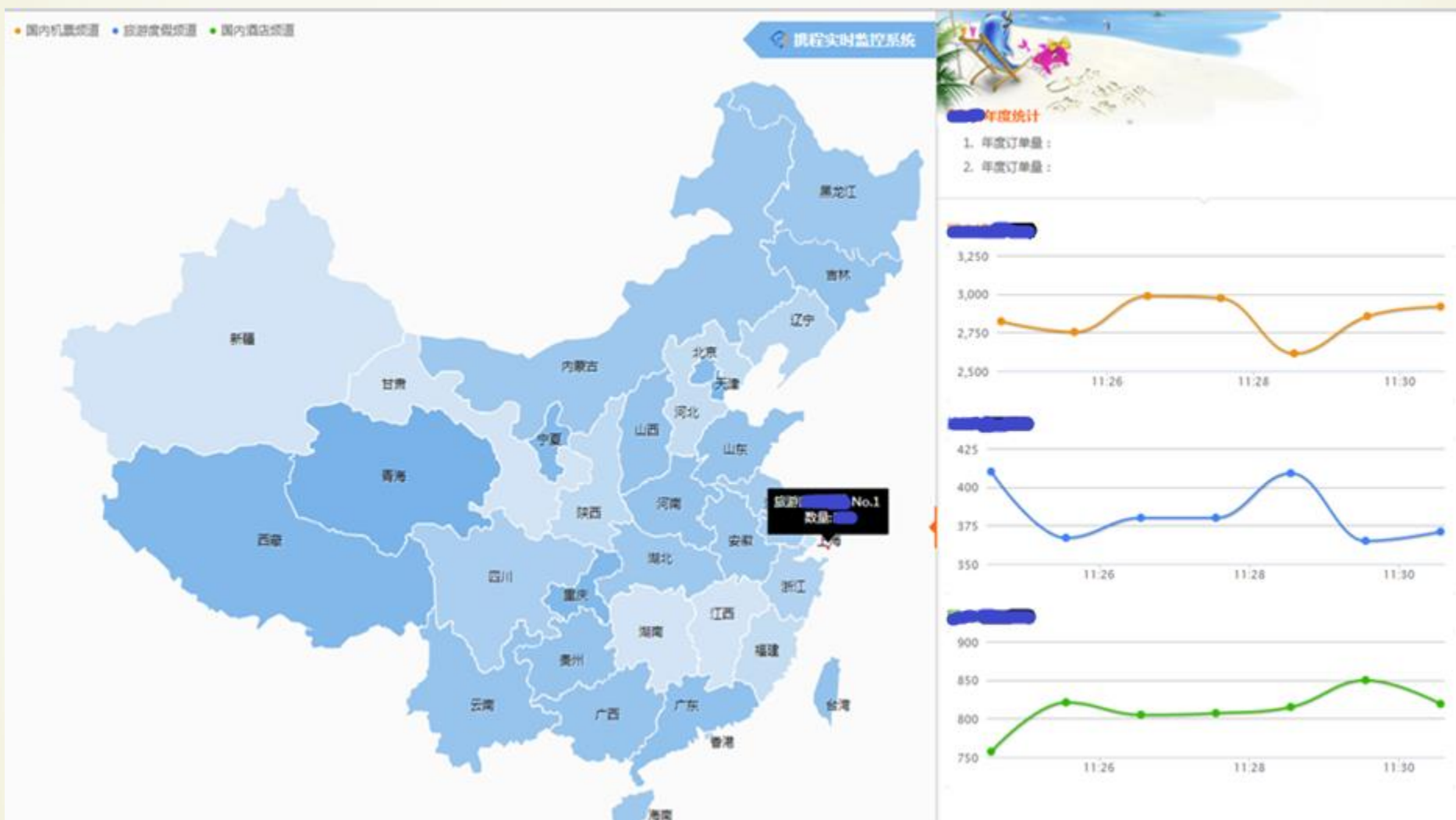
速度趋势 同一地区、不同运营商的加载速度

查看: 30日均线 | 7日均线 | 实时

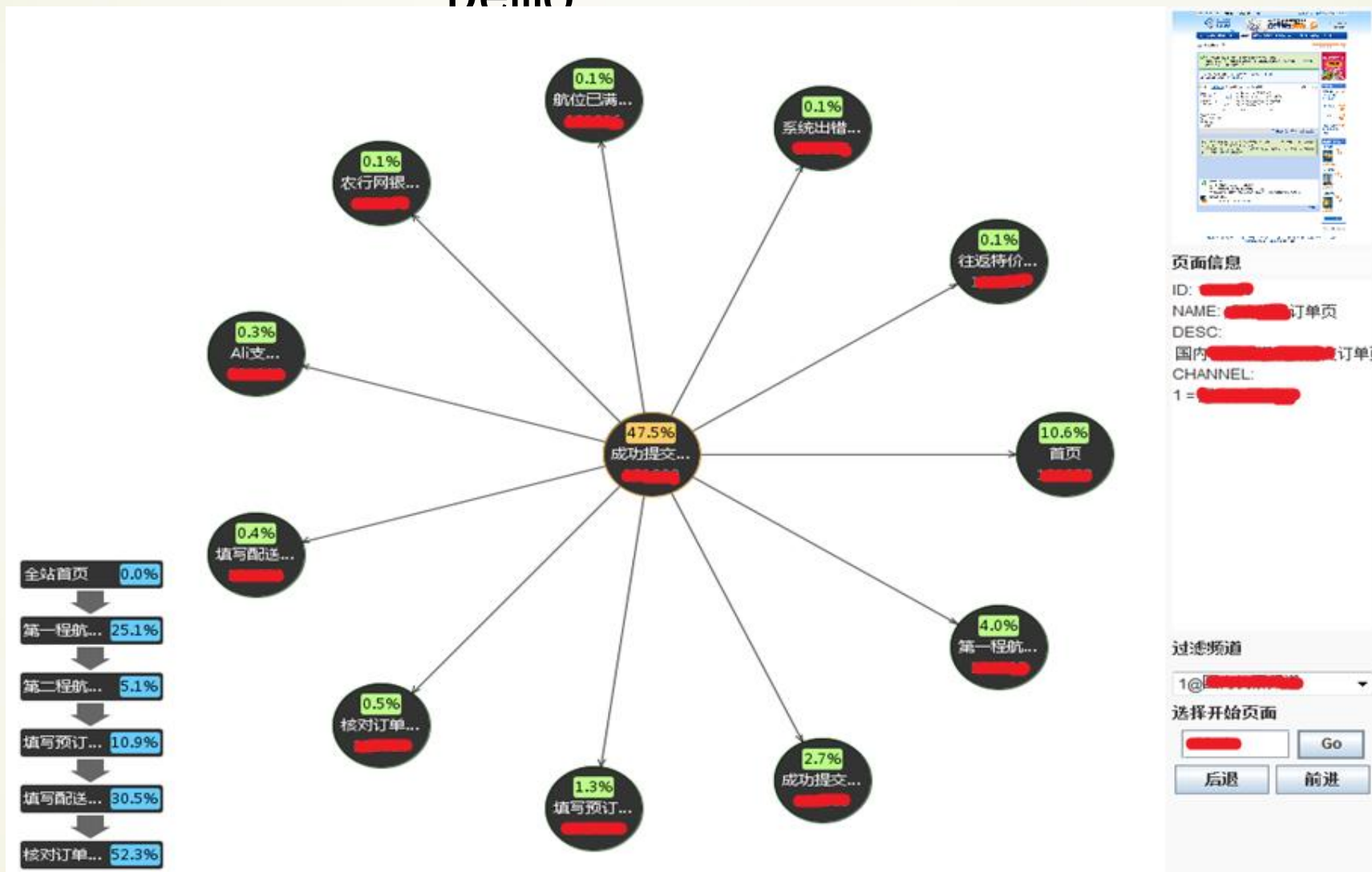
全站首页 | 国内酒店搜索结果页 | 国内机票搜索结果页 | 国际机票搜索结果页 | 度假搜索结果页



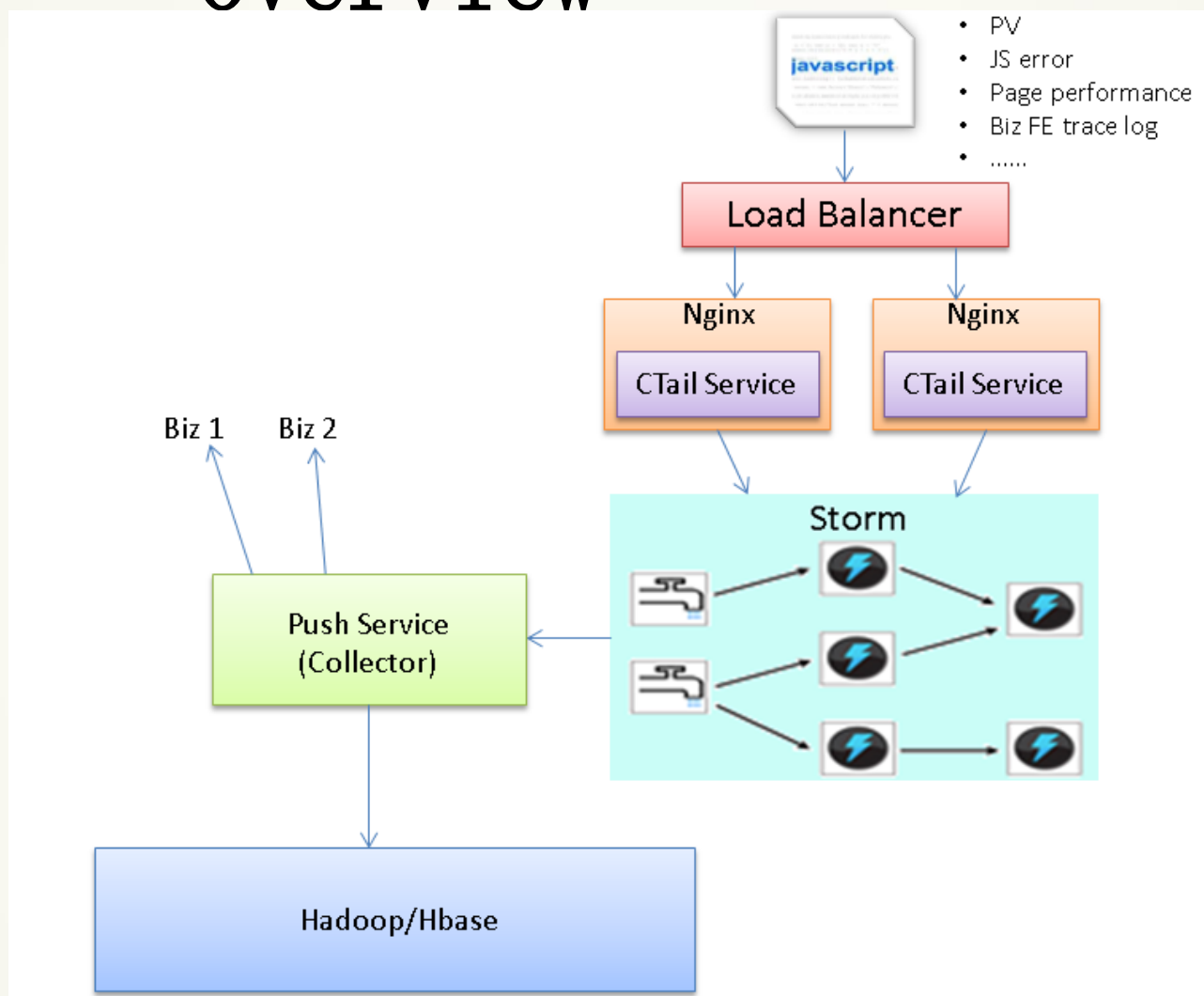
UBT - Real-Time Order Monitoring Demo



UBT - Real-Time Order Monitoring Demo



UBT - Architecture Overview



UBT - Data Collection

- 前端页面JS埋点（JS框架支持）
- Nginx接收回传数据，输出到syslog
- Ctail服务读取syslog，作为spout数据源
 - 支持游标定位
 - 支持数据重放
 - 支持日志轮转
 - 支持压缩传输

UBT - Storm Stream Processing

- 采用transaction方式(Trident Storm)
- 故障自动回复
- 实时统计, metrics写入dashboard
 - 每个session内所有感兴趣的事件
 - 外站来源统计
 - 主要页面的转换率
 - A/B测试结果
 -
- 基础UBT数据实时推送到UBT Push Service
 - 保存到Hadoop/Hbase
 - 推送给某些实时业务服务

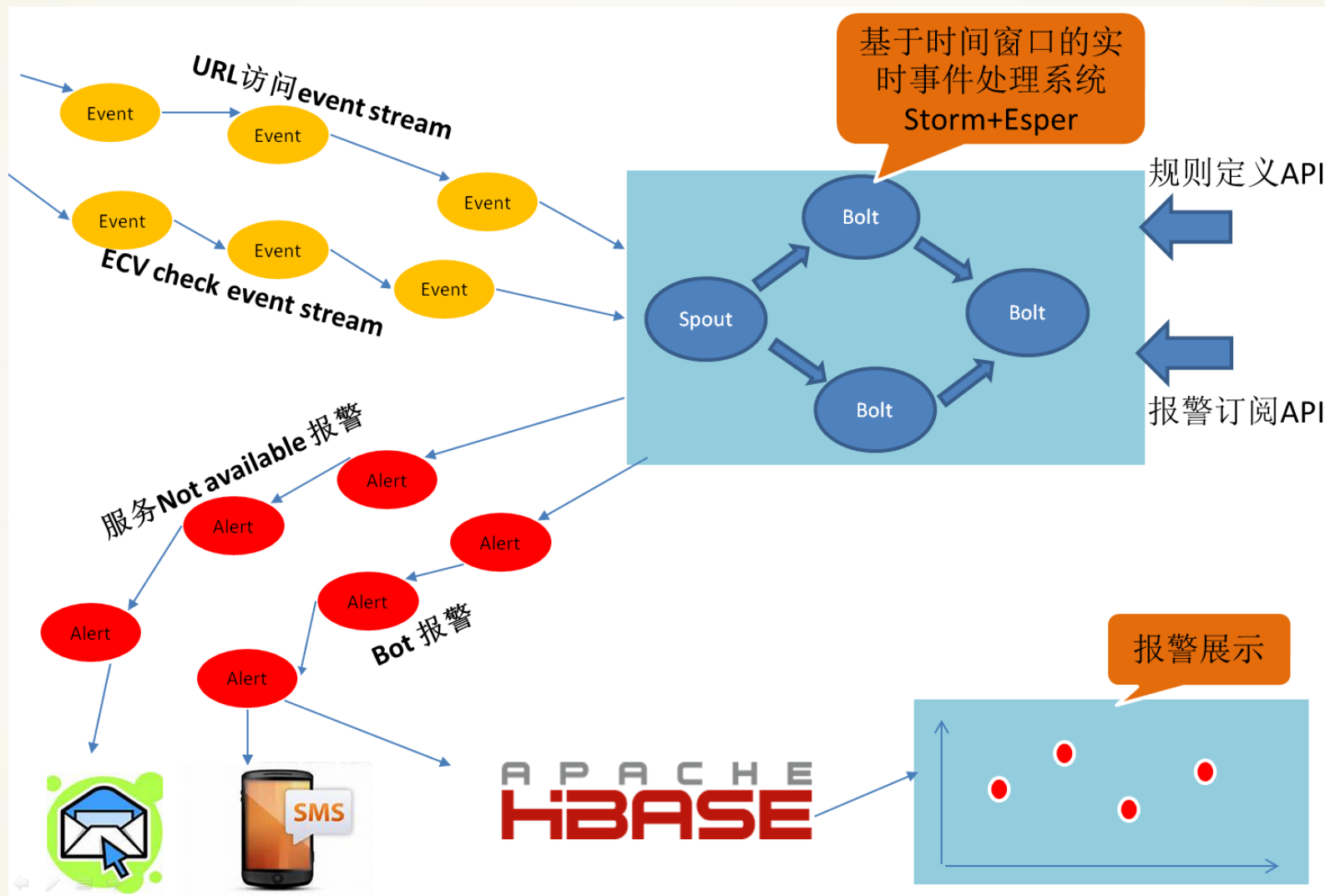
Agenda

1. Brief Introduction
2. Central Logging
3. Log View
4. Dashboard
5. User Behavior Tracking
6. Alerting
7. Data Farm

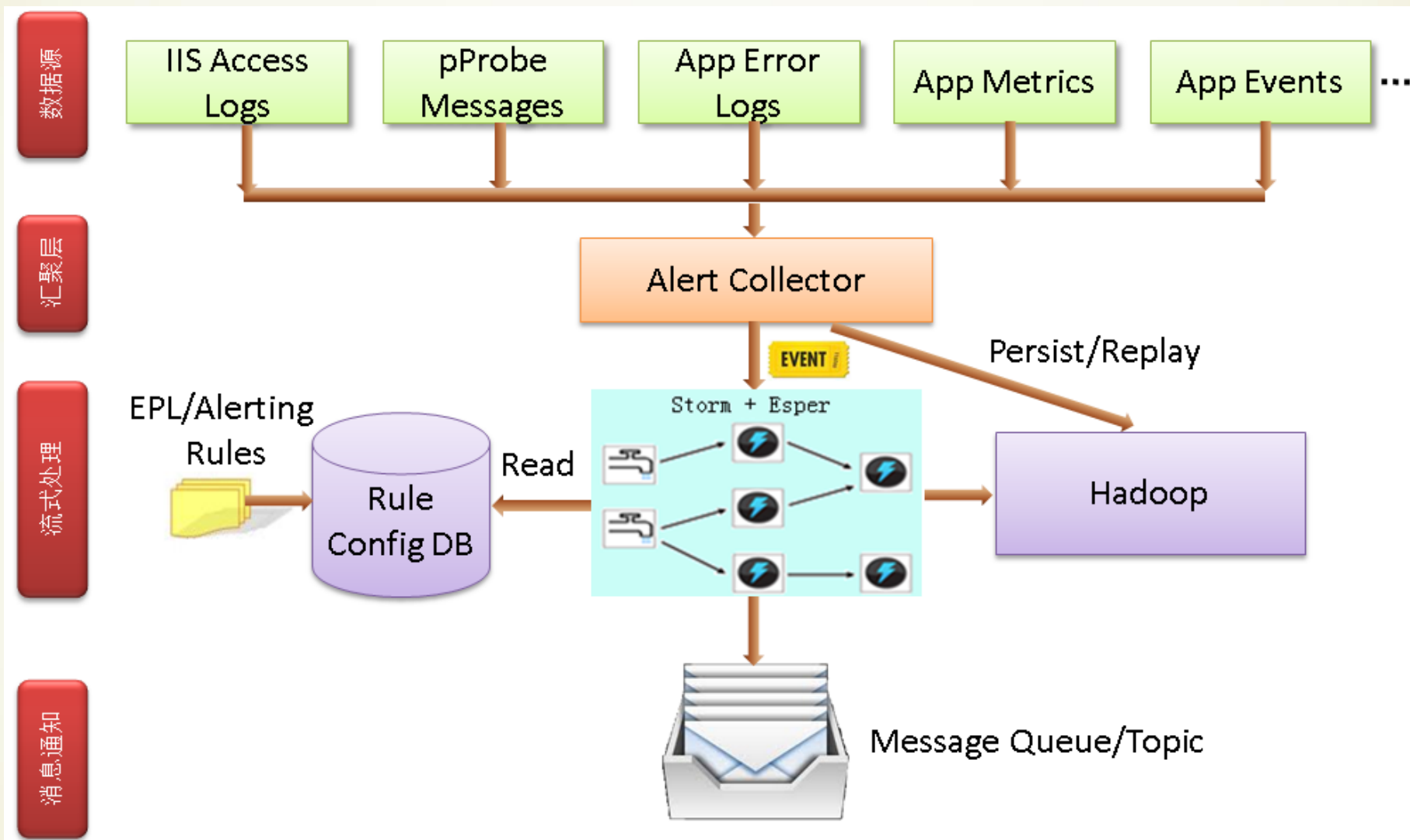
Alerting

- 业务异常报警
- 应用健康报警
- 系统异常报警
- DDoS攻击检测报警
- 恶意爬虫检测报警
-

Alerting - Framework



Alerting - Framework



Alerting - Event Processing Design

- 通用Event定义，支持namespace
- 通用Storm拓扑设计，支持分片逻辑
- 旁路Event数据至Hadoop（replay、dry run）
- EPL规则管理，规则跟新无需重启Storm
- 基于Event时间戳实时排序

Alerting - Rule & Alerts

规则

```
insert into _ipAlarm select ip, urlid, sum(count) as count, min(time) as starttime, max(time) as endtime,
"severe" as level, "ip" as type from Ips.win:time(10000) group by ip, urlid having sum(count) >= 10000
insert into _ipAlarm select ip, urlid, sum(count) as count, min(time) as starttime, max(time) as endtime,
"medium" as level, "ip" as type from Ips.win:time(10000) group by ip, urlid having sum(count) >= 10000 and
sum(count) < 10000
insert into _ipAlarm select ip, urlid, sum(count) as count, min(time) as starttime, max(time) as endtime,
"small" as level, "ip" as type from Ips.win:time(10000) group by ip, urlid having sum(count) >= 10000 and
sum(count) < 10000
```

报警

ip	level	count	url_count	type	urlid	starttime	endtime	urlcategory
1	severe	294	147	ip_category	0	0320-17:42:00	0320-18:11:00	/international
	severe	286	5	ip_category	0	0320-17:27:00	0320-17:50:30	/international
	severe	253	127	ip_category	0	0320-17:12:00	0320-17:41:00	/international
	severe	239	22	ip_category	0	0320-17:28:00	0320-17:56:30	/international
	severe	226	8	ip_category	0	0320-17:19:00	0320-17:35:30	/international
	severe	221	17	ip_category	0	0320-17:31:00	0320-17:56:00	/international
	severe	215	18	ip_category	0	0320-17:24:30	0320-17:52:00	/international
	severe	209	5	ip_category	0	0320-17:42:30	0320-18:11:00	/international
	severe	197	13	ip_category	0	0320-17:29:00	0320-17:58:00	/international
	severe	187	13	ip_category	0	0320-17:27:30	0320-17:54:00	/international
	severe	178	20	ip_category	0	0320-17:14:00	0320-17:42:00	/international
	severe	176	24	ip_category	0	0320-17:39:30	0320-18:08:30	/international
	severe	174	11	ip_category	0	0320-17:27:30	0320-17:56:30	/international
	severe	174	16	ip_category	0	0320-17:37:30	0320-17:56:30	/international
	severe	168	25	ip_category	0	0320-17:30:00	0320-17:54:00	/international
	severe	166	83	ip_category	0	0320-17:16:00	0320-17:45:00	/international
	severe	164	5	ip_category	0	0320-17:12:00	0320-17:41:00	/international
1	severe	150	1	ip_category	0	0320-17:15:30	0320-17:44:00	/international
1	severe	148	14	ip_category	0	0320-17:35:00	0320-17:56:00	/international

Agenda

1. Brief Introduction
2. Central Logging
3. Log View
4. Dashboard
5. User Behavior Tracking
6. Alerting
7. Data Farm

Data Farm - Building

■ 数据仓库

■ 各类数据整合

- 应用日志数据
- 业务数据
- 用户行为数据
- 系统日志数据
- 各类度量数据
-

■ 数据挖掘、分析

Data Farm - Planning

- 各类数据的ETL
- 全面启动ACL，保护数据安全性
- 基于任务依赖调度（Oozie?）
- 执行引擎选择（Hive, Impala, Shark, Phoenix?）
- 优化数据处理



Thanks Q&A

欢迎莅临

2013中国数据库技术大会

Database
BDaaS
flowingdata
DB2
NoSQL MySQL
Oracle Big Data